Cochrane Library

Cochrane Database of Systematic Reviews

Exercise interventions on health-related quality of life for cancer survivors (Review)

Mishra SI, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, Snyder C

Mishra SI, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, Snyder C.
Exercise interventions on health-related quality of life for cancer survivors.
Cochrane Database of Systematic Reviews 2012, Issue 8. Art. No.: CD007566.
DOI: 10.1002/14651858.CD007566.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Exercise interventions on health-related quality of life for cancer survivors

Shiraz I Mishra[1], Roberta W Scherer[2], Paula M Geigle[3], Debra R Berlanstein[4], Ozlem Topaloglu[5], Carolyn C Gotay[6], Claire Snyder[7]

[1]University of New Mexico, Albuquerque, NM, USA. [2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. [3]Rehabilitation Research Center; Department of Neurology, University of Maryland at Kernan Rehabilitation Hospital, Baltimore, MD, USA. [4]Hirsh Health Sciences Library, Tufts University, Boston, MA, USA. [5]Department of Epidemiology, JHU Bloomberg School of Public Health, Baltimore, MD, USA. [6]School of Population and Public Health, Faculty of Medicine, University of British Columbia, Vancouver, Canada. [7]Division of General Internal Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA

Contact address: Shiraz I Mishra, smishra@salud.unm.edu.

Editorial group: Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group.
Publication status and date: Edited (no change to conclusions), published in Issue 5, 2020.

Citation: Mishra SI, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, Snyder C. Exercise interventions on health-related quality of life for cancer survivors. *Cochrane Database of Systematic Reviews* 2012, Issue 8. Art. No.: CD007566. DOI: 10.1002/14651858.CD007566.pub2.

ABSTRACT

Background

Cancer survivors experience numerous disease and treatment-related adverse outcomes and poorer health-related quality of life (HRQoL). Exercise interventions are hypothesized to alleviate these adverse outcomes. HRQoL and its domains are important measures for cancer survivorship.

Objectives

To evaluate the effectiveness of exercise on overall HRQoL and HRQoL domains among adult post-treatment cancer survivors.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, MEDLINE, EMBASE, CINAHL, PsycINFO, PEDRO, LILACS, SIGLE, SportDiscus, OTSeeker, and Sociological Abstracts from inception to October 2011 with no language or date restrictions. We also searched citations through Web of Science and Scopus, PubMed's related article feature, and several websites. We reviewed reference lists of included trials and other reviews in the field.

Selection criteria

We included all randomized controlled trials (RCTs) and controlled clinical trials (CCTs) comparing exercise interventions with usual care or other nonexercise intervention to assess overall HRQoL or at least one HRQoL domain in adults. Included trials tested exercise interventions that were initiated after completion of active cancer treatment. We excluded trials including people who were terminally ill, or receiving hospice care, or both, and where the majority of trial participants were undergoing active treatment for either the primary or recurrent cancer.

Data collection and analysis

Five paired review authors independently extracted information on characteristics of included trials, data on effects of the intervention, and assessed risk of bias based on predefined criteria. Where possible, meta-analyses results were performed for HRQoL and HRQoL domains for the reported difference between baseline values and follow-up values using standardized mean differences (SMD) and a random-effects model by length of follow-up. We also reported the SMDs between mean follow-up values of exercise and control group. Because investigators used many different HRQoL and HRQoL domain instruments and often more than one for the same domain, we selected the

more commonly used instrument to include in the SMD meta-analyses. We also report the mean difference for each type of instrument separately.

Main results

We included 40 trials with 3694 participants randomized to an exercise (n = 1927) or comparison (n = 1764) group. Cancer diagnoses in study participants included breast, colorectal, head and neck, lymphoma, and other. Thirty trials were conducted among participants who had completed active treatment for their primary or recurrent cancer and 10 trials included participants both during and post cancer treatment. Mode of the exercise intervention included strength training, resistance training, walking, cycling, yoga, Qigong, or Tai Chi. HRQoL and its domains were measured using a wide range of measures.

The results suggested that exercise compared with control has a positive impact on HRQoL and certain HRQoL domains. Exercise resulted in improvement in: global HRQoL at 12 weeks' (SMD 0.48; 95% confidence interval (CI) 0.16 to 0.81) and 6 months' (0.46; 95% CI 0.09 to 0.84) follow-up, breast cancer concerns between 12 weeks' and 6 months' follow-up (SMD 0.99; 95% CI 0.41 to 1.57), body image/self-esteem when assessed using the Rosenberg Self-Esteem scale at 12 weeks (MD 4.50; 95% CI 3.40 to 5.60) and between 12 weeks' and 6 months' (mean difference (MD) 2.70; 95% CI 0.73 to 4.67) follow-up, emotional well-being at 12 weeks' follow-up (SMD 0.33; 95% CI 0.05 to 0.61), sexuality at 6 months' follow-up (SMD 0.40; 95% CI 0.11 to 0.68), sleep disturbance when comparing follow-up values by comparison group at 12 weeks' follow-up (SMD -0.46; 95% CI -0.72 to -0.20), and social functioning at 12 weeks' (SMD 0.45; 95% CI 0.02 to 0.87) and 6 months' (SMD 0.49; 95% CI 0.11 to 0.87) follow-up.

Further, exercise interventions resulted in decreased anxiety at 12 weeks' follow-up (SMD -0.26; 95% CI -0.07 to -0.44), fatigue at 12 weeks' (SMD -0.82; 95% CI -1.50 to -0.14) and between 12 weeks' and 6 months' (SMD -0.42; 95% CI -0.02 to -0.83) follow-up, and pain at 12 weeks' follow-up (SMD -0.29; 95% CI -0.55 to -0.04) when comparing follow-up values by comparison group.

Positive trends and impact of exercise intervention existed for depression and body image (when analyzing combined instruments); however, because few studies measured these outcomes the robustness of findings is uncertain.

No conclusions can be drawn regarding the effects of exercise interventions on HRQoL domains of cognitive function, physical functioning, general health perspective, role function, and spirituality.

Results of the review need to be interpreted cautiously owing to the risk of bias. All the trials reviewed were at high risk for performance bias. In addition, the majority of trials were at high risk for detection, attrition, and selection bias.

Authors' conclusions

This systematic review indicates that exercise may have beneficial effects on HRQoL and certain HRQoL domains including cancer-specific concerns (e.g. breast cancer), body image/self-esteem, emotional well-being, sexuality, sleep disturbance, social functioning, anxiety, fatigue, and pain at varying follow-up periods. The positive results must be interpreted cautiously due to the heterogeneity of exercise programs tested and measures used to assess HRQoL and HRQoL domains, and the risk of bias in many trials. Further research is required to investigate how to sustain positive effects of exercise over time and to determine essential attributes of exercise (mode, intensity, frequency, duration, timing) by cancer type and cancer treatment for optimal effects on HRQoL and its domains.

PLAIN LANGUAGE SUMMARY

Can exercise interventions enhance health-related quality of life among cancer survivors?

Cancer survivors often have many psychological and physical adverse events as a result of the cancer and treatment for it. They also suffer from poorer quality of life (QoL) than people without cancer. Some studies have suggested that exercise may be helpful in reducing negative outcomes and improving the QoL of people who have finished cancer treatment. Also, a better QoL may predict longer life. This review looked at the effect of exercise on QoL and areas of life that make up QoL (e.g. tiredness, anxiety, emotional health) among people who had finished all cancer treatment.

The review included 40 trials with a total of 3694 people. The results suggest that exercise may improve overall QoL right after the exercise program is completed. Exercise may also reduce the person's worry about his or her cancer, and affect the way the person views his or her body. Exercise may also help the way the person deals with emotions, sexuality, sleep problems, or functions in society. Exercise also reduced anxiety, tiredness, and pain at different times during and after the exercise program. No effect of exercise was found on the person's ability to think clearly or his or her role function in society. Also, no effect of exercise was found on the way the person views his or her spiritual or physical health, or physical ability.

However, these findings need to be viewed with caution because this review looked at many different kinds of exercise programs, which varied by type of exercise, length of the program, and how hard the trial participants had to exercise. Also, the investigators used a number of different ways to measure QoL.

More research is needed to see how to maintain the effects of exercise over a longer period of time after the exercise program is completed, and to determine which parts of the exercise program are necessary (i.e. when to start the program, type of exercise, length of program

or exercise session, how hard to exercise). It is also important to find out if one type of exercise is better for a specific cancer type than another for the maximum effect on QoL.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings

Exercise intervention compared with usual care on HRQoL and HRQoL domains for cancer survivors

Patient or population: Cancer survivors who have completed active cancer treatment

Settings: Varied

Intervention: Exercise interventions (varied)

Comparison: Usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Comparison group	Exercise intervention group				
Overall quality of life change score - up to 12 weeks' follow-up	The standardized mean change from baseline to up to 12 weeks' follow-up in overall quality of life ranged across control groups from -0.59 to 0.56 standard deviation units	The standardized mean change from baseline to up to 12 weeks' follow-up in overall quality of life in the exercise groups was 0.48 standard deviation units higher (0.16 to 0.81 standard deviation units higher)		826 (11 studies)	⊕⊕⊝⊝ **low[1,2]**	(SMD 0.48; 95% CI 0.16 to 0.81) A standard deviation units is equivalent to about a 14.5-point change using the FACT-G HRQoL form or a 18.5-point change using the QLQ-C30 HRQoL form
Overall quality of life change score - 6 months' follow-up	The standardized mean change from baseline to 6 months' follow-up in overall quality of life ranged across control groups from -0.32 to 0.15 standard deviation units	The standardized mean change from baseline to 6 months' follow-up in overall quality of life in the exercise groups was 0.46 standard deviation units higher (0.09 to 0.84 standard deviation units higher)		115 (2 studies)	⊕⊕⊕⊝ **moderate[1,3]**	(SMD 0.46; 95% CI 0.09 to 0.84)
Overall anxiety change - Up to 12 weeks' follow-up	The standardized mean change from baseline to up to 12 weeks' follow-up in overall anxiety ranged across control groups from -0.25 to 0.04 standard deviation units	The standardized mean change from baseline to up to 12 weeks' follow-up in overall anxiety in the exercise groups was -0.26 standard deviation units lower (-0.44 to -0.07 standard deviation units lower)		455 (4 studies)	⊕⊕⊝⊝ **low[1,4,5]**	(SMD -0.26; 95% CI -0.44 to -0.07) A standard deviation unit is equivalent to about a 3.4-point change using the HADS scale or about a 11.5-point change using the STAI scale

Overall emotional well-being/mental health change - up to 12 weeks' follow-up	The standardized mean change from baseline to up to 12 weeks' follow-up in overall emotional well-being/mental health ranged across control groups from -0.48 to 0.46 standard deviation units	The standardized mean change from baseline to up to 12 weeks' follow-up in overall emotional well-being/mental health in the exercise groups was 0.33 standard deviation units higher (0.05 to 0.61 standard deviation units higher)	632 (8 studies)	⊕⊕⊖⊖ **low[1,2,4,5]**	(SMD 0.33; 95% CI 0.05 to 0.61) A standard deviation unit is equivalent to about a 24-point change on the QLQ-C30 emotional function sub-scale or about a 5-point change on the FACT-emotion sub-scale
Overall fatigue change - Up to 12 weeks' follow-up	The standardized mean change from baseline to up to 12 weeks' follow-up in overall fatigue ranged across control groups from -0.29 to 0.44 standard deviation units	The standardized mean change from baseline to up to 12 weeks' follow-up in overall fatigue in the exercise groups was -0.82 standard deviation units lower (-1.50 to -0.14 standard deviation units lower)	745 (10 studies)	⊕⊕⊕⊖ **moderate[1]**	(SMD -0.82; 95% CI -1.50 to -0.14) A standard deviation unit is equivalent to about a 21-point change on the QLQ-C30 fatigue subscale or about a 11-point change on the FACT-F sub-scale
Overall fatigue change - More than 12 weeks less than 6 months' follow-up	The standardized mean change from baseline to between 12 weeks and 6 month follow-up in overall fatigue ranged across control groups from -0.27 to 0.74 standard deviation units	The standardized mean change from baseline to between 12 weeks and 6 month follow-up in overall fatigue in the exercise groups was -0.42 standard deviation units lower (-0.83 to -0.02 standard deviation units lower)	246 (3 studies)	⊕⊕⊖⊖ **low[1,2,3]**	(SMD -0.42; 95% CI -0.83 to -0.02)
Overall pain follow-up values - up to 12 weeks' follow-up	The standardized mean follow-up values in overall pain for up to 12 weeks' follow-up ranged across control groups from 0.94 to 9.67 standard deviation units	The standardized mean follow-up values in overall pain for up to 12 weeks' follow-up in the exercise groups was -0.29 standard deviation units lower (-0.55 to -0.04 standard deviation units lower)	289 (4 studies)	⊕⊕⊕⊖ **moderate[1,3]**	(SMD -0.29; 95% CI -0.55 to -0.04) A standard deviation unit is equivalent to about a 28-point change on the QLQ-C30 pain sub-scale
Overall sexuality change - 6 months' follow-up	The standardized mean change from baseline to 6 months' follow-up in overall sexuality change ranged across control groups from -0.01 to 0.04 standard deviation units	The standardized mean change from baseline to 6 months' follow-up in overall sexuality change in the exercise groups was 0.40 standard deviation units higher (0.11 to 0.68 standard deviation units higher)	193 (2 studies)	⊕⊕⊕⊖ **moderate[1,3]**	(SMD 0.40; 95% CI 0.11 to 0.68)
Overall sleep dis-	The standardized mean change from baseline to up to 12 weeks' follow-up in overall sleep	The standardized mean change from baseline to up to 12 weeks' follow-up	438 (8 studies)	⊕⊕⊕⊖	(SMD -0.46; 95% CI -0.72 to -0.20)

turbance follow-up values - up to 12 weeks' follow-up	disturbance ranged across control groups from 1.02 to 9.71 standard deviation units	in overall sleep disturbance in the exercise groups was -0.46 standard deviation units lower (-0.72 to -0.20 standard deviation units lower)		**moderate**[1]	A standard deviation unit is equivalent to about a 6-point change on the PSQI
Overall social functioning change - up to 12 weeks' follow-up	The standardized mean change from baseline to up to 12 weeks' follow-up in overall social functioning ranged across control groups from -0.44 to 0.11 standard deviation units	The standardized mean change from baseline to up to 12 weeks' follow-up in overall social functioning in the exercise groups was 0.45 standard deviation units higher (0.02 to 0.87 standard deviation units higher)	386 (5 studies)	⊕⊝⊝⊝ **very low**[1,2,3,4,5]	(SMD 0.45; 95% CI 0.02 to 0.87) A standard deviation unit is equivalent to about a 25-point change on the QLQ-C30 social functioning subscale or about a 6-point change on the FACT-Social subscale
Overall social functioning change - 6 months' follow-up	The standardized mean change from baseline to 6 month follow-up in overall social functioning ranged across control groups from -0.59 to -0.31 standard deviation units	The standardized mean change from baseline to 6 month follow-up in overall social functioning in the exercise groups was 0.49 standard deviation units higher (0.11 to 0.87 standard deviation units higher)	110 (2 studies)	⊕⊕⊕⊝ **moderate**[1,3]	(SMD 0.49; 95% CI 0.11 to 0.87)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core Module; FACT-emotion: Functional Assessment of Cancer Therapy - emotion; FACT-F: Functional Assessment of Cancer Therapy - Fatigue; FACT-G: Functional Assessment of Cancer Therapy - General; FACT-Social: Functional Assessment of Cancer Therapy - Social; HADS: Hospital Anxiety and Depression Scale; HRQoL: health-related quality of life; PSQI: Pittsburgh Sleep Quality Index; SMD: standardized mean difference; STAI: State-Trait Anxiety Index.

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] It was not possible to blind study participants or people administering treatment.
[2] Statistical heterogeneity was moderate to high.
[3] The small total population sample size represents a small effect.
[4] Random sequence generation was unclear in half or more of the trials.
[5] Allocation concealment was unclear in half or more the trials.

BACKGROUND

There is a steady increase in the number of cancer survivors, that is people diagnosed with cancer (Aziz 2003), worldwide. This is due, in a large part, to the dramatic advances in cancer treatment and management (Aziz 2002; Aziz 2003), growing attention to multidisciplinary post-treatment care (Demark-Wahnefried 2000; Stull 2007), and healthier lifestyles (Demark-Wahnefried 2005; Stull 2007). These factors and trends, especially when considered in light of an aging population (Aziz 2008; Stewart 2003), suggest that we can continue to expect increasing numbers of cancer survivors with greater expected length of survival. Ensuring the quality of that survival thus becomes a key priority.

There are approximately 22 million cancer survivors worldwide (Stewart 2003), 11.7 million (in 2007) of whom are estimated to be present in the US alone (Rowland 2011). Cancer survivors represent 3% to 4% of the US population and the mean age at diagnosis is about 68 years for men and 64 years for women (National Cancer Institute 2008). The relative five-year survival rates (all cancers combined) increased steadily between 1960 and 2003. Among cancer survivors, as of January 2007, an estimated 64.8% had lived with a diagnosis of cancer for years or more and nearly 10% had lived with a cancer diagnosis for 25 years or longer (Rowland 2011). The majority (60%) of cancer survivors are ages 65 years or older (Rowland 2011) and five-year survival rates show a decrease with increasing age-at-diagnosis (Ries 2007). These findings lend support to the need for pre-emptive management strategies targeting this age group. Mortality rates (all cancers combined) appear to show stabilization at this point and even a decline for some sites (Jemal 2004; Jemal 2006), supporting estimated trends of growing numbers of cancer survivors in the future.

Description of the condition

Cancer survivors experience numerous disease- or treatment-related adverse outcomes (physiologic or psychosocial, or both) (Aziz 2002; Aziz 2003; Aziz 2007; Aziz 2008; Gotay 1998; Rao 2006) and poorer health-related quality of life (HRQoL) (Ahn 2007; Ganz 1998; Ganz 2002; Ganz 2004; Gotay 1998; Hammerlid 2001). Some of the adverse outcomes include cardiotoxicity, neurotoxicity, lymphedema, premature menopause, sexual dysfunction, infertility, and fatigue (Aziz 2002; Aziz 2003; Aziz 2007; Rao 2006), all with a negative impact on HRQoL long after the completion of active treatment. Many of the adverse outcomes carry the potential to decrease the length and quality of survival and need to be prevented or managed. Intervention strategies capable of mitigating or preventing these adverse outcomes, especially those based on exercise, energy balance, or lifestyle factors, are of great interest as these are modifiable factors. Exercise interventions are particularly relevant because they influence both the physiologic and psychosocial adverse outcomes, including HRQoL (Courneya 2007; Ingram 2007; Schmitz 2005; Warburton 2006). Further, HRQoL and its domains are important measures for cancer survivorship as they provide prognostic (Gotay 2008) and predictive (Efficace 2006; Osoba 1999; Osoba 2007) information and the survivors' subjective experiences (Bottomley 2002) to therapeutic and lifestyle interventions.

Although HRQoL has no commonly accepted definition, there is broad consensus that it is a patient-reported, multidimensional construct. Ferrans provided a comprehensive review of definitions of HRQoL and concluded that, "the literature contains a bewildering array of characterizations" (Ferrans 2005). Nonetheless, the review indicated that there is broad consensus among experts (Bottomley 2002; Gotay 1992; Lipscomb 2007; Osoba 1994) regarding the major domains of HRQoL. These domains comprise subjective assessments of physical, psychological, economic, social, and spiritual well-being. Physical function includes performance of self-care activities, mobility, and physical activities. Psychological functions include emotional well-being, anxiety, body image, and depression. Social and economic functions include work or household responsibilities and social interactions. Spiritual well-being includes perspectives on one's life as a whole. HRQoL also encompasses the negative aspects of the disease or treatment, or both, such as sexual functioning, neuropathy or cognitive changes, and chronic fatigue. Lastly, it is also important to assess positive aspects of HRQoL (Diener 2000). Our selection of the primary outcomes for this review reflects these theoretical perspectives, in that we included both all the well-agreed upon domains of HRQoL and positive aspects of well-being.

Description of the intervention

The benefits of exercise on health status, length of survival, promotion of HRQoL, and mitigating premature death are gaining wide attention (Warburton 2006). Evidence indicates exercise increases physical functioning among cancer survivors (Ingram 2007; Stevinson 2004), and facilitates positive physiologic and psychological benefits in cancer survivors during and after treatment (Galvao 2005; Ingram 2007; Knols 2005; Schmitz 2005). In addition, evidence suggests exercise enhances HRQoL in breast (McNeely 2006; Milne 2008a; Valenti 2008), ovarian (Stevinson 2007), prostate (Thorsen 2008), head and neck (Rogers 2006), bladder (Karvinen 2007), endometrial (Courneya 2005a), multiple myeloma (Jones 2004), and colorectal (Courneya 2003b) cancer survivors. Further, exercise leads to improvements in physical functioning and a reduction in fatigue symptoms in breast (McNeely 2006) and prostate (Thorsen 2008) cancer survivors. The participation in exercise programs by cancer survivors varies by factors such as age (Courneya 2007a); cancer type and stage at diagnosis (Knols 2005); site of cancer diagnosis (Knols 2005), especially for multiple myeloma (Jones 2004) and cancers of the head and neck (Rogers 2006), bladder (Karvinen 2007), and ovary (Stevinson 2007); type of medical treatment received (Knols 2005); and, the survivors' current lifestyle (Knols 2005). However, despite the growing body of literature documenting the beneficial effects of exercise in cancer survivors (Courneya 2007), several studies documented lower levels of exercise behavior among people diagnosed with cancer (Blanchard 2003; Valenti 2008; Vallance 2005).

How the intervention might work

There is tremendous interest in the association between exercise and physiologic and psychological well-being in general and HRQoL in particular. Systematic reviews on the effects of exercise interventions on cancer survivors documented improvements, during and after treatment, in cardiorespiratory fitness (McNeely 2006; Schmitz 2005), physical function (McNeely 2006; Stevinson 2004; Thorsen 2008), psychological well-being (Galvao 2005; Knols 2005), overall HRQoL (Knols 2005), and fatigue (Cramp 2008; McNeely 2006; Mustian 2007). In addition, exercise-related improvements are documented for physiologic outcomes among cancer survivors undergoing treatment (Galvao 2005; Knols 2005;

Schmitz 2005) and vigor and vitality among cancer survivors in the post treatment period (Schmitz 2005).

Why it is important to do this review

There is no systematic review examining the effect of exercise on: (a) overall HRQoL and HRQoL domains (e.g. physical, psychological, economic, social, and spiritual well-being); and, (b) disease- or treatment-related symptoms, or both (e.g. sexual functioning, neuropathy or cognitive changes, and chronic fatigue) among adult post-treatment cancer survivors. Moreover, there is little evidence on the long-term benefits of exercise on survival (Warburton 2006) or HRQoL (Knols 2005; McNeely 2006; Schmitz 2005; Stevinson 2004) and the benefits of exercise among older cancer survivors (Courneya 2004). This review is different from previous systematic reviews in the number of databases searched (Galvao 2005; Schmitz 2005) and on the inclusion criteria for the trials. Several of the previous reviews included trials with non-randomized controlled trial (non-RCT) designs (Galvao 2005; Schmitz 2005; Stevinson 2004; Thorsen 2008), cancer survivors during active treatment and in the immediate post-treatment phase (Cramp 2008; Galvao 2005; Knols 2005; McNeely 2006; Schmitz 2005; Stevinson 2004), and site-specific cancers (McNeely 2006; Thorsen 2008). This lack of documentation and evidence coupled with limitations of the previous reviews necessitated a systematic review to determine the effectiveness of exercise on HRQoL among adult cancer survivors who are beyond the active treatment period. This review complements a previously published protocol that described a systematic review determining the effectiveness of exercise interventions on HRQoL among adult cancer survivors under active treatment (Mishra 2010).

OBJECTIVES

To evaluate the effectiveness of exercise on overall HRQoL outcomes and specific HRQoL domains (e.g. physical, psychological, economic, social, and spiritual well-being, and key disease and treatment (or both) symptoms such as sexual functioning, neuropathy or cognitive changes, and chronic fatigue) among adult post-treatment cancer survivors (i.e. people with a history of cancer who are beyond active treatment, excluding those who are terminally ill and receiving hospice). We focused on post-treatment cancer survivors so that we could evaluate the effectiveness of exercise on HRQoL without adjusting for the adverse effects of cancer and/or its treatment on HRQoL.

A secondary objective, where data were available, examined the effectiveness of exercise on HRQoL outcomes among adult post-treatment cancer survivors stratified by the following:

1. age at diagnosis (i.e. less than 65 years or greater than or equal to 65 years);
2. age at trial enrolment (i.e. less than 65 years or greater than or equal to 65 years);
3. sex;
4. type of prescribed exercise (i.e. aerobic, anaerobic, combination);
5. physical condition prior to cancer treatment (i.e. obesity, heart disease, smoking status, asthma);
6. intensity of exercise (i.e. mild, moderate, vigorous); and
7. format of exercise (i.e. individual or group, professionally led or not, home or group facility).

METHODS

Criteria for considering studies for this review

Types of studies

We included only RCTs and controlled clinical trials (CCTs). The included trials assessed exercise interventions initiated after completion of active cancer treatment (i.e. surgery, chemotherapy, radiation therapy, or hormone therapy).

Types of participants

This review included cancer survivors diagnosed as adults (18 years and over) regardless of age, sex, tumor site, tumor type, tumor stage, and type of anticancer treatment received. We only included cancer survivors diagnosed with cancer as adults (18 years and over). We excluded trials including people who were terminally ill or receiving hospice care, or both, and where the majority of trial participants were undergoing active cancer treatment for either the primary or a recurrent cancer.

Types of interventions

We included trials evaluating and reporting the effects of exercise (excluding dance). We excluded trials only evaluating dance as an intervention because there is a Cochrane review on dance movement therapy for improving psychological and physical outcomes in patients with cancer (Bradt 2011). Included trials compared exercise with no exercise, another intervention, or usual care (e.g. with no specific exercise program prescribed).

We defined exercise as any physical activity causing an increase in energy expenditure, and involving a planned or structured movement of the body performed in a systematic manner in terms of frequency, intensity, and duration and is designed to maintain or enhance health-related outcomes (American College of Sports Medicine 1998; American College of Sports Medicine 2005). The primary exercise intervention included prescribed, active exercise formats of aerobic, anaerobic, or aerobic/anaerobic combinations focused upon cardiopulmonary, musculoskeletal, neuromuscular, or a combination conditioning: active or active-assisted range of motion, stretching exercises, and strengthening or resistance exercises. The specific prescribed, active exercise included but was not limited to the following methods: walking, aquatic exercise, running, sports, resistance training, yoga, tai chi, and pilates programs. The prescribed, active exercise program was individual or group, professionally led or not, and home or facility based. Exercise intensity was based on the rate of perceived exertion (RPE) or heart rate (HR), or both, with mild exercise defined as RPE of 6 to 11, HR at 30% to 54% of maximum HR, or both; moderate exercise was defined as RPE of 12 to 13, HR at 55% to 70% of maximal HR, or both; and vigorous exercise was defined as RPE of 14 to 20, HR at 71% to 95% of maximal HR, or both (American College of Sports Medicine 1998). We classified the intensity of the exercise based on RPE, HR, or both, or when a quantitative measure of intensity of the exercise intervention was not available, we used the study authors' classification of an intervention as mild, moderate, or vigorous.

Types of outcome measures

The included trials measured self-reported participant measures of HRQoL as primary or secondary end points.

Primary outcomes

1. Overall HRQoL, at four follow-up intervals: up to 12 weeks; more than 12 weeks but less than 6 months, 6 months, and more than 6 months following the exercise intervention.
2. HRQoL domains, at the four time intervals described above including, but not limited to:
   a. physical function (e.g. performance of self-care activities, mobility, physical activities);
   b. psychological function (e.g. emotional well-being, anxiety, body image, depression, negative affect);
   c. social and economic role function (e.g. performance of work or household responsibilities, social interactions);
   d. spiritual well-being;
   e. pain;
   f. vitality (e.g. energy and fatigue);
   g. general health perceptions; and
   h. positive attributes (e.g. positive affect, sense of coherence, interpersonal relationships, philosophy of life, spirituality).
3. Disease- or treatment-related symptoms, or both (e.g. sexual functioning, neuropathy or cognitive changes, chronic fatigue).

The adverse outcomes of interest included:

1. any harm associated with the exercise intervention;
2. cancer recurrence or new cancer;
3. decrease in overall HRQoL or HRQoL domain.

Search methods for identification of studies

Electronic searches

We used, at the minimum, the following databases and searches to obtain relevant studies for this review. We searched all databases from inception to the present. There were no language or date restrictions in the electronic search for trials. We utilized the search strategy for MEDLINE for the review using text and indexing terms in each database, combined with filters for RCT and CCT, and human studies (Glanville 2006). The MEDLINE search strategy was developed for precision and sensitivity and was then appropriately modified for the other databases.

1. MEDLINE (Appendix 1)
2. The Cochrane Central Register of Controlled Trials (CENTRAL) (Appendix 2)
3. EMBASE (Appendix 3)
4. CINAHL (Appendix 4)
5. PsycINFO (Appendix 5)
6. PEDRO (Appendix 6)
7. LILACS (Appendix 6)
8. SIGLE (Appendix 6)
9. SportDiscus (Appendix 6)
10. OTSeeker (Appendix 6)
11. Sociological Abstracts (Appendix 6)

We also searched citations of key authors through Web of Science and Scopus, and searched PubMed's related article feature.

Searching other resources

We performed an expanded search in order to identify additional studies for this review, including unpublished trials and references in the "gray literature". This included the following:

1. review of the reference list of all retrieved articles and other reviews on the topic;
2. contacting experts in the field of exercise and HRQoL in order to identify unpublished research;
3. searching the following websites:
   a. World Health Organization (WHO) International Clinical Trials Registry Platform (www.who.int/ictrp/en)
   b. Current Controlled Trials (www.controlled-trials.com)
   c. CenterWatch (www.centerwatch.com)
   d. CliniclTrials.gov (www.clinicaltrials.gov)
4. We did not handsearch any journals specifically for this review.

Data collection and analysis

Selection of studies

Assessment of search results

Two review authors (SM, RS), working independently, screened all the titles and abstracts resulting from the searches, and excluded articles that were clearly not relevant. We retrieved full-text copies of all trials if either author determined a trial possibly or definitely met the inclusion criteria. We translated into English, where possible, all non-English language articles. Five paired review authors (SM, RS, PG, OT, CG) independently reviewed the retrieved full-text articles and, using the defined eligibility criteria, determined their eligibility for inclusion. We did not randomly assign articles to review authors neither did we mask trial details such as trial authors, journal of publication, trial location, and institutional affiliations of the trial authors. If a need arose for clarification of any detail of a trial, we contacted the trial authors to obtain such clarification for a complete assessment of the trial's relevance for the review. We resolved by consensus any disagreement between review authors on classification of an article, either between the two review authors or through use of a third review author.

Data extraction and management

Extraction of study characteristics

For each trial, we extracted:

1. characteristics of the studies:
   a. the trial sponsors and the authors' affiliations.
   b. trial methods: study design, method of sequence generation, method of allocation concealment, masking (participant, researcher, outcome), exclusions after randomization, selective outcome reporting, loss to follow-up, and compliance.

2. characteristics of trial population:
   a. country where participants enrolled;
   b. trial inclusion and exclusion criteria;
   c. number randomized in each arm;
   d. type of control group;
   e. demographic characteristics, including age at trial enrolment, sex;
   f. type of cancer, including primary site, stage at diagnosis, and hormone dependency;
   g. age at diagnosis;
   h. physical condition prior to cancer treatment;
   i. time since diagnosis;
   j. time beyond active treatment;
3. characteristics of the intervention:
   a. type of exercise intervention in each intervention group: aerobic, anaerobic, combination;
   b. description/details of the exercise intervention: frequency, duration, intensity, total number of exercise sessions, duration of follow-up, exercise format (i.e. individual or group, professionally led or not, home or facility based);
   c. description/details of control/comparison intervention;
   d. adherence and contamination;
   e. co-intervention (e.g. medication use);
4. characteristics of the outcomes
   a. self-reported HRQoL measure or HRQoL domain measures, or both (e.g. physical, psychological, economic, social, and spiritual well-being, pain, vitality, health perceptions, positive attributes);
   b. disease or treatment symptoms, or both (e.g. sexual functioning, neuropathy or cognitive changes, and chronic fatigue);
   c. length of time between end of intervention and outcome measurement;
   d. adverse outcomes (e.g. exercise associated harm, cancer recurrence, new cancer, noncompliance with exercise program, trial attrition);
   e. economic data on cost and cost-benefit of the exercise intervention.

Data extraction and entry

Five paired review authors (SM, RS, PG, OT, CG) independently extracted data, using a standardized form, from each article. Disagreements between the review authors on the data abstracted were resolved through consensus or, when necessary, there was a meeting with a third review author not involved in the particular extraction (SM, RS). In addition, we attempted to contact all trial authors (using e-mail, letter, fax, or combination) to search for additional articles, seek clarity and additional information about trials, confirm data extraction, and obtain missing data using a structured instrument with standardized questions. If the trial authors could not provide the requested information or were unable to comply with the request within two weeks, we proceeded with the review without the information. If available, we extracted similar data for each outcome from each trial included in the review. For the primary and secondary HRQoL outcomes, if more than two time points were reported during a single interval, the one closest to 12 weeks (for the 12 weeks' follow-up time point), or the longest time interval (for the other follow-up time points) was selected for analyses. We also collected information on any harm reported in the included trials. We collected data, if reported, on cost and cost-benefit of the exercise interventions. The unit of analysis was the individual cancer survivor randomized to each arm of the trial. We entered and combined the trial data using Review Manager (RevMan 5.1) (RevMan 2011). One review author (RS) entered the data into RevMan 5.1, and another review author (SM) worked independently to verify the data entry.

Assessment of risk of bias in included studies

We (SM and RS assessed the risk of bias of all the included trials by evaluating the parameters listed on the RevMan 5.1 'Risk of bias' table, including adequate sequence generation, allocation concealment, masking or blinding, methods of addressing incomplete outcome data, selective reporting and other bias, including attrition from and adherence with the exercise intervention. We assessed and graded each trial quality parameter as high risk, low risk, or unclear risk based on recommendations for judging risk of bias provided in Chapter 8 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Data synthesis

Measurement of intervention intensity

We reported the authors' classification of the intensity of the exercise based on RPE, HR, or both, or on study authors' classification of the exercise intervention's intensity as mild, moderate, or vigorous.

Measurement of intervention effect

Trials reported data on HRQoL, HRQoL domains, or both in different ways or used different instruments to measure the same construct, but all reported continuous (versus dichotomous) outcomes. If necessary, we planned to transform outcome data to achieve consistency of results, but did not need to do so for this review. We combined data using a weighted mean difference (WMD) and a random-effects model when trials measured HRQoL or HRQoL domains using either the same measurement method or scale to generate continuous data. We used a standardized mean difference (SMD) analysis and random-effects model to combine data from different instruments measuring the same domain. When there was significant clinical or statistical heterogeneity, we performed subgroup analyses or provided a qualitative analysis rather than a quantitative analysis of HRQoL or HRQoL domains.

Wherever possible, we conducted subgroup analysis of treatment effect based on:

1. grouping of the exercise intervention on:
   a. type (i.e. aerobic, anaerobic, combination);
   b. intensity (i.e. mild, moderate, vigorous); and
   c. format (i.e. individual or group, professionally led or not, home or facility based);

2. grouping of cancer survivors on:
   a. sex;
   b. age at trial enrolment (i.e. less than 65 years or 65 years and over);
   c. age at diagnosis (i.e. less than 65 years or 65 years and over); and
   d. physical condition prior to cancer treatment; and
   e. cancer type.

To investigate publication bias, we prepared funnel plots and visually examined them for signs of asymmetry. We followed the recommendations in Chapter 10 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Sterne 2011) for any statistical testing for funnel plot asymmetry. If there was statistically significant asymmetry, we considered interpretations other than publication bias.

We combined data from trials in a meta-analysis when appropriate to pool for a meta-analysis, that is, those data showing no significant clinical heterogeneity. We evaluated clinical heterogeneity by examining differences in type of cancer, exercise intervention, and overall HRQoL or HRQoL domains among trials. When there was moderate clinical heterogeneity, we conducted prespecified subgroup analyses (i.e. cancer type, intensity of exercise, etc. as mentioned above). We also checked for statistical heterogeneity by visual inspection of forest plots and by using the Chi^2 and I^2 tests. When there was significant heterogeneity as demonstrated by a statistically significant Chi^2 test or I^2 above 50%, we investigated sources for heterogeneity and if possible, conducted a quantitative meta-analysis by subgroups only. We pooled all trials (or all similar trials) for a random-effects meta-analysis to determine the pooled intervention effect estimate (odds ratio (OR) and 95% confidence interval (CI)).

We also conducted sensitivity analysis to assess the effects of including studies with a high risk of bias.

RESULTS

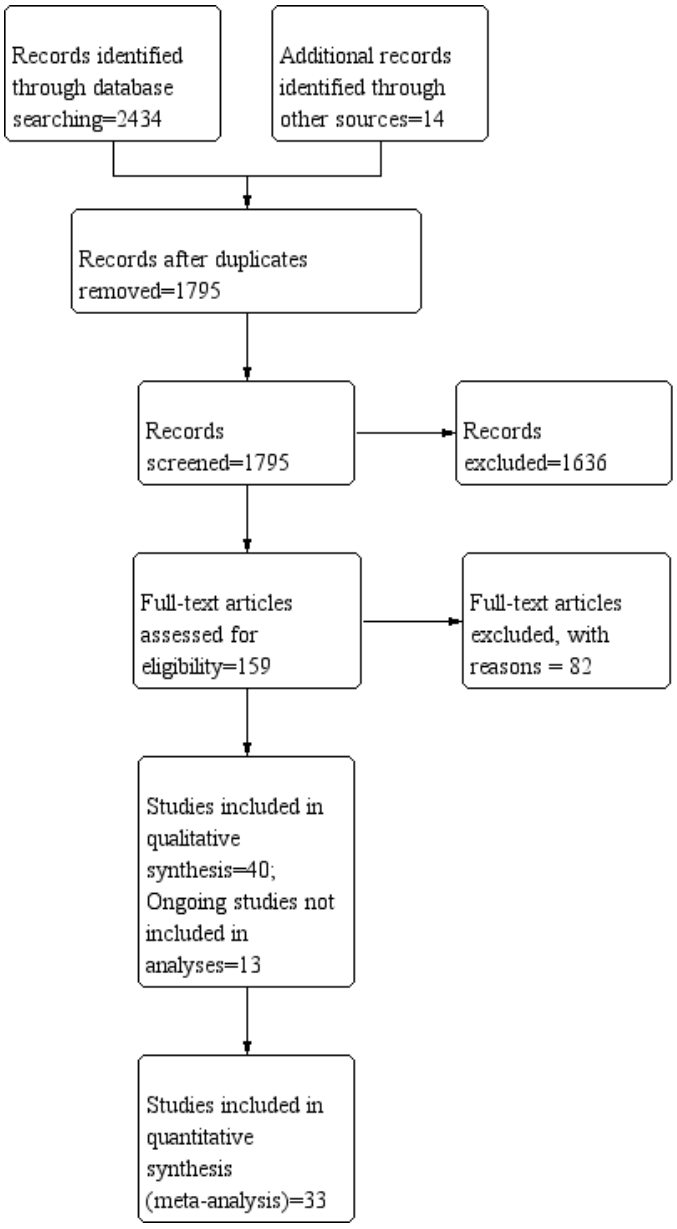
Description of studies

Results of the search

See: Characteristics of included studies; Characteristics of excluded studies; Characteristics of ongoing studies

Through a comprehensive literature search, we identified and screened for retrieval 1795 non-duplicate potentially relevant references. We excluded a total of 1636 references based on the title and abstract and retrieved 159 references for more detailed evaluation. From these, 82 trials were excluded as they did not meet the inclusion criteria and 40 trials were identified as appropriate for inclusion in the current review. In addition, 12 trials (Devonish 2007; Galvao 2009; Hayes 2011; Jones 2010; Jones 2010a; Kampshoff 2010; Persoon 2010; Saxton 2006; Sekse 2011; Spence 2007; Vardy 2010; Walsh 2010) were ongoing and one trial (Utz-Billing 2010) was awaiting classification and these trials were not included in the analysis presented below but will be considered in future updates of this review. All searches were completed in October 2011. See Figure 1 for a flowchart of the search process based on the PRISMA template (Moher 2009).

Figure 1. Study flow diagram.



Included studies

The final selection based on consensus resulted in 40 trials being included in this review (Bai 2004; Banasik 2011; Berglund 1994; Bourke 2011; Burnham 2002; Cadmus 2009; Cho 2006; Cohen 2004; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Culos-Reed 2006; Daley 2007a; Danhauer 2009; Dimeo 2004; Dodd 2010; Donnelly 2011; Fillion 2008; Heim 2007; Herrero 2006; Knols 2011; McNeely 2008a; Mehnert 2011; Milne 2008a; Moadel 2007; Mustian 2004; Oh 2008; Oh 2010; Ohira 2006; Payne 2008; Penttinen 2011; Pinto 2003; Pinto 2005; Rogers 2009; Segar 1998; Speck 2010; Tang 2010; Targ 2002; Thorsen 2005). We also reviewed and included information on trial characteristics and outcome-related data from an additional 24 publications that were secondary publications to several of the 40 trials. We corresponded with and requested additional data from 11 trial authors (Daley 2007a; Dodd 2010; Heim 2007; Herrero 2006; Mehnert 2011; Oh 2008; Payne 2008; Penttinen 2011; Rogers 2009; Tang 2010; Thorsen 2005), and six of these trials authors were able to provide additional data. We were unable to find the correct corresponding address for one author (Berglund 1994). For trial characteristics and outcomes see the Characteristics of included studies table.

Overall study characteristics

Of the 40 included trials, 38 were RCTs, although three trials used variations of the RCT design in that two used a cross-over design (Culos-Reed 2006; Milne 2008a) and one randomized clusters (Courneya 2003a), where clusters were psychotherapy classes. Two trials (Cho 2006; Heim 2007) used a quasi-randomized design to allocate participants to treatment. All trials, except for four (Burnham 2002; Daley 2007a; Dodd 2010; Segar 1998), randomized eligible participants to either the exercise or comparison arm. The additional study group in these four trials comprised variations in

the exercise arm, such as low-intensity exercise group or moderate-intensity exercise group (Burnham 2002), exercise-therapy group or exercise-placebo group (Daley 2007a), group that began exercise during treatment or group that began exercise after treatment (Dodd 2010), and an exercise group or an exercise and behavioral modification group (Segar 1998). In all, 3694 (range: 18 to 573) participants were randomized to an exercise intervention(s) (n = 1927, range: 9 to 302) or the comparison group (n = 1764, range: 7 to 271). In one trial, the number of participants randomized to the exercise and comparison arms did not add up to the number of participants randomized in the trial (Dimeo 2004). For detailed information on overall study characteristics see Characteristics of included studies table.

Participants

Participants enrolled in the trials had various cancer diagnoses including breast, colorectal, head and neck, and other. Twenty-two trials investigated participants with breast cancer only (Banasik 2011; Cadmus 2009; Cho 2006; Courneya 2003c; Daley 2007a; Danhauer 2009; Fillion 2008; Heim 2007; Herrero 2006; Mehnert 2011; Milne 2008a; Moadel 2007; Mustian 2004; Ohira 2006; Payne 2008; Penttinen 2011; Pinto 2003; Pinto 2005; Rogers 2009; Segar 1998; Speck 2010; Targ 2002) and an additional 12 trials investigated participants with a range of cancer diagnoses (Berglund 1994; Burnham 2002; Courneya 2003a; Culos-Reed 2006; Dimeo 2004; Dodd 2010; Donnelly 2011; Knols 2011; Oh 2008; Oh 2010; Tang 2010; Thorsen 2005). Thirty trials were conducted among participants who had completed active treatment for their cancer, and the remaining 10 trials included participants both during and post cancer treatment (Cohen 2004; Courneya 2003a; Courneya 2009; Danhauer 2009; Donnelly 2011; Moadel 2007; Oh 2008; Oh 2010; Tang 2010; Targ 2002). One of these reported data separately on trial participants who completed treatment (Moadel 2007) and we included only these data in this review. Twenty-three trials reported the time beyond active treatment, which ranged from immediate end of treatment to years beyond the end of active cancer treatment (Bai 2004; Banasik 2011; Bourke 2011; Burnham 2002; Cadmus 2009; Cho 2006; Courneya 2003b; Culos-Reed 2006; Daley 2007a; Danhauer 2009; Dimeo 2004; Fillion 2008; Herrero 2006; Knols 2011; McNeely 2008a; Mehnert 2011; Milne 2008a; Mustian 2004; Ohira 2006; Penttinen 2011; Pinto 2003; Segar 1998; Thorsen 2005). Seventeen trials reported the time since cancer diagnosis and it ranged across the trials from immediately after surgery to about 15 years (Cadmus 2009; Cho 2006; Courneya 2003a; Courneya 2009; Culos-Reed 2006; Danhauer 2009; Dimeo 2004; Donnelly 2011; Fillion 2008; Milne 2008a; Moadel 2007; Ohira 2006; Pinto 2003; Pinto 2005; Speck 2010; Tang 2010; Targ 2002).

Twenty-four trials were conducted among females only (Banasik 2011; Cadmus 2009; Cho 2006; Courneya 2003c; Daley 2007a; Danhauer 2009; Dodd 2010; Donnelly 2011; Fillion 2008; Heim 2007; Herrero 2006; Mehnert 2011; Milne 2008a; Moadel 2007; Mustian 2004; Ohira 2006; Payne 2008; Penttinen 2011; Pinto 2003; Pinto 2005; Rogers 2009; Segar 1998; Speck 2010; Targ 2002) and 15 trials included a mixed sample of males and females (Bai 2004; Bourke 2011; Burnham 2002; Cohen 2004; Courneya 2003a; Courneya 2003b; Courneya 2009; Culos-Reed 2006; Dimeo 2004; Knols 2011; McNeely 2008a; Oh 2008; Oh 2010; Tang 2010; Thorsen 2005), with one trial not reporting on the gender of the participants (Berglund 1994). The mean age of the participants ranged between 39 and 68 years, with one trial not reporting on the age of the participants (Berglund 1994). The ethnicity of the participants was reported by 18 trials and 27 trials reported on the education level of the participants, with the majority of trials reporting educational attainment of more than high school. Fifteen trials reported on the socio-demographic status of the participants and 19 trials reported on the employment status of the participants. Fifteen trials reported on the past exercise history of the participants (Cadmus 2009; Cohen 2004; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Daley 2007a; Danhauer 2009; Dodd 2010; Heim 2007; Knols 2011; McNeely 2008a; Mehnert 2011; Penttinen 2011; Targ 2002). For detailed information on trial participants see the Characteristics of included studies.

Interventions

Mode of exercise differed across trials. Six trials prescribed strength training by itself (Ohira 2006; Speck 2010) or in combination with walking, stretching, cardiovascular activity, or resistance training (Heim 2007; Herrero 2006; McNeely 2008a; Pinto 2003); five trials prescribed resistance training by itself (Bourke 2011) or in combination with cycling or strength training (Burnham 2002; Knols 2011; McNeely 2008a; Milne 2008a), four trials prescribed walking (Fillion 2008; Payne 2008; Rogers 2009; Tang 2010), and one trial prescribed walking with strength training (Heim 2007); three trials prescribed cycling (Courneya 2003c; Courneya 2009; Dimeo 2004) and four trials prescribed cycling with resistance training (Burnham 2002; Knols 2011; Milne 2008a) or strength training (Herrero 2006). Four trials prescribed yoga (Banasik 2011; Cohen 2004; Culos-Reed 2006; Danhauer 2009) and three trials incorporated practices of Qigong (Oh 2008; Oh 2010) or Tai Chi (Mustian 2004). Seventeen trials incorporated a range of modalities or allowed participants to choose from a range of preferred modalities (Bai 2004; Berglund 1994; Cadmus 2009; Cho 2006; Courneya 2003a; Courneya 2003b; Daley 2007a; Dodd 2010; Donnelly 2011; Mehnert 2011; Milne 2008a; Penttinen 2011; Pinto 2003; Pinto 2005; Segar 1998; Targ 2002; Thorsen 2005).

In the majority of trials (n = 32) the comparison arm did not receive an exercise prescription (i.e. 'usual care' or 'no intervention') during the course of the trial and for 14 of these trials (Banasik 2011; Cho 2006; Cohen 2004; Courneya 2003b; Courneya 2003c; Courneya 2009; Culos-Reed 2006; Danhauer 2009; Milne 2008a; Moadel 2007; Ohira 2006; Pinto 2003; Speck 2010; Tang 2010), the comparison arm was a 'waiting list' control wherein participants were offered either a portion or the full exercise program at the completion of the trial. The comparison group in eight trials received an intervention that included educational program, physical therapy, group exercise, and psycho-oncological interventions (Heim 2007); group psychotherapy (Courneya 2003a); information and coping skill training (Berglund 1994); unstructured psycho-educational support groups (Targ 2002); psychosocial support therapy (Mustian 2004); progressive relaxation training (Dimeo 2004); light-intensity body conditioning/stretching (e.g. flexibility and passive stretching) exercises (Daley 2007a); and supervised active and passive range of motion/stretching exercises, postural exercises, and basic strengthening exercises with light weights (1 to 5 kg) and elastic resistance bands (McNeely 2008a).

Length of the exercise intervention varied greatly between trials with a range from three weeks (Dimeo 2004) to one year (Penttinen 2011; Speck 2010), with a modal exercise intervention period of 12 weeks (n = 13 trials). The majority of trials (n = 26) had no follow-up period between the end of the exercise intervention and the postexercise assessment. Among the 14 trials with a follow-up

period, this period ranged from two months (Tang 2010) to one year (Berglund 1994), with a modal length of three months from the end of the intervention (n = 6). Thirty trials implemented an aerobic exercise program and an additional nine trials implemented a combined (aerobic and anaerobic) exercise program. The nature of exercise program for one trial was unclear (Courneya 2003b).

The intensity of the exercise varied substantially between trials as did the methods used to measure and monitor intensity. Methods used to measure intensity of the exercise included relatively objective measures such as percentage of the maximum heart rate, percentage of maximum oxygen consumption, heart rate and ratings of perceived exertion, and perceived effort to reach a value on the Borg scale (Bourke 2011; Cadmus 2009; Cho 2006; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Daley 2007a; Dimeo 2004; Dodd 2010; Knols 2011; McNeely 2008a; Mehnert 2011; Pinto 2003; Pinto 2005; Segar 1998; Thorsen 2005). Sixteen trials used a relatively subjective assessment of intensity by documenting a rating of mild, low -o moderate, mild to moderate, or vigorous (Bai 2004; Banasik 2011; Burnham 2002; Cohen 2004; Danhauer 2009; Donnelly 2011; Milne 2008a; Moadel 2007; Mustian 2004; Oh 2008; Oh 2010; Payne 2008; Penttinen 2011; Rogers 2009; Tang 2010; Targ 2002).

The frequency and duration of individual exercise sessions, and the total number of exercise sessions varied greatly across the trials. Frequency of the exercise program ranged between once per week to daily and in some trials participants attended exercise sessions at a facility such as a gym, community center, or university or hospital facility and were advised to practice at home. Duration of exercise sessions ranged from 20 minutes to more than 90 minutes, with a modal duration of 90 minutes (n = 7). In some trials the frequency of the exercise program and duration of each exercise session increased during the course of the trial. The total number of exercise sessions ranged between 7 and 12 sessions. In terms of the format of implementing the exercise program, 12 trials each used a group or individual format and an additional 11 trials used a combined group and individual format. The majority of trials (n = 20) implemented the exercise program in a facility such as a gym, community center, yoga studio, or university or hospital facility (Banasik 2011; Berglund 1994; Burnham 2002; Courneya 2003c; Courneya 2009; Culos-Reed 2006; Daley 2007a; Danhauer 2009; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Milne 2008a; Mustian 2004; Oh 2008; Oh 2010; Ohira 2006; Segar 1998; Speck 2010; Targ 2002), 12 trials implemented the exercise program at both a facility and the participant's home (Bourke 2011; Cadmus 2009; Cho 2006; Cohen 2004; Courneya 2003a; Fillion 2008; Heim 2007; McNeely 2008a; Moadel 2007; Penttinen 2011; Pinto 2003; Rogers 2009), seven trials implemented the exercise program only at the participant's home (Courneya 2003b; Dodd 2010; Donnelly 2011; Payne 2008; Pinto 2005; Tang 2010; Thorsen 2005), and one trial did not report the location of implementation of the exercise program (Bai 2004). The majority of the trials (n = 28) enlisted the services of exercise physiologists, sports trainers, yoga instructors, or other professionals to lead the exercise program.

For detailed information on interventions see Characteristics of included studies table.

Outcome measures

See: Table 1 for a summary of instruments, the HRQoL domains assessed, and trials using each scale.

HRQoL assessment included a wide range of measures including the European Organization for Research and Treatment of Cancer (EOTRC) Quality of Life Questionnaire-C30 (QLQ-C30) (Culos-Reed 2006; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011; Thorsen 2005), Functional Assessment of Cancer Therapy - General (FACT-G) (Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Daley 2007a; Donnelly 2011; Heim 2007; McNeely 2008a; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009), Functional Assessment of Cancer Therapy - Breast (FACT-B) (Banasik 2011; Cadmus 2009; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Rogers 2009), Functional Assessment of Cancer Therapy - Colorectal (FACT-C) (Bourke 2011; Courneya 2003b), Functional Assessment of Cancer Therapy - Fatigue (FACT-F) (Heim 2007), Cancer Rehabilitation Evaluation System Short Form (CARES-SF) (Ohira 2006), Chae and Cho (Cho 2006), Functional Assessment of Cancer Therapy - Anemia (FACT-An) (Courneya 2009; McNeely 2008a), Quality of Life for Cancer Patients (QoL Index) (Burnham 2002), Medical Outcomes Study Short Form-36 (MOS SF-36) (Cadmus 2009; Mehnert 2011), and Functional Assessment of Chronic Illness Therapy - Fatigue (FACIT-F) (Mustian 2004; Targ 2002). Some trials incorporated condition-specific HRQoL measures such as Functional Assessment of Cancer Therapy - lymphoma (FACT-Lym) (Courneya 2009) for a measure of lymphoma symptoms and the Neck Dissection Impairment Index (McNeely 2008a) for treatment-specific quality of life (QoL) for head and neck cancer survivors.

In addition to measuring overall HRQoL, trials measured HRQoL domains including:

- anxiety measured using the Hospital Anxiety and Depression Scale (HADS) (Berglund 1994; Heim 2007; Mehnert 2011; Thorsen 2005), State-Trait Anxiety Scale (STAI) (Cadmus 2009; Cohen 2004; Segar 1998), Linear Analog Self Assessment (LASA) (Burnham 2002), and Profile of Mood Scale (POMS) (Culos-Reed 2006; Moadel 2007; Oh 2010; Pinto 2003; Targ 2002);
- body image/self-esteem measured using the Body Esteem Scale (BES) (Pinto 2003; Pinto 2005), Physical Self-Perception Profile (PSPP) (Daley 2007a), Body Image Questionnaire (BIQ) (Mehnert 2011), Body Image and Relationships Scale (BIRS) (Speck 2010), Social Physique Anxiety Scale (SPAS) (Milne 2008a), and the Rosenberg Self-Esteem scale (Cadmus 2009; Courneya 2003c; Mustian 2004; Segar 1998);
- cognitive function measured using the QLQ-C30 (Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011), Symptoms of Stress Inventory (SOSI) (Culos-Reed 2006), Functional Assessment of Cancer Therapy - Cognitive (FACT-C) (Oh 2010; Rogers 2009), POMS (Culos-Reed 2006; Moadel 2007; Oh 2010; Pinto 2003; Targ 2002), and LASA (Burnham 2002);
- depression measured using the Centers for Epidemiologic Studies - Depression Scale (CES-D) (Cadmus 2009; Cohen 2004; Courneya 2003a; Courneya 2003b; Courneya 2009; Danhauer 2009; Dodd 2010; Payne 2008), HADS (Berglund 1994; Mehnert 2011; Thorsen 2005), Beck Depression Inventory-II (BDI) (Daley 2007a; Donnelly 2011; Segar 1998), Finnish Version of BDI (Penttinen 2011), LASA (Burnham 2002), and POMS (Culos-Reed 2006; Oh 2010; Pinto 2003; Targ 2002),
- emotional function/mental health measured using the FACT-B (Banasik 2011; Cadmus 2009; Courneya 2003b; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009), FACIT-F (Targ 2002), POMS (mood)

(Culos-Reed 2006; Moadel 2007; Oh 2010; Pinto 2003; Pinto 2005; Targ 2002), POMS (anger-hostility) (Oh 2010; Pinto 2003; Targ 2002), POMS (anxiety and depression scales) (Fillion 2008), POMS (irritability) (Moadel 2007; Oh 2010), QLQ-C30 (Bai 2004; Culos-Reed 2006; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011; Thorsen 2005), SOSI (Culos-Reed 2006), LASA (Burnham 2002), Psychosocial scale (Cho 2006), Fordyce Happiness Measure (Fordyce) (Cadmus 2009), Happiness Measure (HM) (Courneya 2003c; Courneya 2009), Medical Outcomes Study Short Form-12 (MOS SF-12) (Danhauer 2009; Fillion 2008), MOS SF-36 (Cadmus 2009; Speck 2010; Tang 2010), Positive and Negative Affect Scale (PANAS) (Danhauer 2009; Donnelly 2011; Pinto 2003), CARES-SF (Ohira 2006), FACT-B (Banasik 2011; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Rogers 2009), Satisfaction with Life Scale (SWLS) (Courneya 2003a; Courneya 2003b; Daley 2007a), Cohen's Perceived Stress Scale (Cadmus 2009), SOSI (Berglund 1994; Mehnert 2011; Thorsen 2005), and Symptom Checklist-90 Revised (SCL-90R) (Mehnert 2011).

- fatigue, a HRQoL domain, was measured using the QLQ-C30 (Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Thorsen 2005), FACT-F (Bourke 2011; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Danhauer 2009; Donnelly 2011; Heim 2007; Knols 2011; McNeely 2008a; Rogers 2009), POMS (fatigue-inertia) (Oh 2010; Pinto 2003; Targ 2002), POMS (vigor-activity) (Fillion 2008; Oh 2010; Pinto 2005; Targ 2002), LASA (Burnham 2002), Schwartz Cancer Fatigue Scale (SCFS) (Milne 2008a), Multidimensional Fatigue Inventory (MFI) (Donnelly 2011; Fillion 2008; Heim 2007), Revised Piper Fatigue Scale (PFS) (Daley 2007a; Dodd 2010; Payne 2008), FACIT-F (Penttinen 2011), Linear Analog Scale for Fatigue (Pinto 2005), MOS SF-36 (Cadmus 2009), and Brief Fatigue Inventory (BFI) (Cohen 2004).
- general health perspective, a HRQoL domain, was measured using the QLQ-C30 (Dimeo 2004; Donnelly 2011; Knols 2011; Mehnert 2011), MOS SF-12 (Courneya 2009), and based on a single question on health (Rogers 2009).
- pain measured using the QLQ-C30 (Dimeo 2004; Knols 2011; Mehnert 2011), MOS SF-36 (Cadmus 2009), Shoulder Pain and Disability Index (SPADI) (McNeely 2008a), Worst Pain Intensity Scale (WPIS) (Dodd 2010), and Brief Pain Inventory (BPI) (Fillion 2008).
- physical well-being measured using the CARES-SF (Ohira 2006), QLQ-C30 (Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Thorsen 2005), FACT-B (Banasik 2011; Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Daley 2007a; Danhauer 2009; Heim 2007; Mehnert 2011; Milne 2008a; Moadel 2007; Rogers 2009), FACIT-F (Targ 2002), MOS SF-12 (Danhauer 2009; Fillion 2008), BIRS (Speck 2010), MOS SF-36 (Cadmus 2009; Speck 2010; Tang 2010), and BES (Pinto 2003; Pinto 2005).
- sexuality measured using the BES (Pinto 2003; Pinto 2005), CARES-SF (Ohira 2006), and the BIRS (Speck 2010).
- sleep measured using the QLQ-C30 (Dimeo 2004; Knols 2011; Mehnert 2011; Penttinen 2011), Pittsburg Sleep Quality Index (PSQI) (Cohen 2004; Danhauer 2009; Donnelly 2011; Payne 2008; Rogers 2009), General Sleep Disturbance Scale (GSDS) (Dodd 2010), and the Taiwanese PSQI (Tang 2010).
- role function measured using FACT (Banasik 2011; Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Daley 2007a; Danhauer 2009; Heim 2007; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009), QLQ-C30 (Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011), FACIT-F (Targ 2002), MOS SF-36 (Cadmus 2009), and CARES-SF (Ohira 2006).
- social function measured using the QLQ-C30 (Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011), FACT (Banasik 2011; Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009), FACIT-F (Targ 2002), MOS SF-36 (Cadmus 2009), Social Barriers Scale (Mehnert 2011), BIRS (Speck 2010), and BIQ (Mehnert 2011).
- spiritual function measured using the FACT-B (Courneya 2003a; Rogers 2009; Targ 2002), Functional Assessment of Chronic Illness Therapy - Spiritual (FACIT-Sp) (Danhauer 2009; Moadel 2007), and Principles of Living Survey (Targ 2002).

Several trials measured HRQoL and non-HRQoL outcomes. The most frequently measured non-HRQoL outcomes included body composition or anthropometric measures (n = 11), fitness (n = 10), physiologic measures (n = 9), physical function measures (n = 8), and physical activity (n = 8). Other non-HRQoL outcomes assessed included physical outcomes, flexibility, nausea, and vomiting.

Twelve trials measured both HRQoL and non-HRQoL outcomes but did not identify a primary outcome (Bai 2004; Berglund 1994; Bourke 2011; Cadmus 2009; Cho 2006; Courneya 2003a; Dodd 2010; Heim 2007; Payne 2008; Penttinen 2011; Pinto 2003; Rogers 2009). Among the 28 trials that identified a primary outcome, 10 trials measured only HRQoL outcomes that were designated as primary (Banasik 2011; Cohen 2004; Danhauer 2009; Dimeo 2004; Donnelly 2011; Mehnert 2011; Milne 2008a; Segar 1998; Tang 2010; Targ 2002). Of the remaining 18 trials, 15 trials measured both HRQoL and non-HRQoL outcomes and identified HRQoL outcomes as primary (Burnham 2002; Courneya 2003b; Courneya 2003c; Culos-Reed 2006; Daley 2007a; Fillion 2008; Herrero 2006; Knols 2011; McNeely 2008a; Moadel 2007; Mustian 2004; Oh 2008; Oh 2010; Ohira 2006; Pinto 2005), and three trials identified non-HRQoL outcomes as primary (Courneya 2009; Speck 2010; Thorsen 2005).

For detailed information on outcome measures see Characteristics of included studies table.

Excluded studies

The 82 trials retrieved and subsequently excluded did not meet the inclusion criteria for the following reasons: 26 trials did not compare exercise with no exercise, another intervention, or usual care (Basen-Engquist 2006; Carmack Taylor 2004; Carmack Taylor 2006; Cheung 2003; Courneya 2004b; Demark-Wahnefried 2003; Demark-Wahnefried 2003a; Demark-Wahnefried 2006; Demark-Wahnefried 2007; Dincer 2007; Dong 2006; Elliott 2006; Kim 2011; Korstjens 2008; Livingston 2011; McClure 2010; Morey 2009; Poorkiani 2010; Sandel 2005; Snyder 2009; Vallance 2007a; Vallance 2008; van Weert 2005; van Weert 2010; von Gruenigen 2009; Zhang 2006); 20 trials did not measure overall HRQoL or an HRQoL domain as a trial outcome (Bloom 2008; Carson 2009; Courneya 2004c; Courneya 2004a; Courneya 2005; Daley 2007; Duijts 2009; Fairey 2005; Fairey 2005a; Filocamo 2005; Kim 2010; Lazowski 1999; Ligibel 2008; May 2008; Milne 2008; Mustian 2006; Nikander 2007; Pinto 2009; Twiss 2009; Wall 2000); eight trials included participants all, or a majority, of whom were undergoing active treatment for their cancer (Carmack Taylor 2007; Courneya 2008; Griffith 2009; Houborg 2006; Jarden 2009; Mutrie 2007; Segal 2001; Segal 2003);

two trials were not an RCT or CCT (Blanchard 2001; Gordon 2005); and one included participants below 18 years of age (Braam 2010). Additionally, six trials were excluded because they focused on complications due to treatment (e.g. lymphedema or menopause) rather than improving whole body function or HRQoL (Beurskens 2007; Cinar 2008; Kilbreath 2006; McKenzie 2003; McNeely 2004) or participants were not cancer survivors or undergoing treatment for cancer (Osei-Tutu 2005). The remaining 19 trials were excluded for meeting more than one of the reasons for exclusion (Cheema 2006; Elkin 1998; Emslie 2007; Hayes 2004; Hughes 2004; Hughes 2008; Jones 2004a; Jones 2008; Kolden 2002; Mansky 2006; Mathewson-Chapman 1997; Matthews 2007; Midtgaard 2006; Rabin 2006; Schneider 2007; Sprod 2005; Vallance 2007; Vallance 2008a; Yeh 2011). For detailed information on reasons for exclusion of retrieved studies see the Characteristics of excluded studies table.

Risk of bias in included studies

The included studies were assessed for risk of bias using the 'Risk of Bias' assessment tool and recommendations for judging risk of bias provided in Chapter 8 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). For each trial the risk of bias was detailed in the 'risk of bias' tables included with the Characteristics of included studies and the 'risk of bias' summary (Figure 2). In addition, an overall assessment of risk of bias is presented in Figure 3.

Figure 2. Risk of bias summary: review authors' judgments about each risk of bias item for each included study.

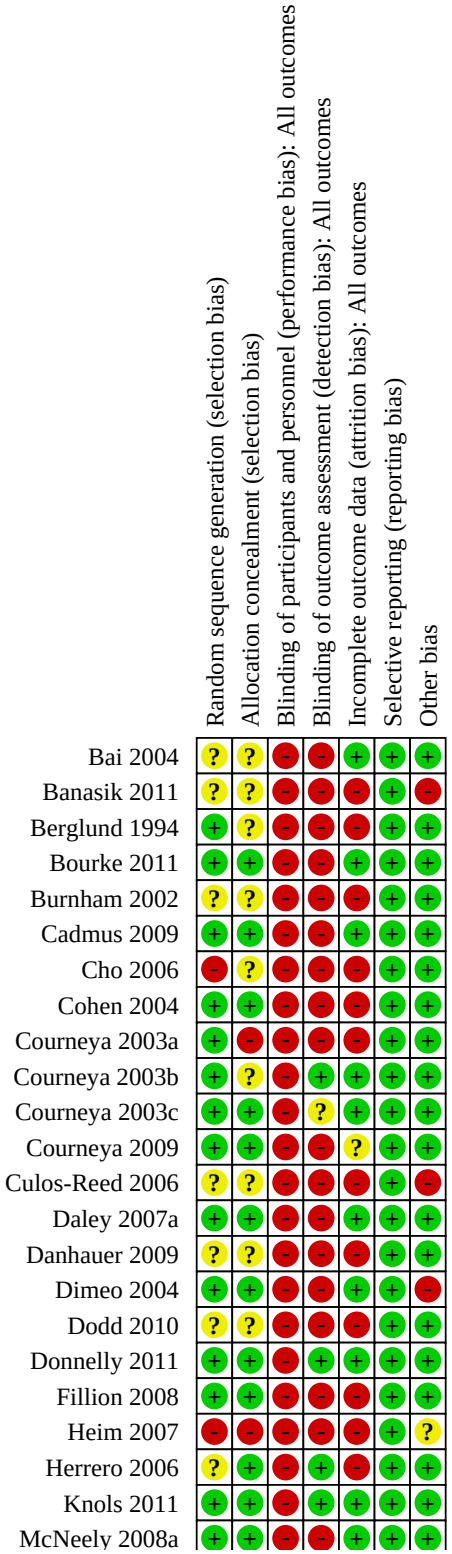


Figure 2. (Continued)

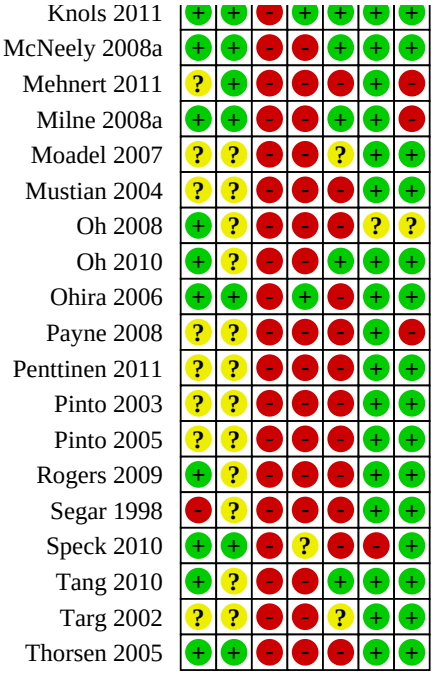
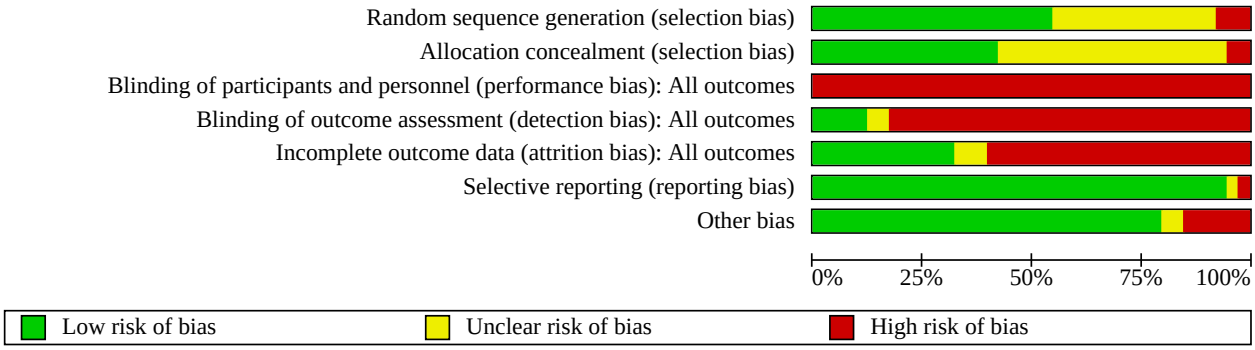


Figure 3. Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies.



Allocation

Twenty-two trials were at a low risk of selection bias due to adequate generation of the randomized sequence as the trials used a random component to generate their sequence. Three trials had a high risk of selection bias as they used a non-random component to generate their sequence (Cho 2006; Heim 2007; Segar 1998). Fifteen trials were considered to have an unclear risk of selection bias, largely because the generation of the random sequence was not described (Bai 2004; Banasik 2011; Burnham 2002; Culos-Reed 2006; Danhauer 2009; Dodd 2010; Herrero 2006; Mehnert 2011; Moadel 2007; Mustian 2004; Payne 2008; Penttinen 2011; Pinto 2003; Pinto 2005; Targ 2002).

Seventeen trials were at a low risk of selection bias owing to inadequate concealment of allocation to the intervention as the participants and investigators could not foresee assignment to the study groups. Two trials had a high risk of selection bias as the participants or investigators might have foreseen assignment to the study groups (Courneya 2003a; Heim 2007). Twenty-one trials were considered to have an unclear risk of selection bias owing to allocation concealment, largely because the method of concealment either was not described or not described in detail to allow a definite judgment (Bai 2004; Banasik 2011; Berglund 1994; Burnham 2002; Cho 2006; Courneya 2003b; Culos-Reed 2006; Danhauer 2009; Dodd 2010; Moadel 2007; Mustian 2004; Oh 2008; Oh 2010; Payne 2008; Penttinen 2011; Pinto 2003; Pinto 2005; Rogers 2009; Segar 1998; Tang 2010; Targ 2002).

Blinding

All trials included in this review were at high risk for performance bias because, owing to the nature of the intervention (exercise), it was not possible to blind the trial personnel and participants.

Five trials were at a low risk for detection bias because the outcome assessors were unaware about the allocation of the participants to the study groups (Courneya 2003b; Donnelly 2011; Herrero 2006; Knols 2011; Ohira 2006), although this was typically for outcome assessors measuring physiologic outcomes rather than HRQoL outcomes. Two trials were considered to have unclear risk for detection bias (Courneya 2003c; Speck 2010). Thirty-three trials were at high risk for detection bias.

Incomplete outcome data

Twelve trials were at a low risk of attrition bias due to the amount, nature, or handling of incomplete outcome data (Bai 2004; Bourke 2011; Cadmus 2009; Courneya 2003b; Courneya 2003c; Daley 2007a; Dimeo 2004; Donnelly 2011; McNeely 2008a; Milne 2008a; Oh 2010; Tang 2010) and three trials were considered to have an unclear risk for attrition bias (Courneya 2009; Moadel 2007; Targ 2002). Twenty-five trials were at high risk for attrition bias.

Selective reporting

Thirty-eight trials were at a low risk of reporting bias as, based on the information provided by the trial authors, there was no reason to believe that there was selective reporting of the primary and secondary outcomes. One trial each was considered at high risk (Speck 2010) or unclear risk (Oh 2008) for reporting bias.

Other potential sources of bias

Thirty-two trials were at a low risk for other biases such as sample size, description of study sample, and generalizability of findings. Six trials were considered to be at high risk for other biases (Banasik 2011; Culos-Reed 2006; Dimeo 2004; Mehnert 2011; Milne 2008a; Payne 2008) and two trials were at unclear risk for other biases (Heim 2007; Oh 2008).

Effects of interventions

See: **Summary of findings 1** Summary of findings

Authors reported trial results either as change in score from baseline to follow-up or follow-up values. We completed meta-analyses for both types of outcomes and for each follow-up time period. Because the change in scores from baseline to follow-up take into account baseline variability, we preferentially report those pooled results here. We also include pooled analyses of follow-up values, both combining all forms within a domain, but also within individual instruments. We use these latter data to show consistency across results within domains. In general, when we observed a significant effect, it was usually at 12 weeks or at only one follow-up period, although there were fewer studies at later time points. When we found heterogeneity, we investigated subgroups by cancer type or intensity of the exercise intervention and usually found similar estimates of treatment as within the entire group. Although all studies showed a relatively high risk of bias, we conducted a sensitivity analysis of studies where the allocation concealment scored as low risk of bias versus unclear or with a high risk of bias. We did not complete subgroup analyses when there was only one trial in a subgroup. We combined data using a WMD and a random-effects model when trials measured HRQoL or HRQoL domains using either the same measurement method or scale to generate continuous data. We used an SMD analysis and random-effects model to combine data from different instruments measuring the same domain.

For detailed information on HRQoL and HRQoL domain outcomes, number of trials reporting the outcomes, number of participants on whom the outcomes were reported, statistical methods used for analysis, and effect estimates see the Data and analyses table.

Overall health-related quality of life

Change in HRQoL from baseline following an exercise intervention showed a significant improvement compared with control in 826 trial participants at 12 weeks (SMD 0.48; 95% CI 0.16 to 0.81); no difference at follow-up between 3 and 6 months in 181 participants (SMD 0.14; 95% CI -0.38 to 0.66); and improvement at 6 months in 115 participants (SMD 0.46; 95% CI 0.09 to 0.84) (Analysis 1.1). At 12 weeks' follow-up, subgroups by cancer type breast (SMD 0.57; 95% CI, 0.20 to 0.95) versus all other (SMD 0.27; 95% CI -0.00 to 0.55) showed similar results. There appeared to be an effect if the exercise was moderate to vigorous as defined by the author (SMD 0.29; 95% CI -0.00 to 0.58) but not if the exercise had been defined by the author as mild to moderate (SMD 0.46; 95% CI -0.62 to 1.53). The effect of exercise was still significant when we excluded studies that included participants who were still undergoing treatment (SMD 0.51; 95% CI 0.10 to 0.92). There were too few studies at longer follow-up times to complete subgroup analyses.

All studies showed a relatively high risk of bias, so we conducted a sensitivity analysis of studies where the allocation concealment scored as low risk of bias versus unclear or with a high risk of bias. We found that the effect of exercise on the change from baseline to 12 weeks' follow-up remained significant (SMD 0.49; 95% CI 0.08 to 0.90), with no change to the results found at 6 months' follow-up (SMD 0.46; 95% CI 0.09 to 0.84).

Because there was significant clinical and statistical heterogeneity when combining all studies in an SMD model, we also examined the treatment effect by individual HRQoL instrument. The most commonly used instruments included those in the FACT series; including FACT-An, FACT-B, FACT-C, and FACIT; and the QLQ-C30. A significant change in the HRQoL score from baseline to 12 weeks compared with change in the control group was seen at 12 weeks' follow-up with the FACT-An (MD 7.10; 95% CI 1.50 to 12.71) but not with the FACT-B (MD 9.29; 95% CI -3.73 to 22.30); FACT-G (MD 4.94; 95% CI -0.08 to 9.95), FACIT (MD 6.80; 95% CI -9.51 to 23.10), or QLQ-C30 (MD 15.66; 95% CI -7.78 to 39.09). However, in most casesonly a few studies were included in the analysis for each instrument. Similar results were seen at longer follow-up periods. No investigator reported change in HRQoL score from baseline to 12 weeks' follow-up using the FACT-C instrument.

We found similar results when we looked at the follow-up values reported rather than the differences between baseline and follow-up (Analysis 1.2). Again, we found a significant effect at 12 weeks (SMD 0.49; 95% CI 0.24 to 0.74), but not at longer follow-up periods (between 12 weeks and 6 months: SMD 0.11; 95% CI -0.10 to 0.32; 6 months: SMD 0.25; 95% CI -0.12 to 0.62). Subgroup analyses at 12 weeks' follow-up showed a significant effect for both breast (SMD 0.53; 95% CI 0.17 to 0.89) and other types of cancer (SMD 0.43; 95% CI 0.08 to 0.79), and for studies in which authors reported

that the exercise was moderate to vigorous (SMD 0.34; 95% CI 0.10 to 0.58) but not when authors reported that the exercise was mild or moderate (SMD 0.44; 95% CI -0.02 to 0.89). The effect of exercise was still significant when we excluded studies that included participants who were still undergoing treatment (SMD 0.56; 95% CI 0.26 to 0.87). Looking at the treatment effect by the individual instrument administered, we found significant effects with the FACT-C (MD 14.00; 95% CI 2.59 to 25.41) and QLQ-C30 (MD 16.41; 95% CI 1.89 to 30.93), but not for the FACT-An (MD 4.25; 95% CI -3.28 to 11.78), FACT-B (MD 9.82; 95% CI -0.76 to 20.40), FACT-G (MD 5.90; 95% CI -0.36 to 12.16), or FACIT (MD -2.60; 95% CI -21.19 to 15.99). Again, few studies contributed to each analysis neither were there sufficient studies to complete subgroup analyses or look at longer times of follow-up. Limiting the analyses to studies with a low risk of bias for allocation concealment did not change the results at 12 weeks (SMD 0.39; 95% CI 0.09 to 0.70) or longer follow-up time points.

Two trials for which we were unable to extract data also reported on HRQoL (Heim 2007; Oh 2008). These studies reported that exercise resulted in an increase in HRQoL, although the trial by Heim 2007 also showed an increase in HRQoL in the control group, while that by Oh 2008 observed no change in the control group.

Cancer-specific health-related quality of life

Although there was a significant improvement in the exercise group compared with the control group from baseline to follow-up between 12 weeks and 6 months in breast cancer concerns (SMD 0.99; 95% CI 0.41 to 1.57), we did not see an effect at either 12 weeks (SMD -0.13; 95% CI -0.41 to 0.14) or 6 months (SMD 0.14; 95% CI -0.24 to 0.51). Similar findings were obtained when we examined follow-up values rather than the difference between baseline and follow-up. In addition, concerns about lymphoma, colorectal cancer, or head and neck cancers did not demonstrate consistent significant effects when comparing the difference between baseline and follow-up or follow-up values.

Anxiety

There was a significant reduction in anxiety in the group exposed to exercise compared with the control group at 12 weeks (SMD -0.26; 95% CI -0.44 to -0.07), although this finding was not consistent at all follow-up periods (between 12 weeks and 6 months: SMD 0.06; 95% CI -0.23 to 0.35; 6 months: SMD -0.15; 95% CI -0.61 to 0.30) (Analysis 3.1). Including only studies with a low risk of bias for allocation concealment resulted in this results becoming non-significant (SMD -0.26; 95% CI -0.55 to 0.03), although only two studies contributed to this analysis. There was little statistical heterogeneity across studies, but looking at subgroups, we did not find a significant effect at 12 weeks' follow-up for breast cancer only (SMD -0.15; 95% CI -0.61 to 0.30) or for vigorous to moderate exercise (SMD -0.26; 95% CI -0.55 to 0.03), although an effect on anxiety was observed when the exercise intervention was mild to moderate (SMD -0.26; 95% CI -0.02, -0.50). There were insufficient numbers of studies to compare subgroups at longer follow-up time periods. In general, similar results were seen when we compared follow-up values rather than looking at the change from baseline to follow-up. Examination by individual instrument assessing anxiety showed a significant effect at 12 weeks' follow-up only when the POMS anxiety and tension subscale was used to assess anxiety (MD -3.20; 95% CI -5.40 to -1.00).

In addition, Berglund 1994 reported a reduction in anxiety in both exercise and control group, but did not observe a difference between groups. We could not include this data in the meta-analysis because the variances for the outcome measures were not reported.

Body image

No significant effect of exercise was observed on body image when comparing an exercise with a control intervention and looking at differences in scores between baseline and 12 weeks (SMD -1.09; 95% CI -2.29 to 0.11) or 6 months (SMD -0.05; 95% CI -0.51 to 0.40), although a significant effect was seen at follow-up between 12 weeks and 6 months (SMD -0.74; 95% CI -1.30 to -0.18) and longer than 6 months (SMD -0.49; 95% CI -0.86 to -0.13) (Analysis 4.1). There was significant heterogeneity when we analyzed the results by including all types of instruments and so we evaluated change scores for each instrument separately. We found a significant effect when 'body image' was assessed using the Rosenberg Self-Esteem scale at 12 weeks (MD 4.50; 95% CI 3.40 to 5.60), between 12 weeks and 6 months (MD 2.70; 95% CI 0.73 to 4.67), but not at 6 months (MD 0.20; 95% CI -1.50 to 1.90). No significant effect was observed for any other body image instrument.

One additional trial reported a positive change in body image for both the exercise and control groups, but did not report that there was a difference between treatment groups (Berglund 1994).

Cognitive function

We observed no significant effect of exercise on any measure of cognitive function, including subgroup analyses by type cancer or intensity of exercise. One small trial whose data were not extracted reported a significant effect on cognitive function with exercise without a corresponding effect in the control group (Oh 2008) (Analysis 5.1).

Depression

We observed no significant effect of exercise on depression in 455 participants looking at the change in score across instruments from baseline to follow-up (Analysis 6.1). Because there was heterogeneity, we examined results by subgroup including type cancer (breast versus other) and observed a significant effect for other types of cancer (SMD -0.46, 95% CI -0.72, -0.19) but not for breast cancer survivors. No differences were noted when we looked at studies by intensity of exercise (vigorous to moderate versus mild to moderate) or excluding two studies that included patients still receiving therapy (Oh 2010; Targ 2002). In contrast, we did observe a significant effect of the exercise intervention at 12 weeks looking at follow-up values (SMD -0.41; 95% CI -0.65 to -0.17) but not at longer follow-up time points, and this effect was still significant after excluding studies that included individuals still undergoing treatment (SMD -0.72; 95% CI -1.03 to -0.41). We also looked at the effect of the exercise intervention for each instrument, and observed a significant treatment effect looking at change at 12 weeks in the CES-D (MD -2.40; 95% CI -4.05 to -0.75), follow-up values in the Beck Depression Inventory (BDI) at 12 weeks (MD -4.28; 95% CI -6.01 to -2.55), and change from baseline in a visual analog scale (VAS) from baseline to 12 weeks (MD -4.28; 95% CI -6.01 to -2.55). No significant effect of the exercise intervention was noted at later time points when comparing exercise with control intervention for change in score over time.

We were unable to extract data from three trials that reported on depression (Berglund 1994; Dodd 2010; Payne 2008). Dodd 2010 and Payne 2008 both reported that exercise had no effect on depression, and Berglund 1994 observed an improvement in depression in the exercise group with a worsening of depression in the control group.

Emotional well-being

A meta-analysis of the change in score from baseline to follow-up and comparing exercise with control intervention showed a significant improvement in emotional well-being at 12 weeks' follow-up in 617 trial participants (SMD 0.33; 95% CI 0.05 to 0.61), but not at other follow-up time points (Analysis 7.1). A subgroup of breast cancer survivors did not show a significant effect (SMD 0.30, 95% CI -0.15, 0.75) whereas a subgroup of survivors of other types of cancer did (SMD 0.43, 95% CI 0.16 to 0.69). There was no significant difference when we looked at subgroups by reported exercise intensity or after excluding studies that included participants still undergoing treatment. We also found a significant effect when we compared follow-up scores between the exercise and control groups at 12 weeks (SMD 0.24; 95% CI 0.12 to 0.37).

Looking at each type of instrument separately, we found a significant difference between the exercise and the control interventions in change in the POMS total mood disturbance score from baseline to 12 weeks' follow-up (MD -8.08; 95% CI -15.03 to -1.12). There was also significant differences in change score using the CARES-SF instrument and the Lee Psychosocial scale, although these latter instruments were only used in one trial each.

One trial without extractable data reported no change over time in emotional well-being in either the exercise or control group (Oh 2008).

Fatigue

We observed a significant effect of exercise on decrease in fatigue scores in cancer survivors at follow-up of 12 weeks (SMD -0.82; 95% CI -1.50 to -0.14) and between 12 weeks and 6 months (SMD -0.42; 95% CI -0.83 to -0.02), but not at 6 months or longer (Analysis 8.1). We also observed no treatment difference whether within a subgroup of breast cancer survivors (SMD -0.28, 95% CI -0.77, 0.20) or survivors of other types of cancer (-1.47, 95% CI -3.12, 0.19), or whether the exercise intervention was reported as moderate to vigorous (SMD -1.35, 95% CI -3.08 to 0.38) or mild to moderate (SMD -0.51, 95% CI -1.27 to 0.25). The effect of exercise was not significant when we excluded studies with participants undergoing treatment (SMD -0.24; 95% CI -0.69 to 0.22). Including only change scores for individual instruments showed a significant improvement in fatigue score change for the FACT-F subscale at 12 weeks' follow-up (MD 4.33; 95% CI 2.43 to 6.22). In single studies, an effect of the exercise intervention on fatigue was observed in change from baseline in the Schwartz Cancer Fatigue scale (SCFS) (MD -2.20; 95% CI -4.32 to -0.08) or a VAS scale (MD -13.14; 95% CI -23.32 to -2.96) at 12 weeks' follow-up.

Among trials for which we could not extract data, three found no group differences in fatigue (Dodd 2010; Oh 2008; Payne 2008), while one reported an initial improvement in both exercise and control groups with the exercise group improving slightly more over time, and the control group becoming worse (Heim 2007).

General health perspective

We observed no significant effect of exercise on any reported measure of general health (Analysis 9.1). Similarly, Berglund 1994 reported no difference in general health perspective between groups.

Pain

Few trials reported on pain or change in pain related to the exercise intervention. No significant effect was obtained when pooling trials that did report change in pain over time when looking at change in reports from baseline to follow-up; however, only one trial reported pain in this way (Analysis 10.1). Looking at follow-up scores in 289 trial participants, a significant reduction in pain was observed at 12 weeks (SMD -0.29; 95% CI -0.55 to -0.04) but not at longer follow-up periods. Two trials for which we could not extract data also reported on pain (Berglund 1994; Dodd 2010). In the trial by Berglund 1994, no change in pain score was observed immediately after treatment, although it was observed that the control group experienced more pain at later time points. Dodd 2010 observed a reduction in pain in both exercise and control groups, but no difference between groups in the level of pain.

Physical functioning

We did not observe any change in physical functioning at any time point looking either at change from baseline to follow-up or at follow-up values (Analysis 11.1). It was not possible to compare by intensity of exercise because of the limited number of studies. However, looking at individual instruments, a significant effect was observed in differences in scores from baseline to 12 weeks in the QLQ-C30 (MD 6.23; 95% CI 1.74 to 10.72), to 6 months in a single trial using the CARES-SF (MD -3.30; 95% CI -5.54 to -1.06), and in two trials looking at follow-up values of the physical condition subscale of the BES (MD 4.41; 95% CI 0.57 to 8.25).

One trial without extractable data reported improvement in physical functioning in both groups (Heim 2007), while another reported no change in either group (Oh 2008).

Role function

We observed no significant effect of exercise on any reported measure of role function (Analysis 12.1). Similar results were reported by Oh 2008 who found no change in role function in either the exercise or control group.

Sexuality

A positive effect of the exercise intervention compared with the control intervention was observed at 6 months (SMD 0.40; 95% CI 0.11 to 0.68) (Analysis 13.1). No trials reported a change in scores at earlier time points. In one trial without extractable data, a reduction in sexual problems was observed for both exercise and control groups but no difference between groups was noted (Berglund 1994).

Sleep disturbance

We observed no significant effect of exercise on any reported measure of sleep disturbance when we looked at the change from baseline to follow-up scores, but did see a significant effect showing improvement when comparing follow-up values by comparison group at 12 weeks (SMD -0.46; 95% CI -0.72 to -0.20), but not at longer follow-up time points (Analysis 14.1). In addition, we

observed an improvement in sleep disturbance when follow-up values were reported using the QLQ-C30 sleep disturbance subscale at 12 weeks (MD -3.11; 95% CI -4.66 to -1.57). Three additional trials reported on sleep disturbances with two not observing any treatment effect on sleep disturbance (Dodd 2010; Oh 2008), but one reporting a large significant effect of exercise on sleep that was not present in the control group (Payne 2008).

Social functioning

Pooling results of trials evaluating change from baseline to follow-up in HRQoL instruments assessing social functioning showed significant improvement following an exercise intervention compared with a control intervention in 386 trial participants at 12 weeks (SMD 0.45; 95% CI 0.02 to 0.87) and in 110 participants at 6 months' follow-up (SMD 0.49; 95% CI 0.11 to 0.87), but this effect was not found at other follow-up periods (Analysis 15.1). A positive treatment effect was still present after excluding results from two studies that included trial participants still receiving treatment (Oh 2010; Targ 2002). At 12 weeks' follow-up, when comparing follow-up scores between the exercise and control groups, a significant effect was seen within the subgroup of breast cancer survivors (SMD 0.42; 95% CI 0.20, 0.64). No treatment effect was observed when comparing change or follow-up scores as assessed in subscales of individual instruments.

One small trial whose data were not extracted reported a significant effect on social functioning with exercise without a corresponding effect in the control group (Oh 2008).

Spirituality

Only one trial reported change in scores from baseline to follow-up in the domain of spirituality and did not find a significant difference (Analysis 16.1). We also observed no treatment effect of exercise when we looked at follow-up scores.

DISCUSSION

Summary of main results

We included 40 trials with a total of 3694 participants randomized to the exercise intervention (n = 1927) or the comparison (n = 1764) groups. Participants enrolled in the trials had various cancer diagnoses including breast, colorectal, head and neck, and other. Thirty trials were conducted among participants who had completed active treatment for their primary or recurrent cancer and 10 trials included participants both during and post cancer treatment. Mode of the exercise intervention differed across trials and included strength training, resistance training, walking, cycling, yoga, Qigong, or Tai Chi. HRQoL and its domains were measured using a wide range of measures.

The results suggest that exercise interventions compared with control interventions have a positive impact on HRQoL and certain HRQoL domains. Exercise interventions resulted in improvement in: global HRQoL at 12 weeks' (SMD 0.48; 95% CI 0.16 to 0.81) and 6 months' (SMD 0.46; 95% CI 0.09 to 0.84) follow-up, breast cancer concerns between 12 weeks' and 6 months' follow-up (SMD 0.99; 95% CI 0.41 to 1.57), body image/self-esteem when assessed using the Rosenberg Self-Esteem scale at 12 weeks' (MD 4.50; 95% CI 3.40 to 5.60) and between 12 weeks' and 6 months (MD 2.70; 95% CI 0.73 to 4.67) follow-up, emotional well-being at 12 weeks' follow-up (SMD 0.33; 95% CI 0.05 to 0.61), sexuality at 6 months' follow-up (SMD 0.40; 95% CI 0.11 to 0.68), sleep disturbance when comparing follow-up values by comparison group at 12 weeks' follow-up (SMD -0.46; 95% CI -0.72 to -0.20), and social functioning at 12 weeks (SMD 0.45; 95% CI 0.02 to 0.87) and 6 months' follow-up (SMD 0.49; 95% CI 0.11 to 0.87).

Further, exercise interventions resulted in decrease in anxiety at 12 weeks' follow-up (SMD -0.26; 95% CI -0.07 to -0.44), fatigue at 12 weeks' follow-up (SMD -0.82; 95% CI -1.50 to -0.14) and between 12 weeks' and 6 months' follow-up (SMD -0.42; 95% CI -0.02 to -0.83), and pain at 12 weeks' follow-up (MD -7.06; 95% CI -13.91 to -0.21) when comparing follow-up values by comparison group.

There were positive trends and impact of exercise intervention for depression and body image (when analyzing combined instruments); however, because only a few studies measured these outcomes, the robustness of findings is uncertain. We observed no effect of the exercise interventions on HRQoL domains of cognitive function, physical functioning, general health perspective, role function, and spirituality.

The positive results must be interpreted with caution owing to the heterogeneity of exercise programs tested and measures used to assess HRQoL and HRQoL domains, and the risk of bias in many trials. Further research is required to investigate how to sustain positive effects of exercise over time and to determine essential attributes of exercise (mode, intensity, frequency, duration, timing) by cancer type and cancer treatment for optimal effects on HRQoL and its domains.

The Summary of findings 1 provides a summary of the main results with associated risks.

Overall completeness and applicability of evidence

This systematic review included 40 trials, 38 of which were RCTs and two were CCTs. These trials allocated 3694 participants to either the exercise or comparison groups. Participants enrolled in the trials had various cancer diagnoses including breast, colorectal, head and neck, and other. All trials included participants who had completed active cancer treatment; however, some trials also included participants who were currently undergoing treatment. Exercise interventions tested in the trials varied greatly and included strength training, resistance training, yoga, walking, cycling, Tai Chi, and Qigong. HRQoL and HRQoL domains were assessed using a wide range of measures. The Characteristics of included studies table provides detailed information on the trial attributes.

The review draws upon worldwide studies. The comprehensive search strategy obtained information from several electronic databases, citations through Web of Science and Scopus, PubMed's related article feature, and several websites; and review of reference list of other reviews in the field and reference list of all included trials. There were no language or date restrictions in the search strategy. See Search methods for identification of studies for details.

In terms of applicability of evidence, the majority of trials were conducted among women who were breast cancer survivors. Further, many trials did not provide socio-demographic information of participants (race/ethnicity, education level, employment status, annual income, social and health benefits) that would enable comparisons between trials and assess

generalizability of findings. Based on socio-demographic data presented in trials, participants were generally white people and with more than high school level education. These characteristics would limit applicability of evidence to a broader cancer survivor population. Further, the majority of trials measured effects of the intervention at the end of the intervention. Thus it is unclear about how sustainable the positive effects of the intervention would be.

The exercise programs varied greatly in their mode, frequency, duration, and intensity. These variations and the lack of understanding about important elements of exercise programs (mode, frequency, duration of sessions and programs, and intensity) for optimal effects on HRQoL and HRQoL domains would preclude informed decision-making in clinical settings and limit applicability of findings.

The HRQoL and HRQoL domains were assessed using a diverse range of instruments with varying psychometric properties. Further, reliance on self-report measures, without triangulation of findings with objectively measured outcomes, can open interpretation of findings to bias.

Because of the variability across interventions, outcome measures, and follow-up times, we looked for treatment effects that were consistent across time and across the different instruments used to assess a specific domain. Although we found some significant effects, they tended to be in subgroups or only at one time point, undermining our confidence in the observed effect. When we observed a significant effect, it was usually at the 12 weeks' follow-up period, which typically equates to the end of the intervention. We frequently found that a positive effect at 12 weeks was not observed at later time periods (i.e. improvement in global HRQoL, reduction in anxiety, etc.), but it is unclear if this finding is because of lack of effect of the exercise intervention at later times, or because there were so few studies measuring outcomes at longer times of follow-up.

The trials provided no data on cost or cost-effectiveness of exercise program on HRQoL and HRQoL domains among cancer survivors.

Quality of the evidence

Results of the review need to be interpreted with caution owing to the risk of bias. All the trials reviewed were at high risk for performance bias because blinding of participants is not possible in exercise intervention unless more rigorously controlled comparative designs are utilized to test the effects of exercise interventions. Performance bias becomes accentuated in trials where participants are asked to provide subjective assessments of outcomes such as HRQoL and HRQoL domains. In addition, the majority of trials were at high risk for detection bias as the outcome assessors were not blinded, were at high risk for attrition bias owing to inadequate handling of incomplete data, and were at high or unclear risk for selection bias owing to inadequate concealment of allocation to the intervention.

The Summary of findings 1 provides a summary on the quality of evidence.

Potential biases in the review process

The strength of this review is the comprehensive search strategy that included a search of 11 electronic databases, citations through Web of Science and Scopus, PubMed's related article feature, and several websites; and, review of reference lists of other reviews in the field and reference lists of all included trials. The comprehensive search strategy was designed and implemented to ensure the identification and retrieval of the maximum number of available published trials and trials in the gray literature. The search strategy also ensured no language restrictions. Trials published in non-English language were assessed for eligibility and, if eligible, had data abstracted by native speakers of the language in which the trial was published.

In spite of such a comprehensive search, it is still possible that this review may have a publication bias. We prepared two funnel plots to assess publication bias for change in global QoL from baseline to follow-up (Figure 4) and for follow-up values for global QoL (Figure 5). Visually both figures showed some slight asymmetry indicating that there is some publication bias in this area of research. We did not complete funnel plots for the other outcomes, because too few studies contributed to the outcome measures. It is possible this review missed some potentially eligible trials in the gray literature, but given the trial results, it is unclear whether the addition of trials only in the gray literature would have a significant impact on results of the review if, as been suggested, trials reported only in the gray literature includes trials have small sample sizes and inconclusive results (McAuley 2000). Further, we corresponded with and requested additional data from 11 trial authors (Daley 2007a; Dodd 2010; Heim 2007; Herrero 2006; Mehnert 2011; Oh 2008; Payne 2008; Penttinen 2011; Rogers 2009; Tang 2010; Thorsen 2005), and six of these trials authors were able to provide additional data. We were unable to find the correct corresponding address for one author (Berglund 1994). Obtaining additional data allowed inclusion of these trials in the quantitative meta-analyses, which made the analyses and findings more robust and complete.

Figure 4. Funnel plot of comparison: 1 Health-related quality of life, outcome: 1.1 Overall quality of life change score.

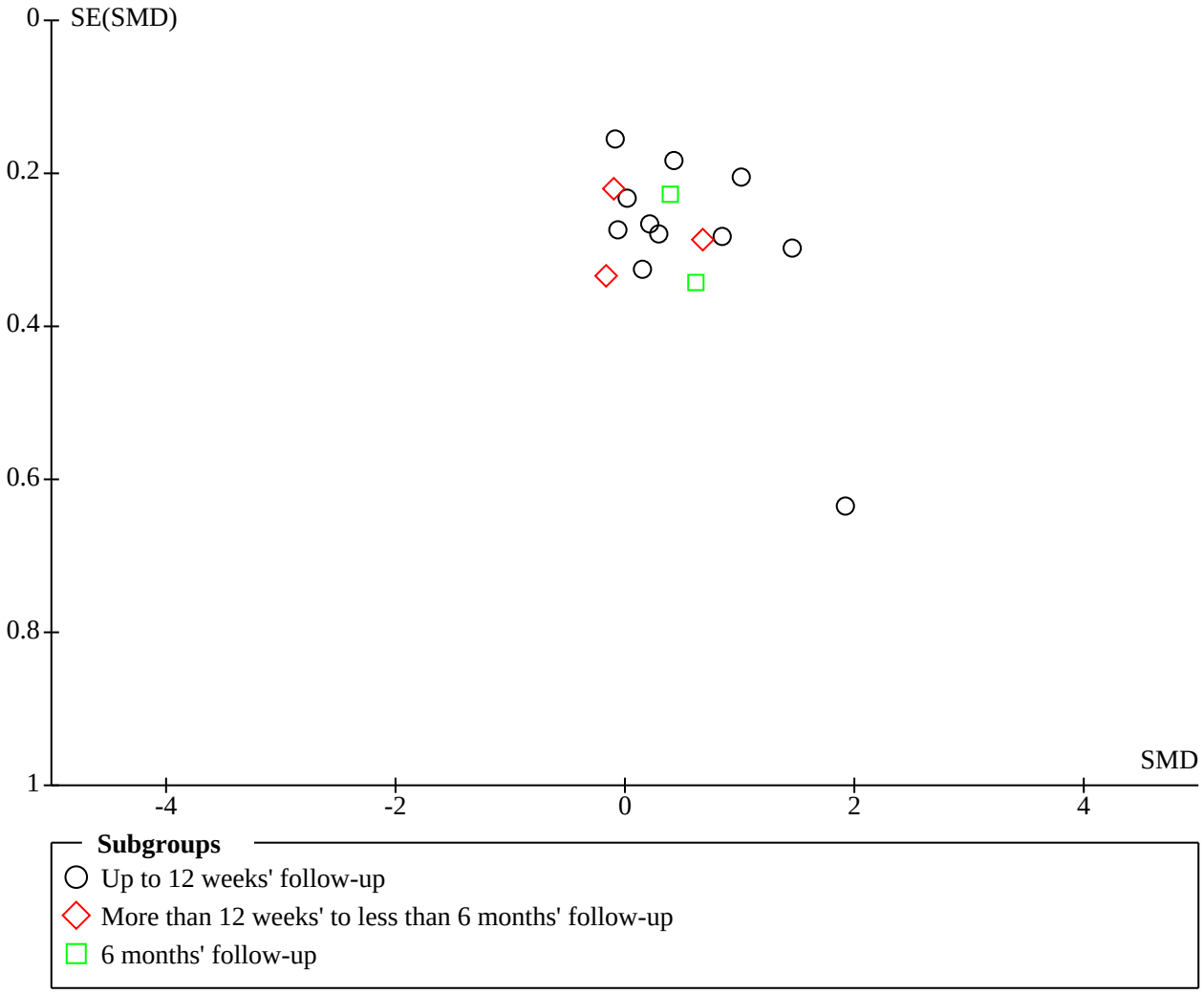
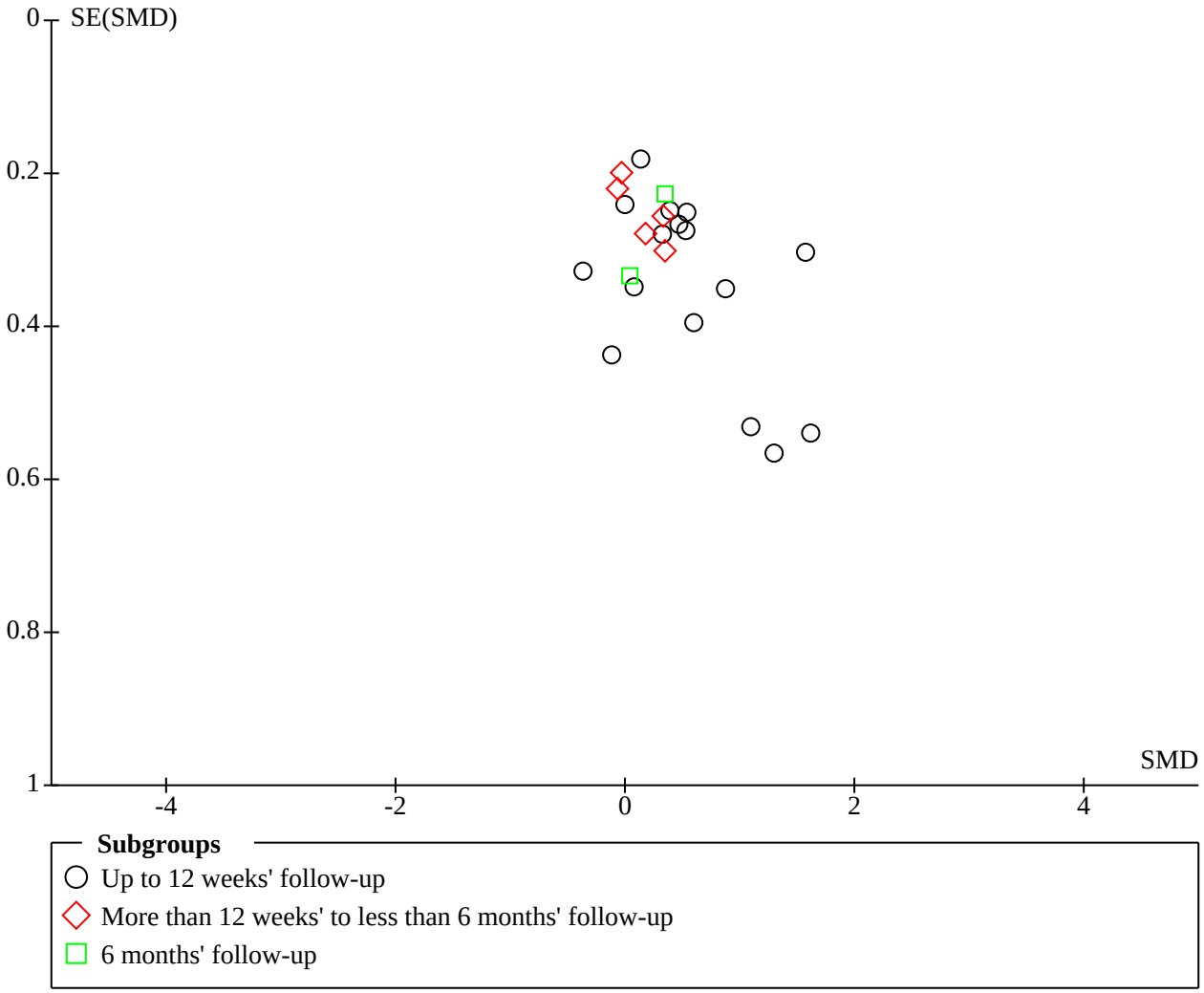


Figure 5. Funnel plot of comparison: 1 Health-related quality of life, outcome: 1.2 Overall quality of life values.



Agreements and disagreements with other studies or reviews

Some systematic reviews have evaluated the effectiveness of exercise interventions on HRQoL or HRQoL domains (Craft 2011; Cramp 2008; Cramp 2010; Duijts 2011; Ferrer 2011; Speck 2010a). All of these reviews included people both during and after active cancer and only one review (Speck 2010a) presented findings by treatment status. Cramp 2008 examined the effect of exercise on cancer-related fatigue and reported that overall exercise was beneficial in the management of cancer-related fatigue and that exercise was beneficial in the management of cancer-related fatigue among breast cancer survivors, a finding that is similar to what we report here. In another review, Cramp 2010, based on a review of the effect of resistance (strength) training on HRQoL, reported no significant benefit of resistance training on global HRQoL and on anxiety and depression. In contrast we found an effect of exercise on global HRQoL, but not on depression. Two of the four trials reviewed by Cramp 2010 reported a significant improvement in fatigue. Duijts 2011 evaluated the effects of exercise (and behavioral) interventions on fatigue, depression, body image, stress, and HRQoL in breast cancer survivors both during and after cancer treatment. Physical exercise interventions had moderate statistically significant effects for fatigue, depression, body image, and HRQoL, but the effect on anxiety, although in the expected direction, did not reach statistical significance. Again, these findings are generally consistent with those presented here although we did not find an effect on depression. Ferrer 2011, based on a meta-analysis of the efficacy of exercise interventions in improving HRQoL in cancer survivors during and after cancer treatment, documented increased HRQoL scores but the effect was more pronounced for interventions that had intense aerobic exercises and that targeted women. Speck 2010a evaluated the effects of physical activity across the cancer control continuum (including during and after cancer treatment). Physical activity interventions among people who had completed active treatment for their cancer had moderate effects on fatigue- and breast-cancer-specific concerns and had small to moderate effects on overall HRQoL. Craft 2011 reviewed the effects of exercise on depression and documented that exercise had a modest positive effect on depressive symptoms. Our review did not find a consistent effect on depression in contrast to these reviews,

but our findings are congruent with respect to global HRQoL and fatigue observed by other reviewers. These differences may be owing to differences in the trial population in that our review only included individuals who had completed active cancer treatment.

AUTHORS' CONCLUSIONS

Implications for practice

This systematic review finds that exercise interventions may have beneficial effects on overall HRQoL and HRQoL domains including cancer-specific concerns (e.g. breast cancer), body image/self-esteem, emotional well-being, sexuality, sleep disturbance, social functioning, anxiety, fatigue, and pain at varying follow-up periods among cancer survivors who are beyond active treatment for their primary or recurrent cancer. Exercise programs could be considered as an integral component for the management of HRQoL among cancer survivors.

The positive results must be interpreted cautiously owing to the heterogeneity of mode of exercise programs, measures used to assess HRQoL and HRQoL domains, and the risk of bias in many trials. Further, a lack of understanding about important elements of exercise programs (mode, frequency, duration of sessions and programs, and intensity) for optimal effects on HRQoL and HRQoL domains would preclude informed decision making in clinical settings and limit practical applicability of findings.

No evidence of effect was found for HRQoL domains such as cognitive function, physical functioning, general health perspective, role function, and spirituality. The lack of evidence may be due to few trials assessing these outcomes, small number of participants in trials measuring these outcomes, and substantial heterogeneity between trials measuring these outcomes on the exercise programs implemented and measures used to assess the outcomes. Owing to these limitations, no conclusions can be drawn at this time regarding the effects of exercise interventions on these HRQoL domains.

From a practice perspective, it would be important to understand whether certain exercise attributes have more or less optimal effects on HRQoL and HRQoL domains among survivors of certain types of cancers. Further, it would be important to understand which mode of exercise program (strength; resistance; Tai Chi; yoga; and aerobic, anaerobic, or a combination) coupled with what levels of essential attributes (frequency of program, duration of program and each session) is optimal for which cancer type and cancer treatment.

Implications for research

This systematic review and meta-analysis of 40 trials on the effects of exercise on HRQoL and HRQoL domains for cancer survivors provides evidence that exercise interventions may have beneficial effects on overall HRQoL and HRQoL domains including cancer-specific concerns (e.g. breast cancer), body image/self-esteem, emotional well-being, sexuality, sleep disturbance, social functioning, anxiety, fatigue, and pain at varying follow-up periods among cancer survivors who are beyond active treatment for their primary or recurrent cancer. Further, findings of this review suggest that exercise interventions may have minimal or no effects on HRQoL domains such as cognitive function, physical functioning, general health perspective, role function, and spirituality among cancer survivors.

Further research is required to investigate whether the effect of an exercise intervention can be maintained beyond the active intervention period, and if so, how to sustain changes in exercise behaviors and positive effects of exercise on HRQoL and HRQoL domains. Empirical evidence is also needed to determine the optimal follow-up period from end of the intervention. To further this understanding, rigorous RCTs could include qualitative research components in trials to benefit from the contextually rich insights gained from engaging participants about their experiences in exercise interventions.

More research is needed to determine essential attributes of exercise (mode, intensity, frequency, duration, timing) by cancer type and cancer treatment for optimal effects on HRQoL and its domains.

HRQoL and HRQoL domains are important measures of cancer survivorship. However, the heterogeneous range of measures used to assess HRQoL and HRQoL domains, make comparisons of findings between trials extremely difficult. Efforts such as the Patient-Reported Outcomes Measurement Information System (PROMIS) may help address these issues (Cella 2010; National Cancer Institute 2012).

ACKNOWLEDGEMENTS

The authors would like to thank and acknowledge Clare Jess (Managing Editor) and Editorial Base of the Cochrane Gynecological Cancer Review Group for their help and editorial advice during the preparation of the review. Moreover, the authors would like to thank the peer reviewers and consumer reviewer for their invaluable feedback during the peer review process, and the authors of primary trials for additional information about their trials. This research was supported in part by the National Institute for Health Research (NIHR) Health Technology Assessment program. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the funding agency.

REFERENCES

References to studies included in this review

Bai 2004 *{published data only}*

* Bai S-M, Ma C, Liu Y-M, Xue W-P, Luo M, Ou Z-H. Effects of cognitive behavior intervention and cinesiateics on the quality of life of patients with nasopharyngeal carcinoma after radiotherapy. *Chinese Journal of Clinical Rehabilitation* 2004;**8**(29):6312-3.

Banasik 2011 *{published data only}*

* Banasik J, Williams H, Haberman M, Blank SE, Bendel R. Effect of Iyengar yoga practice on fatigue and diurnal salivary cortisol concentration in breast cancer survivors. *Journal of the American Academy of Nurse Practitioners* 2011;**23**(3):135-42.

Berglund 1994 *{published data only}*

* Berglund G, Bolund C, Gustafsson UL, Sjoden PO. One-year follow-up of the 'Starting Again' group rehabilitation programme for cancer patients. *European Journal of Cancer* 1994;**30A**(12):1744-51.

Bourke 2011 *{published data only}*

* Bourke L, Thompson G, Gibson D J, Daley A, Crank H, Adam I, et al. Pragmatic lifestyle intervention in patients recovering from colon cancer: a randomized controlled pilot study. *Archives of Physical Medicine and Rehabilitation* 2011;**92**(5):749-5.

Burnham 2002 *{published data only}*

* Burnham TR, Wilcox A. Effects of exercise on physiological and psychological variables in cancer survivors. *Medicine and Science in Sports and Exercise* 2002;**34**(12):1863-7.

Cadmus 2009 *{published data only}*

* Cadmus LA, Salovey P, Yu H, Chung G, Kasl S, Irwin ML. Exercise and quality of life during and after treatment for breast cancer: results of two randomized controlled trials. *Psychooncology* 2009;**18**(4):343-52.

Cadmus LA. Exercise and quality of life during and after treatment for breast cancer. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2008;**68**:12-B.

Irwin ML, Varma K, Alvarez-Reeves M, Cadmus L, Wiley A, Chung GG, et al. Randomized controlled trial of aerobic exercise on insulin and insulin-like growth factors in breast cancer survivors: the Yale Exercise and Survivorship study. *Cancer Epidemiology, Biomarkers & Prevention* 2009;**18**(1):306-13.

Latka RN, Alvarez-Reeves M, Cadmus L, Irwin ML. Adherence to a randomized controlled trial of aerobic exercise in breast cancer survivors: the Yale exercise and survivorship study. *Journal of Cancer Survivorship* 2009;**3**(3):148-57.

Cho 2006 *{published data only}*

* Cho OH, Yoo YS, Kim NC. Efficacy of comprehensive group rehabilitation for women with early breast cancer in South Korea. *Nursing and Health Sciences* 2006;**8**(3):140-6.

Cohen 2004 *{published data only}*

* Cohen L, Warneke C, Fouladi RT, Rodriguez MA, Chaoul-Reich A. Psychological adjustment and sleep quality in a randomized trial of the effects of a Tibetan yoga intervention in patients with lymphoma. *Cancer* 2004;**100**(10):2253-60.

Courneya 2003a *{published data only}*

* Courneya KS, Friedenreich CM, Sela RA, Quinney HA, Rhodes RE, Handman M. The group psychotherapy and home-based physical exercise (group-hope) trial in cancer survivors: physical fitness and quality of life outcomes. *Psychooncology* 2003;**12**(4):357-74.

Courneya KS, Friedenreich CM, Sela RA, Quinney HA, Rhodes RE. Correlates of adherence and contamination in a randomized controlled trial of exercise in cancer survivors: an application of the theory of planned behavior and the five factor model of personality. *Annals of Behavioral Medicine* 2002;**24**(4):257-68.

Courneya 2003b *{published data only}*

* Courneya KS, Friedenreich CM, Quinney HA, Fields AL, Jones LW, Fairey AS. A randomized trial of exercise and quality of life in colorectal cancer survivors. *European Journal of Cancer Control* 2003;**12**(4):347-57.

Courneya KS, Quinney HA. Baseline relationships between exercise and quality of life in colorectal cancer patients participating in a randomized exercise trial. *Journal of Sport and Exercise Psychology* 2001;**2**(Suppl. 2):S40.

Courneya 2003c *{published data only}*

* Courneya KS, Mackey JR, Bell GJ, Jones LW, Field CJ, Fairey AS. Randomized controlled trial of exercise training in postmenopausal breast cancer survivors: cardiopulmonary and quality of life outcomes. *Journal of Clinical Oncology* 2003;**21**(9):1660-8.

Fairey AS, Courneya KS, Field CJ, Bell GJ, Jones LW, Mackey JR. Effects of exercise training on fasting insulin, insulin resistance, insulin-like growth factors, and insulin-like growth factor binding proteins in postmenopausal breast cancer survivors: a randomized controlled trial. *Cancer Epidemiology, Biomarkers & Prevention* 2003;**12**(8):721-7.

Courneya 2009 *{published data only}*

Courneya KS, Sellar CM, Stevinson C, McNeely ML, Friedenreich CM, Peddle CJ, et al. Moderator effects in a randomized controlled trial of exercise training in lymphoma patients. *Cancer Epidemiology, Biomarkers & Prevention* 2009;**18**(10):2600-7.

* Courneya KS, Sellar CM, Stevinson C, McNeely ML, Peddle CJ, Friedenreich CM, et al. Randomized controlled trial of the effects of aerobic exercise on physical functioning and quality of life in lymphoma patients. *Journal of Clinical Oncology* 2009;**27**(27):4605-12.

Courneya KS, Stevinson C, McNeely ML, Sellar CM, Peddle CJ, Friedenreich CM, et al. Predictors of adherence to supervised

exercise in lymphoma patients participating in a randomized controlled trial. *Annals of Behavioral Medicine* 2010;**40**(1):30-9.

Culos-Reed 2006 *{published data only}*

* Culos-Reed SN, Carlson LE, Daroux LM, Hately-Aldous S. A pilot study of yoga for breast cancer survivors: physical and psychological benefits. *Psychooncology* 2006;**15**(10):891-7.

Daley 2007a *{published data only}*

* Daley AJ, Crank H, Saxton JM, Mutrie N, Coleman R, Roalfe A. Randomized trial of exercise therapy in women treated for breast cancer. *Journal of Clinical Oncology* 2007;**25**(13):1713-21.

Daley AJ, Mutrie N, Crank H, Coleman R, Saxton J. Exercise therapy in women who have had breast cancer: design of the Sheffield women's exercise and well-being project. *Health Education Research* 2004;**19**(6):686-97.

Danhauer 2009 *{published data only}*

* Danhauer SC, Mihalko SL, Russell GB, Campbell CR, Felder L, Daley K, et al. Restorative yoga for women with breast cancer: findings from a randomized pilot study. *Psychooncology* 2009;**18**(4):360-8.

Dimeo 2004 *{published data only}*

* Dimeo FC, Thomas F, Raabe-Menssen C, Propper F, Mathias M. Effect of aerobic exercise and relaxation training on fatigue and physical performance of cancer patients after surgery. A randomised controlled trial. *Supportive Care in Cancer* 2004;**12**(11):774-9.

Dodd 2010 *{published data only}*

* Dodd MJ, Cho M H, Miaskowski C, Painter PL, Paul SM, Cooper BA, et al. A randomized controlled trial of home-based exercise for cancer-related fatigue in women during and after chemotherapy with or without radiation therapy. *Cancer Nursing* 2010;**33**(4):245-57.

Donnelly 2011 *{published data only}*

* Donnelly CM, Blaney JM, Lowe-Strong A, Rankin JP, Campbell A, McCrum-Gardner E, et al. A randomised controlled trial testing the feasibility and efficacy of a physical activity behavioural change intervention in managing fatigue with gynaecological cancer survivors. *Gynecologic Oncology* 2011;**122**(3):618-24.

Fillion 2008 *{published data only}*

* Fillion L, Gagnon P, Leblond F, Gelinas C, Savard J, Dupuis R, et al. A brief intervention for fatigue management in breast cancer survivors. *Cancer Nursing* 2008;**31**(2):145-59.

Heim 2007 *{published data only}*

* Heim ME, v d Malsburg ML, Niklas A. Randomized controlled trial of a structured training program in breast cancer patients with tumor-related chronic fatigue. *Onkologie* 2007;**30**(8-9):429-34.

Herrero 2006 *{published data only}*

* Herrero F, San Juan AF, Fleck SJ, Balmer J, Perez M, Canete S, et al. Combined aerobic and resistance training in breast cancer survivors: a randomized, controlled pilot trial. *International Journal of Sports Medicine* 2006;**27**(7):573-80.

Knols 2011 *{published data only}*

* Knols RH, de Bruin ED, Uebelhart D, Aufdemkampe G, Schanz U, Stenner-Liewen F, et al. Effects of an outpatient physical exercise program on hematopoietic stem-cell transplantation recipients: a randomized clinical trial. *Bone Marrow Transplantation* 2011;**46**(9):1245-55.

McNeely 2008a *{published data only}*

* McNeely ML, Parliament MB, Seikaly H, Jha N, Magee DJ, Haykowsky MJ, et al. Effect of exercise on upper extremity pain and dysfunction in head and neck cancer survivors: a randomized controlled trial. *Cancer* 2008;**113**(1):214-22.

McNeely ML. Exercise rehabilitation for breast and head and neck cancer survivors. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2008;**68**(10-B):6627.

Mehnert 2011 *{published data only}*

* Mehnert A, Veers S, Howaldt D, Braumann KM, Koch U, Schulz KH. Effects of a physical exercise rehabilitation group program on anxiety, depression, body image, and health-related quality of life among breast cancer patients. *Onkologie* 2011;**34**(5):248-53.

Milne 2008a *{published data only}*

* Milne HM, Wallman KE, Gordon S, Courneya KS. Effects of a combined aerobic and resistance exercise program in breast cancer survivors: a randomized controlled trial. *Breast Cancer Research and Treatment* 2008;**108**(2):279-88.

Moadel 2007 *{published data only}*

* Moadel AB, Shah C, Wylie-Rosett J, Harris MS, Patel SR, Hall CB, et al. Randomized controlled trial of yoga among a multiethnic sample of breast cancer patients: effects on quality of life. *Journal of Clinical Oncology* 2007;**25**(28):4387-95.

Mustian 2004 *{published data only}*

Janelsins MC, Davis PG, Wideman L, Katula JA, Sprod LK, Peppone LJ, et al. Effects of Tai Chi Chuan on insulin and cytokine levels in a randomized controlled pilot study on breast cancer survivors. *Clinical Breast Cancer* 2011;**11**(3):161-70.

* Mustian KM, Katula JA, Gill DL, Roscoe JA, Lang D, Murphy K. Tai Chi Chuan, health-related quality of life and self-esteem: a randomized trial with breast cancer survivors. *Supportive Care in Cancer* 2004;**12**(12):871-6.

Mustian KM, Palesh OG, Flecksteiner SA. Tai Chi Chuan for breast cancer survivors. *Medicine and Sport Science* 2008;**52**:209-17.

Oh 2008 *{published data only}*

* Oh B, Butow P, Mullan B. Medical Qigong for cancer patients: pilot study of impact on quality of life, side effects of treatment and inflammation. *American Journal of Chinese Medicine* 2008;**36**(3):459-72.

Oh 2010 *{published data only}*

Oh B, Butow P, Mullan AB, Clarke SJ, Beale PJ, Pavlakis N, et al. Effect of Qigong on cognitive function, quality of life and a

biomarker of inflammation in cancer patients: a randomized controlled trial. *Support Care in Cancer* 2012;**20**(6):1235-42. [DOI: 10 1007/s00520-011-1209-6]

* Oh B, Butow P, Mullan B, Clarke S, Beale P, Pavlakis N, et al. Impact of medical Qigong on quality of life, fatigue, mood and inflammation in cancer patients: a randomized controlled trial. *Annals of Oncology* 2010;**21**(3):608-14.

Oh B, Butow P, Mullan B, Clarke S, Beale P, Pavlakis N, et al. Randomized clinical trial of medical Qigong on quality of life, fatigue, side effects, mood status, and inflammation of cancer patients. *Journal of Clinical Oncology* 2009;**27**(15):Suppl 1.

Oh B, Butow PN, Clarke J, Beale PJ, Pavlakis N. Impact of medical qigong on quality of life, fatigue, symptoms, mood and inflammation in cancer patient: a randomized controlled trial. *Asia-Pacific Journal of Clinical Oncology* 2009;**5**(Suppl. 2):A237.

Ohira 2006 *{published data only}*

Ahmed RL, Thomas W, Yee D, Schmitz KH. Randomized controlled trial of weight training and lymphedema in breast cancer survivors. *Journal of Clinical Oncology* 2006;**24**(18):2765-72.

* Ohira T, Schmitz KH, Ahmed RL, Yee D. Effects of weight training on quality of life in recent breast cancer survivors: the Weight Training for Breast Cancer Survivors (WTBS) study. *Cancer* 2006;**106**(9):2076-83.

Schmitz KH, Ahmed RL, Hannan PJ, Yee D. Safety and efficacy of weight training in recent breast cancer survivors to alter body composition, insulin, and insulin-like growth factor axis proteins. *Cancer Epidemiology, Biomarkers, and Prevention* 2005;**14**(7):1672-80.

Payne 2008 *{published data only}*

* Payne JK, Held J, Thorpe J, Shaw H. Effect of exercise on biomarkers, fatigue, sleep disturbances, and depressive symptoms in older women with breast cancer receiving hormonal therapy. *Oncology Nursing Forum* 2008;**35**(4):635-42.

Penttinen 2011 *{published data only}*

Penttinen H, Nikander R, Blomqvist C, Luoto R, Saarto T. Recruitment of breast cancer survivors into a 12-month supervised exercise intervention is feasible. *Contemporary Clinical Trials* 2009;**30**(5):457-63.

* Penttinen HM, Saarto T, Kellokumpu-Lehtinen P, Blomqvist C, Huovinen R, Kautiainen H, et al. Quality of life and physical performance and activity of breast cancer patients after adjuvant treatments. *Psychooncology* 2011;**20**(11):1211-20.

Saarto T, Sievanen H, Kellokumpu-Lehtinen P, Nikander R, Vehmanen L, Huovinen R, et al. Effect of supervised and home exercise training on bone mineral density among breast cancer patients. A 12-month randomised controlled trial. *Osteoporosis International* 2011;**23**(5):1601-12. [DOI: 10.1007/s00198-011-1761-4]

Pinto 2003 *{published data only}*

* Pinto BM, Clark MM, Maruyama NC, Feder SI. Psychological and fitness changes associated with exercise participation among women with breast cancer. *Psychooncology* 2003;**12**(2):118-26.

Pinto 2005 *{published data only}*

* Pinto BM, Frierson GM, Rabin C, Trunzo JJ, Marcus BH. Home-based physical activity intervention for breast cancer patients. *Journal of Clinical Oncology* 2005;**23**(15):3577-87.

Pinto BM, Rabin C, Papandonatos GD, Frierson GM, Trunzo JJ, Marcus BH. Maintenance of effects of a home-based physical activity program among breast cancer survivors. *Supportive Care in Cancer* 2008;**16**(11):1279-89.

Rogers 2009 *{published data only}*

Rogers LQ, Hopkins-Price P, Vicari S, Markwell S, Pamenter R, Courneya KS, et al. Physical activity and health outcomes three months after completing a physical activity behavior change intervention: persistent and delayed effects. *Cancer Epidemiology Biomarkers and Prevention* 2009;**18**(5):1410-8.

* Rogers LQ, Hopkins-Price P, Vicari S, Pamenter R, Courneya KS, Markwell S, et al. A randomized trial to increase physical activity in breast cancer survivors. *Medicine and Science in Sports and Exercise* 2009;**41**(4):935-46.

Rogers LQ, Markwell S, Hopkins-Price P, Vicari S, Courneya KS, Hoelzer K, et al. Reduced barriers mediated physical activity maintenance among breast cancer survivors. *Journal of Sport & Exercise Psychology* 2011;**33**(2):235-54.

Segar 1998 *{published data only}*

* Segar ML, Katch VL, Roth RS, Garcia AW, Portner TI, Glickman SG, et al. The effect of aerobic exercise on self-esteem and depressive and anxiety symptoms among breast cancer survivors. *Oncology Nursing Forum* 1998;**25**(1):107-13.

Speck 2010 *{published data only}*

Rogerino A, Grant LL, Wilcox H 3rd, Schmitz KH. Geographic recruitment of breast cancer survivors into community-based exercise interventions. *Medicine and Science in Sports and Exercise* 2009;**41**(7):1413-20.

Schmitz KH, Troxel AB, Cheville A, Grant LL, Bryan CJ, Gross CR, et al. Physical Activity and Lymphedema (the PAL trial): assessing the safety of progressive strength training in breast cancer survivors. *Contemporary Clinical Trials* 2009;**30**(3):233-45.

* Speck RM, Gross CR, Hormes JM, Ahmed RL, Lytle LA, Hwang WT, et al. Changes in the Body Image and Relationship Scale following a one-year strength training trial for breast cancer survivors with or at risk for lymphedema. *Breast Cancer Research and Treatment* 2010;**121**(2):421-30.

Tang 2010 *{published data only}*

* Tang MF, Liou TH, Lin CC. Improving sleep quality for cancer patients: benefits of a home-based exercise intervention. *Supportive Care in Cancer* 2010;**18**(10):1329-39.

Targ 2002 *{published data only}*

* Targ EF, Levine EG. The efficacy of a mind-body-spirit group for women with breast cancer: a randomized controlled trial. *General Hospital Psychiatry* 2002;**24**(4):238-48.

Thorsen 2005 *{published data only}*

* Thorsen L, Skovlund E, Stromme SB, Hornslien K, Dahl AA, Fossa SD. Effectiveness of physical activity on cardiorespiratory fitness and health-related quality of life in young and middle-aged cancer patients shortly after chemotherapy. *Journal of Clinical Oncology* 2005;**23**(10):2378-88.

References to studies excluded from this review

Basen-Engquist 2006 *{published data only}*

Basen-Engquist K, Taylor CL, Rosenblum C, Smith MA, Shinn EH, Greisinger A, et al. Randomized pilot test of a lifestyle physical activity intervention for breast cancer survivors. *Patient Education and Counseling* 2006;**64**(1-3):225-34.

Beurskens 2007 *{published data only}*

Beurskens CH, van Uden CJ, Strobbe LJ, Oostendorp RA, Wobbes T. The efficacy of physiotherapy upon shoulder function following axillary dissection in breast cancer, a randomized controlled study. *BMC Cancer* 2007;**7**:166.

Blanchard 2001 *{published data only}*

Blanchard CM, Courneya KS, Laing D. Effects of acute exercise on state anxiety in breast cancer survivors. *Oncology Nursing Forum* 2001;**28**(10):1617-21.

Bloom 2008 *{published data only}*

Bloom JR, Stewart SL, D'Onofrio CN, Luce J, Banks PJ. Addressing the needs of young breast cancer survivors at the 5 year milestone: can a short-term, low intensity intervention produce change? *Journal of Cancer Survivorship* 2008;**2**(3):190-204.

Braam 2010 *{published data only}*

Braam KI, van Dijk EM, Veening MA, Bierings MB, Merks JH, Grootenhuis MA, et al. Design of the Quality of Life in Motion (QLIM) study: a randomized controlled trial to evaluate the effectiveness and cost-effectiveness of a combined physical exercise and psychosocial training program to improve physical fitness in children with cancer. *BMC Cancer* 2010;**10**:624.

Carmack Taylor 2004 *{published data only}*

Carmack Taylor CL, Smith MA, de Moor C, Dunn AL, Pettaway C, Sellin R, et al. Quality of life intervention for prostate cancer patients: design and baseline characteristics of the active for life after cancer trial. *Controlled Clinical Trials* 2004;**25**(3):265-85.

Carmack Taylor 2006 *{published data only}*

Carmack Taylor CL, Demoor C, Smith MA, Dunn AL, Basen-Engquist K, Nielsen I, et al. Active for Life After Cancer: a randomized trial examining a lifestyle physical activity program for prostate cancer patients. *Psychooncology* 2006;**15**(10):847-62.

Carmack Taylor 2007 *{published data only}*

Carmack Taylor CL, de Moor C, Basen-Engquist K, Smith MA, Dunn AL, Badr H, et al. Moderator analyses of participants in the Active for Life after cancer trial: implications for physical activity group intervention studies. *Annals of Behavioral Medicine* 2007;**33**(1):99-104.

Carson 2009 *{published data only}*

Carson JW, Carson KM, Porter LS, Keefe FJ, Seewaldt VL. Yoga of Awareness program for menopausal symptoms in breast cancer survivors: results from a randomized trial. *Supportive Care in Cancer* 2009;**17**(10):1301-9.

Cheema 2006 *{published data only}*

Cheema BS, Gaul CA. Full-body exercise training improves fitness and quality of life in survivors of breast cancer. *Journal of Strength and Conditioning Research* 2006;**20**(1):14-21.

Cheung 2003 *{published data only}*

Cheung YL, Molassiotis A, Chang AM. The effect of progressive muscle relaxation training on anxiety and quality of life after stoma surgery in colorectal cancer patients. *Psychooncology* 2003;**12**(3):254-66.

Cinar 2008 *{published data only}*

Cinar N, Seckin U, Keskin D, Bodur H, Bozkurt B, Cengiz O. The effectiveness of early rehabilitation in patients with modified radical mastectomy. *Cancer Nursing* 2008;**31**(2):160-5.

Courneya 2004a *{published data only}*

Courneya KS, Friedenreich CM, Sela RA, Quinney HA, Rhodes RE, Jones LW. Exercise motivation and adherence in cancer survivors after participation in a randomized controlled trial: an attribution theory perspective. *International Journal of Behavioral Medicine* 2004;**11**(1):8-17.

Courneya 2004b *{published data only}*

Courneya KS, Segal RJ, Reid RD, Jones LW, Malone SC, Venner PM, et al. Three independent factors predicted adherence in a randomized controlled trial of resistance exercise training among prostate cancer survivors. *Journal of Clinical Epidemiology* 2004;**57**(6):571-9.

Courneya 2004c *{published data only}*

Courneya KS, Friedenreich CM, Quinney HA, Fields AL, Jones LW, Fairey AS. Predictors of adherence and contamination in a randomized trial of exercise in colorectal cancer survivors. *Psychooncology* 2004;**13**(12):857-66.

Courneya 2005 *{published data only}*

Courneya KS, Friedenreich CM, Quinney HA, Fields AL, Jones LW, Vallance JK, et al. A longitudinal study of exercise barriers in colorectal cancer survivors participating in a randomized controlled trial. *Annals of Behavioral Medicine* 2005;**29**(2):147-53.

Courneya 2008 *{published data only}*

Courneya KS, Jones LW, Peddle CJ, Sellar CM, Reiman T, Joy AA, et al. Effects of aerobic exercise training in anemic cancer patients receiving darbepoetin alfa: a randomized controlled trial. *Oncologist* 2008;**13**(9):1012-20.

Daley 2007 *{published data only}*

Daley A J, Crank H, Mutrie N, Saxton J M, Coleman R. Determinants of adherence to exercise in women treated for breast cancer. *European Journal of Oncology Nursing* 2007;**11**(5):392-9.

Demark-Wahnefried 2003 *{published data only}*

Demark-Wahnefried W, Clipp E C, McBride C, Lobach DF, Lipkus I, Peterson B, et al. Design of FRESH START: a randomized trial of exercise and diet among cancer survivors. *Medicine and Science in Sports and Exercise* 2003;**35**(3):415-24.

Demark-Wahnefried 2003a *{published data only}*

Demark-Wahnefried W, Morey MC, Clipp EC, Pieper CF, Snyder DC, Sloane R, et al. Leading the Way in Exercise and Diet (Project LEAD): intervening to improve function among older breast and prostate cancer survivors. *Controlled Clinical Trials* 2003;**24**(2):206-23.

Demark-Wahnefried 2006 *{published data only}*

Demark-Wahnefried W, Clipp EC, Morey MC, Pieper CF, Sloane R, Snyder DC, et al. Lifestyle intervention development study to improve physical function in older adults with cancer: outcomes from Project LEAD. *Journal of Clinical Oncology* 2006;**24**(21):3465-73.

Demark-Wahnefried 2007 *{published data only}*

Demark-Wahnefried W, Clipp EC, Lipkus IM, Lobach D, Snyder DC, Sloane R, et al. Main outcomes of the FRESH START trial: a sequentially tailored, diet and exercise mailed print intervention among breast and prostate cancer survivors. *Journal of Clinical Oncology* 2007;**25**(19):2709-18.

Dincer 2007 *{published data only}*

Dincer U, Kaya E, Cakar E, Kiralp MZ, Dursun H. Effectiveness of comprehensive rehabilitation program and home-based exercise in middle and long term mastectomy related disability. *Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi* 2007;**53**(4):138-43.

Dong 2006 *{published data only}*

Dong H-Y, Wang Z-F, Cai L. Correlation between quality of life and rehabilitative guidance education in the postoperative patients with breast cancer. *Chinese Journal of Clinical Rehabilitation* 2006;**10**(42):28-30.

Duijts 2009 *{published data only}*

Duijts SF, Oldenburg HS, van Beurden M, Aaronson NK. Cognitive behavioral therapy and physical exercise for climacteric symptoms in breast cancer patients experiencing treatment-induced menopause: design of a multicenter trial. *BMC Women's Health* 2009;**9**:15.

Elkin 1998 *{published data only}*

Elkin TD, Tyc VL, Hudson M, Crom D. Participation in sports by long-term survivors of childhood cancer. *Journal of Psychosocial Oncology* 1998;**16**(1):63-73.

Elliott 2006 *{published data only}*

Elliott D, McKinley S, Alison JA, Aitken LM, King MT. Study protocol: home-based physical rehabilitation for survivors of a critical illness [ACTRN12605000166673]. *Critical Care* 2006;**10**(3):R90.

Emslie 2007 *{published data only}*

Emslie C, Whyte F, Campbell A, Nutrie N, Lee L, Ritchie D, Kearney N. 'I wouldn't have been interested in just sitting round a table talking about cancer'; exploring the experiences of women with breast cancer in a group exercise trial. *Health Education Research* 2007;**22**(6):827-838.

Fairey 2005 *{published data only}*

Fairey AS, Courneya KS, Field C J, Bell GJ, Jones LW, Mackey JR. Randomized controlled trial of exercise and blood immune function in postmenopausal breast cancer survivors. *Journal of Applied Physiology* 2005;**98**(4):1534-40.

Fairey 2005a *{published data only}*

Fairey AS, Courneya KS, Field CJ, Bell GJ, Jones LW, Martin BS, et al. Effect of exercise training on C-reactive protein in postmenopausal breast cancer survivors: a randomized controlled trial. *Brain, Behavior, and Immunity* 2005;**19**(5):381-8.

Filocamo 2005 *{published data only}*

Filocamo M T, Li Marzi V, Del Popolo G, Cecconi F, Marzocco M, Tosto A, et al. Effectiveness of early pelvic floor rehabilitation treatment for post-prostatectomy incontinence. *European Urology* 2005;**48**(5):734-8.

Gordon 2005 *{published data only}*

Gordon L G, Battistutta D, Scuffham P, Tweeddale M, Newman B. The impact of rehabilitation support services on health-related quality of life for women with breast cancer. *Breast Cancer Research and Treatment* 2005;**93**(3):217-26.

Griffith 2009 *{published data only}*

Griffith K, Wenzel J, Shang J, Thompson C, Stewart K, Mock V. Impact of a walking intervention on cardiorespiratory fitness, self-reported physical function, and pain in patients undergoing treatment for solid tumors. *Cancer* 2009;**115**(20):4874-84.

Hayes 2004 *{published data only}*

Hayes S, Davies PS, Parker T, Bashford J, Newman B. Quality of life changes following peripheral blood stem cell transplantation and participation in a mixed-type, moderate-intensity, exercise program. *Bone Marrow Transplant* 2004;**33**(5):553-8.

Houborg 2006 *{published data only}*

Houborg KB, Jensen MB, Rasmussen P, Gandrup P, Schroll M, Laurberg S. Postoperative physical training following colorectal surgery: a randomised, placebo-controlled study. *Scandinavian Journal of Surgery* 2006;**95**(1):17-22.

Hughes 2004 *{published data only}*

Hughes DC. Physical activity and stress in Hispanic breast cancer survivors. Dissertation Abstracts International: Section B: The Sciences and Engineering 2004;**65**((10-B)).

Hughes 2008 *{published data only}*

Hughes DC, Leung P, Naus MJ. Using single-system analyses to assess the effectiveness of an exercise intervention on quality of life for Hispanic breast cancer survivors: a pilot study. *Social Work in Health Care* 2008;**47**(1):73-91.

Jarden 2009 *{published data only}*

Jarden M, Baadsgaard MT, Hovgaard DJ, Boesen E, Adamsen L. A randomized trial on the effect of a multimodal intervention on physical capacity, functional performance and quality of

life in adult patients undergoing allogeneic SCT. *Bone Marrow Transplantation* 2009;**43**(9):725-37.

Jones 2004a *{published data only}*

Jones LW, Courneya KS, Fairey AS, Mackey JR. Effects of an oncologist's recommendation to exercise on self-reported exercise behavior in newly diagnosed breast cancer survivors: a single-blind, randomized controlled trial. *Annals of Behavioral Medicine* 2004;**28**(2):105-13.

Jones 2008 *{published data only}*

Jones LW, Eves ND, Peterson BL, Garst J, Crawford J, West MJ, et al. Safety and feasibility of aerobic training on cardiopulmonary function and quality of life in postsurgical nonsmall cell lung cancer patients: a pilot study. *Cancer* 2008;**113**(12):3430-9.

Kilbreath 2006 *{published data only}*

Kilbreath SL, Refshauge KM, Beith JM, Ward LC, Simpson JM, Hansen RD. Progressive resistance training and stretching following surgery for breast cancer: study protocol for a randomised controlled trial. *BMC Cancer* 2006;**6**:273.

Kim 2010 *{published data only}*

Kim do S, Sim Y J, Jeong HJ, Kim GC. Effect of active resistive exercise on breast cancer-related lymphedema: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* 2010;**91**(12):1844-8.

Kim 2011 *{published data only}*

Kim SH, Shin MS, Lee HS, Lee ES, Ro JS, Kang HS, et al. Randomized pilot test of a simultaneous stage-matched exercise and diet intervention for breast cancer survivors. *Oncology Nursing Forum* 2011;**38**(2):E97-106.

Kolden 2002 *{published data only}*

Kolden GG, Strauman TJ, Ward A, Kuta J, Woods TE, Schneider KL, et al. A pilot study of group exercise training (GET) for women with primary breast cancer: feasibility and health benefits. *Psychooncology* 2002;**11**(5):447-56.

Korstjens 2008 *{published data only}*

Korstjens I, May AM, van Weert E, Mesters I, Tan F, Ros WJ, et al. Quality of life after self-management cancer rehabilitation: a randomized controlled trial comparing physical and cognitive-behavioral training versus physical training. *Psychosomatic Medicine* 2008;**70**(4):422-9.

Lazowski 1999 *{published data only}*

Lazowski DA, Ecclestone NA, Myers AM, Paterson DH, Tudor-Locke C, Fitzgerald C, et al. A randomized outcome evaluation of group exercise programs in long-term care institutions. *Journal of Gerontology* 1999;**54**(12):M621-8.

Ligibel 2008 *{published data only}*

Ligibel JA, Campbell N, Partridge A, Chen WY, Salinardi T, Chen H, et al. Impact of a mixed strength and endurance exercise intervention on insulin levels in breast cancer survivors. *Journal of Clinical Oncology* 2008;**26**(6):907-12.

Livingston 2011 *{published data only}*

Livingston PM, Salmon J, Courneya KS, Gaskin CJ, Craike M, Botti M, et al. Efficacy of a referral and physical activity program for survivors of prostate cancer [ENGAGE]: rationale and design for a cluster randomised controlled trial. *BMC Cancer* 2011;**11**:237.

Mansky 2006 *{published data only}*

Mansky P, Sannes T, Wallerstedt D, Ge A, Ryan M, Johnson LL, et al. Tai chi chuan: mind-body practice or exercise intervention? Studying the benefit for cancer survivors. *Integrative Cancer Therapies* 2006;**5**(3):192-201.

Mathewson-Chapman 1997 *{published data only}*

Mathewson-Chapman M. Pelvic muscle exercise/biofeedback for urinary incontinence after prostatectomy: an education program. *Journal of Cancer Education* 1997;**12**(4):218-23.

Matthews 2007 *{published data only}*

Matthews CE, Wilcox S, Hanby CL, Der Ananian C, Heiney SP, Gebretsadik T, et al. Evaluation of a 12-week home-based walking intervention for breast cancer survivors. *Supportive Care in Cancer* 2007;**15**(2):203-11.

May 2008 *{published data only}*

May A M, Van Weert E, Korstjens I, Hoekstra-Weebers J E, Van Der Schans C P, Zonderland M L, et al. Improved physical fitness of cancer survivors: a randomised controlled trial comparing physical training with physical and cognitive-behavioural training. *Acta Oncologica* 2008;**47**(5):825-34.

McClure 2010 *{published data only}*

McClure MK, McClure RJ, Day R, Brufsky AM. Randomized controlled trial of the Breast Cancer Recovery Program for women with breast cancer-related lymphedema. *American Journal of Occupational Therapy* 2010;**64**(1):59-72.

McKenzie 2003 *{published data only}*

McKenzie DC, Kalda AL. Effect of upper extremity exercise on secondary lymphedema in breast cancer patients: a pilot study. *Journal of Clinical Oncology* 2003;**21**(3):463-6.

McNeely 2004 *{published data only}*

McNeely ML, Parliament M, Courneya KS, Seikaly H, Jha N, Scrimger R, et al. A pilot study of a randomized controlled trial to evaluate the effects of progressive resistance exercise training on shoulder dysfunction caused by spinal accessory neurapraxia/neurectomy in head and neck cancer survivors. *Head & Neck* 2004;**26**(6):518-30.

Midtgaard 2006 *{published data only}*

Midtgaard J, Rorth M, Stelter R, Adamsen L. The group matters: an explorative study of group cohesion and quality of life in cancer patients participating in physical exercise intervention during treatment. *European Journal of Cancer Care* 2006;**15**(1):25-33.

Milne 2008 *{published data only}*

Milne HM, Wallman KE, Gordon S, Courneya KS. Impact of a combined resistance and aerobic exercise program on motivational variables in breast cancer survivors: a

randomized controlled trial. *Annals of Behavioral Medicine* 2008;**36**(2):158-66.

Morey 2009 *{published data only}*

Morey MC, Snyder DC, Sloane R, Cohen HJ, Peterson B, Hartman TJ, et al. Effects of home-based diet and exercise on functional outcomes among older, overweight long-term cancer survivors: RENEW: a randomized controlled trial. *Journal of the American Medical Association* 2009;**301**(18):1883-91.

Mustian 2006 *{published data only}*

Mustian KM, Katula JA, Zhao H. A pilot study to assess the influence of tai chi chuan on functional capacity among breast cancer survivors. *Journal of Supportive Oncology* 2006;**4**(3):139-45.

Mutrie 2007 *{published data only}*

Mutrie N, Campbell AM, Whyte F, McConnachie A, Emslie C, Lee L, et al. Benefits of supervised group exercise programme for women being treated for early stage breast cancer: pragmatic randomised controlled trial. *British Medical Journal* 2007;**334**(7592):517.

Nikander 2007 *{published data only}*

Nikander R, Sievanen H, Ojala K, Oivanen T, Kellokumpu-Lehtinen PL, Saarto T. Effect of a vigorous aerobic regimen on physical performance in breast cancer patients - a randomized controlled pilot trial. *Acta Oncologica* 2007;**46**(2):181-6.

Osei-Tutu 2005 *{published data only}*

Osei-Tutu KB, Campagna PD. The effects of short- vs. long-bout exercise on mood, VO2max, and percent body fat. *Preventive Medicine* 2005;**40**(1):92-8.

Pinto 2009 *{published data only}*

Pinto BM, Rabin C, Dunsiger S. Home-based exercise among cancer survivors: adherence and its predictors. *Psychooncology* 2009;**18**(4):369-76.

Poorkiani 2010 *{published data only}*

Poorkiani M, Abbaszadeh A, Hazrati M, Jafari P, Sadeghi M, Mohammadianpanah M. The effect of rehabilitation on quality of life in female breast cancer survivors in Iran. *Indian Journal of Medical and Paediatric Oncology* 2010;**31**(4):105-9.

Rabin 2006 *{published data only}*

Rabin CS, Pinto BM, Trunzo JJ, Frierson GM, Bucknam LM. Physical activity among breast cancer survivors: regular exercisers vs participants in a physical activity intervention. *Psychooncology* 2006;**15**(4):344-54.

Sandel 2005 *{published data only}*

Sandel SL, Judge JO, Landry N, Faria L, Ouellette R, Majczak M. Dance and movement program improves quality-of-life measures in breast cancer survivors. *Cancer Nursing* 2005;**28**(4):301-9.

Schneider 2007 *{published data only}*

Schneider C M, Hsieh C C, Sprod L K, Carter S D, Hayward R. Exercise training manages cardiopulmonary function and fatigue during and following cancer treatment in male cancer survivors. *Integrative Cancer Therapies* 2007;**6**(3):235-41.

Segal 2001 *{published data only}*

Segal R, Evans W, Johnson D, Smith J, Colletta S, Gayton J, et al. Structured exercise improves physical functioning in women with stages I and II breast cancer: results of a randomized controlled trial. *Journal of Clinical Oncology* 2001;**19**(3):657-65.

Segal 2003 *{published data only}*

Segal RJ, Reid RD, Courneya KS, Malone SC, Parliament MB, Scott CG, et al. Resistance exercise in men receiving androgen deprivation therapy for prostate cancer. *Journal of Clinical Oncology* 2003;**21**(9):1653-9.

Snyder 2009 *{published data only}*

Snyder DC, Morey MC, Sloane R, Stull V, Cohen H J, Peterson B, et al. Reach out to ENhancE Wellness in Older Cancer Survivors (RENEW): design, methods and recruitment challenges of a home-based exercise and diet intervention to improve physical function among long-term survivors of breast, prostate, and colorectal cancer. *Psychooncology* 2009;**18**(4):429-39.

Sprod 2005 *{published data only}*

Sprod LK, Drum SN, Bentz AT, Carter SD, Schneider CM. The effects of walking poles on shoulder function in breast cancer survivors. *Integrative Cancer Therapies* 2005;**4**(4):287-93.

Twiss 2009 *{published data only}*

Twiss JJ, Waltman NL, Berg K, Ott CD, Gross GJ, Lindsey AM. An exercise intervention for breast cancer survivors with bone loss. *Journal of Nursing Scholarship* 2009;**41**(1):20-7.

Vallance 2007 *{published data only}*

Vallance JK, Courneya KS, Plotnikoff RC, Yasui Y, Mackey JR. Randomized controlled trial of the effects of print materials and step pedometers on physical activity and quality of life in breast cancer survivors. *Journal of Clinical Oncology* 2007;**25**(17):2352-9.

Vallance 2007a *{published data only}*

Vallance JK. Promoting physical activity in breast cancer survivors: a randomized controlled trial. Dissertation Abstracts International Section A: Humanities and Social Sciences 2007;**68**(7-A).

Vallance 2008 *{published data only}*

Vallance JK, Courneya KS, Plotnikoff RC, Dinu I, Mackey JR. Maintenance of physical activity in breast cancer survivors after a randomized trial. *Medicine and Science in Sports and Exercise* 2008;**40**(1):173-80.

Vallance 2008a *{published data only}*

Vallance JK, Courneya KS, Plotnikoff RC, Mackey JR. Analyzing theoretical mechanisms of physical activity behavior change in breast cancer survivors: results from the activity promotion (ACTION) trial. *Annals of Behavioral Medicine* 2008;**35**(2):150-8.

van Weert 2005 *{published data only}*

van Weert E, Hoekstra-Weebers J, Grol B, Otter R, Arendzen HJ, Postema K, et al. A multidimensional cancer rehabilitation

program for cancer survivors: effectiveness on health-related quality of life. *Journal of Psychosomatic Research* 2005;**58**(6):485-96.

van Weert 2010 *{published data only}*

van Weert E, May AM, Korstjens I, Post WJ, van der Schans CP, van den Borne B, et al. Cancer-related fatigue and rehabilitation: a randomized controlled multicenter trial comparing physical training combined with cognitive-behavioral therapy with physical training only and with no intervention. *Physical Therapy* 2010;**90**(10):1413-25.

von Gruenigen 2009 *{published data only}*

von Gruenigen VE, Gibbons HE, Kavanagh MB, Janata JW, Lerner E, Courneya KS. A randomized trial of a lifestyle intervention in obese endometrial cancer survivors: quality of life outcomes and mediators of behavior change. *Health and Quality of Life Outcomes* 2009;**7**:17.

Wall 2000 *{published data only}*

Wall LM. Changes in hope and power in lung cancer patients who exercise. *Nursing Science Quarterly* 2000;**13**(3):234-42.

Yeh 2011 *{published data only}*

Yeh CH, Man Wai JP, Lin US, Chiang YC. A pilot study to examine the feasibility and effects of a home-based aerobic program on reducing fatigue in children with acute lymphoblastic leukemia. *Cancer Nursing* 2011;**34**(1):3-12.

Zhang 2006 *{published data only}*

Zhang AY, Strauss GJ, Siminoff LA. Intervention of urinary incontinence and quality of life outcome in prostate cancer patients. *Journal of Psychosocial Oncology* 2006;**24**(2):17-30.

References to studies awaiting assessment

Utz-Billing 2010 *{published data only}*

Utz-Billing I, Rosemeyer M. Yoga supports physical and mental well-being after breast cancer surgery. *Journal of Psychosomatic Obstetrics and Gynecology* 2010;**31**(Suppl. 1):125.

References to ongoing studies

Devonish 2007 *{published data only}*

Devonish J-A, Culos-Reed N, Keats MR, Stephensopn LE. Physical activity for cancer survivors: fitness and quality of life results from a 16-week intervention. *Journal of Sport & Exercise Psychology* 2007;**29**:S157.

Galvao 2009 *{published data only}*

Galvao DA, Spry N, Taaffe DR, Denham J, Joseph D, Lamb DS, et al. A randomized controlled trial of an exercise intervention targeting cardiovascular and metabolic risk factors for prostate cancer patients from the RADAR trial. *BMC Cancer* 2009;**9**:419.

Hayes 2011 *{published data only}*

Hayes S, Rye S, Battistutta D, Yates P, Pyke C, Bashford J, et al. Design and implementation of the Exercise for Health trial -- a pragmatic exercise intervention for women with breast cancer. *Contemporary Clinical Trials* 2011;**32**(4):577-85.

Jones 2010 *{published data only}*

Jones LW, Douglas PS, Eves ND, Marcom PK, Kraus WE, Herndon JE 2nd, et al. Rationale and design of the Exercise Intensity Trial (EXCITE): A randomized trial comparing the effects of moderate versus moderate to high-intensity aerobic training in women with operable breast cancer. *BMC Cancer* 2010;**10**:531.

Jones 2010a *{published data only}*

Jones LW, Eves ND, Kraus WE, Potti A, Crawford J, Blumenthal JA, et al. The lung cancer exercise training study: a randomized trial of aerobic training, resistance training, or both in postsurgical lung cancer patients: rationale and design. *BMC Cancer* 2010;**10**:155.

Kampshoff 2010 *{published data only}*

Kampshoff CS, Buffart LM, Schep G, van Mechelen W, Brug J, Chinapaw MJ. Design of the Resistance and Endurance exercise After ChemoTherapy (REACT) study: a randomized controlled trial to evaluate the effectiveness and cost-effectiveness of exercise interventions after chemotherapy on physical fitness and fatigue. *BMC Cancer* 2010;**10**:658.

Persoon 2010 *{published data only}*

Persoon S, Kersten MJ, Chinapaw MJ, Buffart LM, Burghout H, Schep G, et al. Design of the EXercise Intervention after Stem cell Transplantation (EXIST) study: a randomized controlled trial to evaluate the effectiveness and cost-effectiveness of an individualized high intensity physical exercise program on fitness and fatigue in patients with multiple myeloma or (non-) Hodgkin's lymphoma treated with high dose chemotherapy and autologous stem cell transplantation. *BMC Cancer* 2010;**10**:671.

Saxton 2006 *{published data only}*

Saxton J M, Daley A, Woodroofe N, Coleman R, Powers H, Mutrie N, et al. Study protocol to investigate the effect of a lifestyle intervention on body weight, psychological health status and risk factors associated with disease recurrence in women recovering from breast cancer treatment [ISRCTN08045231]. *BMC Cancer* 2006;**6**:35.

Sekse 2011 *{published data only}*

Sekse RJT, Hausmann F, Iversen V, Ncsheim-Bjorkvik G, Tengesdal E, Bjerke A, et al. Rehabilitation of women following treatment for gynaecological cancer: a randomized, controlled study-a description of a multicenter study and experiences so far. Psychooncology 2011;**20**(Suppl 2):236-7.

Spence 2007 *{published data only}*

Spence RR, Heesch KC, Eakin EG, Brown WJ. Randomised controlled trial of a supervised exercise rehabilitation program for colorectal cancer survivors immediately after chemotherapy: study protocol. *BMC Cancer* 2007;**7**:154.

Vardy 2010 *{published data only}*

Vardy J, Dhillon H, Van Der Ploeg H, Goddard E, Clarke S, Spencer L, et al. Challenge: a Phase III study of a physical activity program on disease-free survival in patients with high risk stage II/III colon cancer. Asia-Pacific Journal of Clinical Oncology 2010;**6**(Suppl 3):186.

Walsh 2010 *{published data only}*

Walsh JM, Hussey J, Guinan E, O'Donnell D. 'Pragmatic randomized controlled trial of individually prescribed exercise versus usual care in a heterogeneous cancer survivor population': a feasibility study PEACH trial: prescribed exercise after chemotherapy. *BMC Cancer* 2010;**10**:42.

Additional references

Ahn 2007

Ahn SH, Park BW, Noh DY, Nam SJ, Lee ES, Lee MK, et al. Health-related quality of life in disease-free survivors of breast cancer with the general population. *Annals of Oncology* 2007;**18**(1):173-82.

American College of Sports Medicine 1998

American College of Sports Medicine. Position Stand: The Recommended Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory and Muscular Fitness, and Flexibility in Health Adults. *Medicine and Science in Sports and Exercise* 1998;**30**(6):975-91.

American College of Sports Medicine 2005

American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. 7th edition. Philadelphia: Lippincott Williams & Wilkins, 2005.

Aziz 2002

Aziz NM. Cancer survivorship research: challenge and opportunity. *Journal of Nutrition* 2002;**132**(11 Suppl):3494S-503S.

Aziz 2003

Aziz NM, Rowland JH. Trends and advances in cancer survivorship research: challenge and opportunity. *Seminars in Radiation Oncology* 2003;**13**(3):248-66.

Aziz 2007

Aziz NM. Cancer survivorship research: state of knowledge, challenges and opportunities. *Acta Oncologica* 2007;**46**(4):417-32.

Aziz 2008

Aziz NM, Bellizzi K. Older survivors and cancer care. Journal of the National Cancer Institute 2008;**100**(1):4-5.

Blanchard 2003

Blanchard CM, Denniston MM, Baker F, Ainsworth SR, Courneya KS, Hann DM, et al. Do adults change their lifestyle behaviors after a cancer diagnosis? *American Journal of Health Behavior* 2003;**27**(3):246-56.

Bottomley 2002

Bottomley A. The cancer patient and quality of life. *Oncologist* 2002;**7**(2):120-5.

Bradt 2011

Bradt J, Dileo C. Dance/movement therapy for improving psychological and physical outcomes in cancer patients. *Cochrane Database of Systematic Reviews* 2011, Issue 10. [DOI: 10.1002/14651858.CD007103]

Cella 2010

Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008.. *Journal of Clinical Epidemiology* 2010;**63**(11):1179-94.

Courneya 2004

Courneya KS, Vallance JK, McNeely ML, Karvinen KH, Peddle CJ, Mackey JR. Exercise issues in older cancer survivors. *Critical Reviews in Oncology/Hematology* 2004;**51**(3):249-61.

Courneya 2005a

Courneya KS, Karvinen KH, Campbell KL, Pearcey RG, Dundas G, Capstick V, et al. Associations among exercise, body weight, and quality of life in a population-based sample of endometrial cancer survivors. *Gynecologic Oncology* 2005;**97**(2):422-30.

Courneya 2007

Courneya KS, Friedenreich CM. Physical activity and cancer control. *Seminars in Oncology Nursing* 2007;**23**(4):242-52.

Courneya 2007a

Courneya KS, Karvinen KH. Exercise, aging, and cancer. *Applied Physiology, Nutrition, and Metabolism* 2007;**32**(6):1001-7.

Craft 2011

Craft LL, Vanlterson EH, Helenowski IB, Rademaker AW, Courneya KS. Exercise effects on depressive symptoms in cancer survivors: a systematic review and meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention* 2011;**21**(1):3-19.

Cramp 2008

Cramp F, Daniel J. Exercise for the management of cancer-related fatigue in adults. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD006145.pub2]

Cramp 2010

Cramp F, James A, Lambert J. The effects of resistance training on quality of life in cancer: a systematic literature review and meta-analysis. *Supportive Care in Cancer* 2010;**18**:1367-1376.

Demark-Wahnefried 2000

Demark-Wahnefried W, Peterson B, McBride C, Lipkus I, Clipp E. Current health behaviors and readiness to pursue life-style changes among men and women diagnosed with early stage prostate and breast carcinomas. *Cancer* 2000;**88**(3):674-84.

Demark-Wahnefried 2005

Demark-Wahnefried W, Aziz NM, Rowland JH, Pinto BM. Riding the crest of the teachable moment: promoting long-term health after the diagnosis of cancer. *Journal of Clinical Oncology* 2005;**23**(24):5814-30.

Diener 2000

Diener E. Subjective well-being. The science of happiness and a proposal for a national index. *American Psychologist* 2000;**55**(1):34-43.

Duijts 2011

Duijts SFA, Faber MM, Oldenburg HSA, van Beurden M. Effectivness of behavioral techniques and physical exercise on psychosocial functioning and health-related quality of life in breast cancer patients and survivors - a meta-analysis. *Psychooncology* 2011;**20**:115-126.

Efficace 2006

Efficace F, Bottomley A, Coens C, Van Steen K, Conroy T, Schoffski P, et al. Does a patient's self-reported health-related quality of life predict survival beyond key biomedical data in advanced colorectal cancer? *European Journal of Cancer* 2006;**42**(1):42-9.

Ferrans 2005

Ferrans CE. Definitions and conceptual models of quality of life. In: Lipscomb J, Gotay CC, Snyder C, editors(s). Outcomes Assessment in Cancer: Measures, Methods, and Applications. Cambridge: Cambridge University Press, 2005:14-30.

Ferrer 2011

Ferrer RA, Huedo-Medina TB, Johnson BT, Ryan S, Pescatello LS. Exercise interventions for cancer survivors: a meta-analysis of quality of life outcomes. *Annals of Behavioral Medicine* 2011;**41**:32-47.

Galvao 2005

Galvao DA, Newton RU. Review of exercise intervention studies in cancer patients. *Journal of Clinical Oncology* 2005;**23**(4):899-909.

Ganz 1998

Ganz PA, Rowland JH, Desmond K, Meyerowitz BE, Wyatt GE. Life after breast cancer: understanding women's health-related quality of life and sexual functioning. *Journal of Clinical Oncology* 1998;**16**(2):501-14.

Ganz 2002

Ganz PA, Desmond KA, Leedham B, Rowland JH, Meyerowitz BE, Belin TR. Quality of life in long-term, disease-free survivors of breast cancer: a follow-up study. *Journal of the National Cancer Institute* 2002;**94**(1):39-49.

Ganz 2004

Ganz PA, Kwan L, Stanton AL, Krupnick JL, Rowland JH, Meyerowitz BE, et al. Quality of life at the end of primary treatment of breast cancer: first results from the moving beyond cancer randomized trial. *Journal of the National Cancer Institute* 2004;**96**(5):376-87.

Glanville 2006

Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. *Journal of the Medical Library Association* 2006;**94**(2):130-6.

Gotay 1992

Gotay CC, Korn EL, McCabe MS, Moore TD, Cheson BD. Quality-of-life assessment in cancer treatment protocols: research issues in protocol development. *Journal of the National Cancer Institute* 1992;**84**(8):575-9.

Gotay 1998

Gotay CC, Muraoka MY. Quality of life in long-term survivors of adult-onset cancers. *Journal of the National Cancer Institute* 1998;**90**(9):656-67.

Gotay 2008

Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *Journal of Clinical Oncology* 2008;**26**(8):1355-63.

Hammerlid 2001

Hammerlid E, Taft C. Health-related quality of life in long-term head and neck cancer survivors: a comparison with general population norms. *British Journal of Cancer* 2001;**84**(2):149-56.

Higgins 2011

Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011.. Available from: www.cochrane-handbook.org.

Hjelm-Karlsson 1991

Hjelm-Karlsson K. Using biased coin designs for randomization in health care research. *Western Journal of Nursing Research* 1991;**13**:284-87.

Ingram 2007

Ingram C, Visovsky C. Exercise intervention to modify physiologic risk factors in cancer survivors. *Seminars in Oncology Nursing* 2007;**23**(4):275-84.

Jemal 2004

Jemal A, Clegg LX, Ward E, Ries LA, Wu X, Jamison PM, et al. Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival. *Cancer* 2004;**101**(1):3-27.

Jemal 2006

Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, et al. Cancer statistics, 2006. *CA: A Cancer Journal for Clinicians* 2006;**56**(2):106-30.

Jones 2004

Jones LW, Courneya KS, Vallance JK, Ladha AB, Mant MJ, Belch AR, et al. Association between exercise and quality of life in multiple myeloma cancer survivors. *Supportive Care in Cancer* 2004;**12**(11):780-8.

Kaminsky 1998

Kaminsky LA, Whaley M. Evaluation of a new standardized ramp protocol: the BSU/Bruce Ramp Protocol. *Journal of Cardiopulmonary Rehabilitation* 1998;**18**:438-44.

Karvinen 2007

Karvinen KH, Courneya KS, North S, Venner P. Associations between exercise and quality of life in bladder cancer survivors: a population-based study. *Cancer Epidemiology, Biomarkers and Prevention* 2007;**16**(5):984-90.

Knols 2005

Knols R, Aaronson NK, Uebelhart D, Fransen J, Aufdemkampe G. Physical exercise in cancer patients during and after medical treatment: a systematic review of randomized and controlled clinical trials. *Journal of Clinical Oncology* 2005;**23**(16):3830-42.

Lipscomb 2007

Lipscomb J, Gotay CC, Snyder CF. Patient-reported outcomes in cancer: a review of recent research and policy initiatives. *CA: A Cancer Journal for Clinicians* 2007;**57**(5):278-300.

McAuley 2000

McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet* 2000;**356**:1228-31.

McNeely 2006

McNeely ML, Campbell KL, Rowe BH, Klassen TP, Mackey JR, Courneya KS. Effects of exercise on breast cancer patients and survivors: a systematic review and meta-analysis. *Canadian Medical Association Journal* 2006;**175**(1):34-41.

Mishra 2010

Mishra SI, Scherer RW, Aziz NM, Gotay CC, Baquet CR, Berlanstein DR, et al. Exercise interventions on health related quality of life for people with cancer during active treatment (Protocol). *Cochrane Database of Systematic Reviews* 2010, Issue 4. [DOI: 10.1002/14651858.CD008465]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* 2009;**6**(7):e1000097.

Mustian 2007

Mustian KM, Morrow GR, Carroll JK, Figuero-Moseley CD, Jean-Pierre P, Williams GC. Integrative nonpharmacologic behavioral interventions for the management of cancer-related fatigue. *The Oncologist* 2007;**12**(suppl 1):52-67.

National Cancer Institute 2008

National Cancer Institute. Estimated US cancer prevalence counts: who are our cancer survivors? National Cancer Institute. Bethesda, MD, http://dccps.nci.nih.gov/ocs/prevalence/index.html; accessed on March 5, 2008.

National Cancer Institute 2012

National Cancer Institute. Patient Reported Outcomes Measurement Information System. Available at: www.nihpromis.org.

Osoba 1994

Osoba D. Lessons learned from measuring health-related quality of life in oncology. *Journal of Clinical Oncology* 1994;**12**(3):608-16.

Osoba 1999

Osoba D. What has been learned from measuring health-related quality of life in clinical oncology. *European Journal of Cancer* 1999;**35**(11):1565-70.

Osoba 2007

Osoba D. Health-related quality of life and predicting survival in cancer: not a simple matter. Supportive Care in Cancer 2007;**15**(4):353-5.

Rao 2006

Rao AV, Demark-Wahnefried W. The older cancer survivor. *Critical Reviews in Oncology/Hematology* 2006;**60**(2):131-43.

RevMan 2011 [Computer program]

The Cochrane Collaboration Review Manager (RevMan) [Computer program]. Version 5.1. The Nordic Cochrane Centre. Copenhagen: The Cochrane Collaboration, 2011.

Ries 2007

Ries LAG, Melbert D, Krapcho M, Mariotto A, Miller BA, Feuer EJ, et al. SEER Cancer Statistics Review, 1975-2004. National Cancer Institute, Bethesda, MD, http://seer.cancer.gov/csr/1975_2004/, based on November 2006 SEER data submission, posted to the SEER web site, 2007; accessed on March 5, 2008.

Rogers 2006

Rogers LQ, Courneya KS, Robbins KT, Malone J, Seiz A, Koch L, et al. Physical activity and quality of life in head and neck cancer survivors. *Supportive Care in Cancer* 2006;**14**(10):1012-9.

Rowland 2011

Rowland JH, Mariotto A, Alfano CM, Pollack LA, Weir HK, White A. Cancer Survivors--United States, 2007. *Morbidity and Mortality Weekly Report* March 11, 2011;**60**(9):269-72.

Schmitz 2005

Schmitz KH, Holtzman J, Courneya KS, Masse LC, Duval S, Kane R. Controlled physical activity trials in cancer survivors: a systematic review and meta-analysis. *Cancer Epidemiology, Biomarkers and Prevention* 2005;**14**(7):1588-95.

Speck 2010a

Speck RM, Courneya KS, Masse LC, Duval S, Schmitz KH. An update of controlled physical activity trials in cancer survivors: a systematic review and meta-analysis. *Journal of Cancer Survivorship* 2010;**4**:87-100.

Sterne 2011

Sterne JAC, Matthias E, Moher D (editors). Chapter 10: Addressing reporting bias. In: Higgins JPT, Green S, editors(s). Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011 Available at: www.cochrane-handbook.org.

Stevinson 2004

Stevinson C, Lawlor DA, Fox KR. Exercise interventions for cancer patients: systematic review of controlled trials. *Cancer Causes and Control* 2004;**15**(10):1035-56.

Stevinson 2007

Stevinson C, Faught W, Steed H, Tonkin K, Ladha AB, Vallance JK, et al. Associations between physical activity and quality of life in ovarian cancer survivors. *Gynecologic Oncology* 2007;**106**(1):244-50.

Stewart 2003

Stewart BW, Kleihues P, eds. World Cancer Report. Lyon, France: World Health Organization, International Agency for Research on Cancer, 2003.

Stull 2007

Stull VB, Snyder DC, Demark-Wahnefried W. Lifestyle interventions in cancer survivors: designing programs that meet the needs of this vulnerable and growing population. *Journal of Nutrition* 2007;**137**(1 Suppl):243S-8S.

Thorsen 2008

Thorsen L, Courneya KS, Stevinson C, Fossa SD. A systematic review of physical activity in prostate cancer survivors: outcomes, prevalence, and determinants. *Supportive Care in Cancer* 2008;**16**(9):987-97.

Valenti 2008

Valenti M, Porzio G, Aielli F, Verna L, Cannita K, Manno R, et al. Physical exercise and quality of life in breast cancer survivors. *International Journal of Medical Sciences* 2008;**5**(1):24-8.

Vallance 2005

Vallance JK, Courneya KS, Jones LW, Reiman T. Differences in quality of life between non-Hodgkin's lymphoma survivors meeting and not meeting public health exercise guidelines. *Psycho-Oncology* 2005;**14**(11):979-91.

Warburton 2006

Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *Canadian Medical Association Journal* 2006;**174**(6):801-9.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Bai 2004

Study characteristics	
Methods	Study design: RCT Number randomized: 45; 24 to the exercise group and 21 to the control group Study start, 2003; stop date, not reported Length of intervention: 3 months Length of follow-up: to end of the intervention
Participants	Type cancer: nasopharyngeal carcinoma Time since cancer diagnosis: unclear Time beyond active treatment: immediately after radiation therapy Inclusion criteria: • treated using same method and dosage of radiation therapy • KPS score > 60 Eligibility criterion related to interest or ability to exercise, or both: • none reported Exclusion criteria: • transfer to distant location • chronic disease • treated using chemotherapy • distant metastasis Gender, n: male, 26, female, 19 Current age: 23 to 65 years old Age at cancer diagnosis: not reported

Bai 2004 *(Continued)*

	Ethnicity/race: 100% Asian Education level: not reported SES: not reported Employment status: not reported Comorbidities: not reported Past exercise history: not reported
Interventions	24 participants assigned to the exercise intervention, including: • exercise of low to medium strength including jogging, swimming, and exercise with equipment every day • relaxation training of body muscles at least once per day • education on the diseases and psychological support Type exercise (aerobic/anaerobic): aerobic and anaerobic Intensity of experimental exercise intervention: low to moderate Frequency: once a day Duration of individual sessions: not reported Duration of exercise program: 3 months Total number of exercise sessions: unclear Participants were monitored every 2 weeks Format: individual Facility: not reported Professionally led: not reported Adherence: not reported Control group: 21 participants assigned to control group, including • no exercise Contamination of control group: none
Outcomes	Outcomes: QoL and physiologic outcomes, including: • QoL, assessed using the EORTC QLQ-C30; subscales included physical well-being, role function, EWB, social function, and cognitive function • nausea and vomiting • fatigue, but unclear how assessed • pain, but unclear how assessed Outcomes were measured at baseline and at 3 months: • exercise group: n = 24 at baseline, n = 24 after the intervention • control group: n = 21 at baseline, n = 21 after intervention Adverse events: none reported
Notes	Country: China Funding: not reported

Bai 2004 *(Continued)*

	Correspondence with investigator sought

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Low risk	No study participants were lost to follow-up and all were included in the analyses
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Banasik 2011

Study characteristics

Methods	Study design: RCT Number randomized: 18; 9 to the exercise group and 9 to the control group Study start and stop dates: not reported Length of intervention: 8 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, stages II to IV Time since cancer diagnosis: not reported Time beyond active treatment: at least 2 months' post-treatment Inclusion criteria: • none Eligibility criterion related to interest or ability to exercise, or both: • physical condition that prevented participation in yoga

Banasik 2011 *(Continued)*

Exclusion criteria:

- receiving Hercepton (trastuzumab) therapy (an immune modifier)
- pregnant or lactating
- past or current history of another neoplasm
- active serious infection or immune deficiency
- history of psychiatric disorders requiring use of psycho-active medications
- documented alcohol or drug abuse
- current steroid therapy or other known immunomodulating medications

Gender: female

Current age, mean (SD) years:

- exercise group: 63.33 (6.9) years
- control group: 62.4 (7.3) years

Age at cancer diagnosis: not reported

Ethnicity/race: 100% Caucasian

Education level: not reported

SES, (n):

- Exercise group: less than USD10,000 (0); USD10,000 to USD30,000 (2); USD30,000 to USD50,000 (2); USD50,000 to USD75,000 (3); > USD75,000 (2)
- Control group: less than USD10,000 (1); USD10,000 to USD30,000 (2); USD30,000 to USD50,000 (4); USD50,000 to USD75,000 (1); > USD75,000 (1)

Employment status: not reported

Comorbidities: not reported

Past exercise history: not reported

On hormone therapy: not reported

Interventions	9 participants assigned to the exercise intervention, including: • iyengar yoga, which balances physical demands with safety. The poses are taken slowly using props as necessary to maintain proper alignment and form. The active yoga used was primarily physical in nature and included poses traditionally found in beginning Iyengar classes. The sessions were more physically demanding than those of restorative or gentle yoga, with progressively difficult poses, including increased duration of weight-bearing on the arms, as individuals abilities improved. The focus of yoga practice was on training and accepting the physical form of the body and there was no specific component of meditation Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: mild Frequency: twice per week Duration of individual sessions: 90 minutes Duration of exercise program: 8 weeks Total number of exercise sessions: 16 Format: group Facility: facility

Banasik 2011 *(Continued)*

	Professionally led: professionally led by expert Iyengar yoga instructors 9 participants assigned to control group, including: • instructions to continue regular routine with offer of an opportunity for yoga program participation at the end of the study period Adherence: 7 women in the yoga group who completed the study attended an average of 14 of 16 possible yoga sessions (87.5%) with a range of 12 to 15 sessions Contamination of control group: not reported
Outcomes	Outcomes include QoL, measured using the FACT-B and subscales, including: • physical, social/family, emotional, and functional subscales • additional breast cancer concerns • fatigue score Outcomes were measured at baseline and at 8 weeks: • exercise group: n = 9 at baseline, n = 7 at 8 weeks • control group: n = 9 at baseline, n = 7 at 8 weeks Subgroup analysis: not reported Adverse events: no cancer recurrences or adverse events reported
Notes	Country: US Funding: University of Washington Center for Women's Health and Gender Research, Washington State University Cancer Prevention and Research Center, and the Washington State University College of Nursing

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	There was no ITT analysis and it is unclear how missing data were handled. 2 participants in each group withdrew and no reason was given for withdrawal
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes

Banasik 2011 *(Continued)*

Other bias	High risk	The small sample size and lack of description of the recruitment and selection of study participants could give rise to additional biases

Berglund 1994

Study characteristics	
Methods	Study design: RCT Number randomized: 199; 98 to the exercise group and 101 to the control group Study start and stop dates: not reported Length of intervention: 7 weeks Length of follow-up: 1 year after end of the intervention
Participants	Type cancer: 80% breast cancer, 7-8% ovarian cancer, remaining were other types Time since cancer diagnosis: not reported Time beyond active treatment: not reported Inclusion criteria: • < 75 years old • curative treatment for a primary tumor • within 2 months after postoperative treatment with radiation therapy or chemotherapy Eligibility criterion related to interest or ability to exercise, or both: • none Exclusion criteria: • none Gender: not reported Current age: not reported Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: not reported SES: not reported Employment status: not reported Comorbidities: not reported Past exercise history: not reported On hormone therapy: not reported
Interventions	98 participants assigned to the exercise intervention, including: • physical training, information, and coping skills training. The physical training component included exercises to increase mobility, muscle strength, general fitness, and relaxation in the form of progres-

Berglund 1994 *(Continued)*

	sive muscle relaxation or deep relaxation with positive images. Patients were given instructions for progressive relaxation at home Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: not reported Frequency: twice per week during the first 4 weeks (once for physical training and once for information), then once per week for coping skills training Duration of individual sessions: not reported Duration of exercise program: 7 weeks Total number of exercise sessions: 11 Format: group Facility: facility Professionally led: professionally led by an oncology nurse 101 participants assigned to control group, including: • information and coping skill training Adherence: the mean absenteeism among participants was 1 session, representing a variation of the number of participants per session between 3 and 7 (mean 4.9). Contamination of control group: not reported
Outcomes	No primary outcome identified. HRQoL outcomes included: • fatigue - measured as part of a nonspecified larger scale • body image - measured as part of a nonspecified larger scale • pain - measured as part of a nonspecified larger scale • global health - measured as part of a nonspecified larger scale • anxiety - measured using the shortened HADS • depression - measured using the shortened HADS • problems with QoL - not specified how this was measured • MAC scale Physical outcomes included: • physical strength • physical training • tiredness • body image • pain • global health Outcomes were measured at baseline; end of the intervention; and 3 months, 6 months, and 12 months: • exercise group: n = 98 at baseline, n = 90 at end of the intervention, n = 90 at 3 months, n = 88 at 6 months, n = 87 at 12 months • control group: n = 101 at baseline, n = 98 at end of the intervention, n = 93 at 3 months, n = 91 at 6 months, n = 89 at 12 months Subgroup analysis: not reported Adverse events: no cancer recurrences or adverse events reported

Berglund 1994 *(Continued)*

Notes	Country: Sweden Funding: Swedish Cancer Foundation

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Used Efron's method of randomization of small samples (Hjelm-Karlsson 1991)
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	There was no ITT analysis, it is unclear how missing values were handled, there were large losses to follow-up
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Bourke 2011

Study characteristics

Methods	Study design: RCT Number randomized: 18; 9 to the exercise group and 9 to the control group Study start and stop dates: not reported Length of intervention: 12 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: colon cancer, Dukes stages A to C Time since cancer diagnosis: not reported Time beyond active treatment: 6 to 24 month • exercise group: mean, 16.4 months • control group: mean, 16.7 month Inclusion criteria:

Bourke 2011 *(Continued)*

- histologically confirmed colon cancer (Dukes stages A to C)
- resected within previous 6 to 24 months

Eligibility criterion related to interest or ability to exercise, or both:

- participation in regular physical activity (purposeful activity of at least a moderate intensity of ≥ 30 minutes, 3 times per week)

Exclusion criteria:

- KPS score < 80
- unstable angina, uncontrolled hypertension, recent myocardial infarction, or pacemaker

Gender, n:

- exercise group: male (5); female (4)
- control group: male (7); female (2)

Current age, mean (SD) years:

- exercise group: 67.9 (5.7) years
- control group: 70.3 (8.7) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level: not reported

SES: not reported

Employment status: not reported

Comorbidities: not reported

Past exercise history: not reported

On hormone therapy: not reported

BMI, mean (SD) kg/m^2:

- exercise group: 26.9 kg/m^2 (3.8 kg/m^2)
- control group: 26.0 kg/m^2 (3.5 kg/m^2)

Interventions	9 participants assigned to the exercise intervention, including: • supervised and home-based exercise sessions, comprised of 2 group-based supervised exercise sessions once per week, including 30 minutes aerobic exercise (e.g. using treadmills, rowing ergometers, and cycling ergometers) and 2 to 4 sets of 8 to 12 repetitions (with a rest of 30 to 90 seconds between sets) of resistance exercises. In addition, participants were asked to complete similar aerobic activities at home once per week. For weeks 6 to 12, participants attended the university facility once per week and were asked to perform 2 home-based exercise sessions per week • dietary advice Type exercise (aerobic/anaerobic): aerobic and anaerobic Intensity of experimental exercise intervention: 55% to 85% of age predicted maximum HR Frequency: twice per week of supervised sessions and once per week at home for first 6 weeks, then once per week supervised session and twice per week at home for last 6 weeks Duration of individual sessions: 30 minutes and time necessary to complete resistance training Duration of exercise program: 12 weeks

Bourke 2011 *(Continued)*

	Total number of exercise sessions: 36 Format: group and individual Facility: Northern General Hospital, Sheffield, UK Professionally led: professionally led by experienced exercise physiologist at facility 9 participants assigned to control group, including: • usual care Adherence: 90% attendance (completed 146 of 162 sessions) and 94% compliance Contamination of control group: reported no significant difference in exercise behavior as assessed by Godin LSI (15; 95% CI 2 to 28)
Outcomes	No primary outcome identified. QoL outcomes included: • fatigue, measured using FACT-F scale • FACT-C scale Physical outcomes included: • exercise behavior, using Godin LSI • diet diaries • surface electromyography • exercise tolerance, using the Bruce Ramp Protocol (Kaminsky 1998) Outcomes were measured at baseline and at 12 weeks: • exercise group: n = 9 at baseline, n = 8 at 12 weeks • control group: n = 9 at baseline, n = 9 at 12 weeks Subgroup analysis: none reported Adverse events: not reported
Notes	Country: UK Funding: Sheffield Hallam University

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Sequence was generated via code numbers using nQuery statistical software
Allocation concealment (selection bias)	Low risk	Allocation was undertaken by a senior academic who was not directly involved in the recruitment or assessment of patients. The randomization sequence was not disclosed to the researcher responsible for the day-to-day running of the trial until patients had completed the baseline assessments
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to mask or blind the participants; however, it is unclear whether the lack of masking or blinding could influence the outcomes

Bourke 2011 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	The researcher responsible for the day-to-day running of the trial was informed of the randomization after collection of the baseline data. Other study personnel were not masked or blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	ITT analysis was used to compare participants in the groups they were randomly assigned and data were carried over from previous visits in cases of withdrawal of participants. One participant in the intervention group withdrew owing to a cerebrovascular accident
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcome
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Burnham 2002

Study characteristics	
Methods	Study design: RCT (participants matched on KPS and QoL) Number randomized: 21; 7 to a low-intensity exercise group, 7 to a moderate-intensity exercise group, and 7 to the control group Study start and stop dates: not reported Length of intervention: 10 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast or colon cancer; 5 breast cancer and 1 colon cancer in each of the 3 treatment groups Time since cancer diagnosis: not reported Time beyond active treatment, mean (SD) months: • low intensity exercise group: 10.3 (5.1) months • moderate intensity exercise group: 9.8 (4.2) months • control group: 9.0 (5.3) months Inclusion criteria: • cleared by physician to participate • surviving breast, colon, or lung cancer • score of 70 or more on the KPS scale Eligibility criterion related to interest or ability to exercise, or both: • none reported Exclusion criteria: • currently taking mood-enhancing medications or herbal remedies Gender: 15 female and 3 male Current age: 40 to 65 years of age, mean (SD) years: • low intensity exercise group: 54.2 (8.1) years

Burnham 2002 *(Continued)*

- moderate intensity exercise group: 50.7 (8.2) years
- control group: 56.0 (10.1) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level: not reported

SES: not reported

Employment status: not reported

Comorbidities: not reported

Past exercise history: not reported

On hormone therapy: not reported

Interventions	7 participants assigned to the low-intensity exercise group 7 participants assigned to the moderate-intensity exercise group Type of exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: low intensity (25% to 35% of HR reserve) or moderate intensity (40% to 50% of HR reserve) Frequency: 3 times per week Duration of individual sessions: initially 14 minutes, divided equally among the 3 exercise modalities (4 minutes and 40 seconds on the treadmill, stair-climber, and stationary bicycle in a rotational order). Increased by 2 minutes per week, up to 32 minutes at week 10 Duration of exercise program: 10 weeks Total number of exercise sessions: 30 Format: group Facility: facility Professionally led: unclear Adherence: 95% 7 participants assigned to the control, including: • no exercise Contamination of control group: 1 control participant increased exercise. That person and the match in the low- and moderate-intensity exercise groups were removed from the study
Outcomes	Primary outcome: QoL outcomes, including: • Quality of Life Index for Cancer Patients (100-mm analog, measuring QoL) • LASA (100-mm visual analog, measuring fatigue, anxiety, confusion, depression, energy, and anger) Secondary outcomes: physiologic measures, including: • peak aerobic capacity (treadmill) • body composition (3-site skinfold) • lower-body flexibility (modified sit and reach) Outcomes were measured at baseline, 5 and 10 weeks:

Burnham 2002 *(Continued)*

	• low-intensity exercise group: n = 6 at baseline, n = 6 at each follow-up • moderate-intensity group: n = 6 at baseline, n = 6 at each follow-up • control group: n = 6 at baseline, n = 6 at each follow-up Subgroup analysis by demographics Adverse events: none reported
Notes	Country: US Funding: not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	"No subject in any group withdrew from the study...", but "One subject was excluded from the control group when a post-study questionnaire revealed that she had engaged in significant exercise training during the course of the study... To maintain the matched group status, the two subjects matched with the excluded control subject were also removed from the analysis"
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Cadmus 2009

Study characteristics

Methods	Study design: RCT Number randomized: 75; 37 to the exercise group and 38 to the control group Study start and stop dates: March 2004 to July 2006 Length of intervention: 6 months Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, stage 0 to IIIA

Cadmus 2009 *(Continued)*

Time since cancer diagnosis, mean (SD) weeks:

- exercise group: 187.5 (114.2) weeks
- control group: 173.2 (135.2) weeks

Time beyond active treatment: at least 12 months

Inclusion criteria:

- cancer survivor
- 34 to 79 years old
- nondiabetic
- inactive (< 90 minutes per week of moderate to vigorous intensity recreational physical activity)
- postmenopausal

Eligibility criteria related to interest or ability to exercise, or both:

- none reported

Exclusion criteria:

- diagnosis of recurrent or other primary cancer event
- current smoker
- current or planned enrolment in a structured weight loss program
- premenopausal
- physically active
- diabetes mellitus

Gender: female

Current age, mean (SD) years:

- exercise group: 56.5 (9.5) years
- control group: 55.1 (7.7) years

Age at cancer diagnosis: not reported

Ethnicity/race: 84% non-Hispanic white for both groups

Education level:

- exercise group: college degree or higher, 60%
- control group: college degree or higher, 41%

SES: not reported

Employment status: not reported

Comorbidities: none

Past exercise history, mean (SD):

- exercise group: Physical Activity Questionnaire score, 13 (24) minutes per day of physical activity; pedometer, 5145 (2312) steps per day
- control group: Physical Activity Questionnaire score, 12 (20) minutes per day of physical activity; pedometer, 5342 (2744) steps per day

On hormone therapy:

- exercise group, 57%
- control group, 70%

Interventions	37 participants assigned to the exercise intervention, including:

Cadmus 2009 *(Continued)*

	• home- and facility-based supervised exercise program Type exercise (aerobic/anaerobic): aerobic Intensity of the experimental exercise intervention: Polar HR monitors to maintain the goal of 60% to 80% of predicted maximal HR Frequency: 5 days per week Duration of individual sessions: 30 minutes Duration of exercise program: 6 months Total number of exercise sessions: 120 Format: individual Facility: 2 days/week at home and 3 days/week at a facility (local health club) Professionally led: professionally led by an exercise physiologist Adherence: average 123 minutes/week (SD 52) of moderate to vigorous intensity sports/recreational activity (range: 0 to 637) • 34% of participants met the study goal of 150 minutes/week • 56% completed at least 120 minutes/week, or 80% of the study goal • 67% attended supervised exercise sessions • 96% reported exercising at least twice per week at home 38 participants assigned to control group, including: • usual exercise Contamination of control group: not reported
Outcomes	Outcomes: QoL and physiologic outcomes, including: • happiness, assessed using the 2-item Fordyce Happiness Measure • self-esteem, assessed using the RSE Scale • depression, assessed using the CES-D • anxiety, assessed using the STAI • stress, assessed using the Cohen's 10-item Perceived Stress Scale • QoL, assessed using FACT-B • QoL, assessed using the MOS SF-36 • physical activity • anthropometric measurements Outcomes were measured at baseline and 6 months: • exercise group: n = 37 at baseline, n = 37 at 6 months • control group: n = 38 at baseline, n = 37 at 6 months Adverse events: none reported
Notes	Country: US Funding: Lance Armstrong Foundation, American Cancer Society, Susan G. Komen. In part by the National Center of Research Resources (NIH)
Risk of bias	

Cadmus 2009 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomization code
Allocation concealment (selection bias)	Low risk	"The randomization code was obtained by the principal investigator (who was not involved in recruitment or data collection) only after baseline measures for that individual had been completed and staff conducting clinic visits did not have access to the randomization program"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Low risk	ITT analysis performed
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Cho 2006

Study characteristics	
Methods	Study design: quasi-RCT Number randomized: 65; 34 to the exercise group and 31 to the control group Study start and stop dates: October 2002 to June 2003 Length of intervention: 10 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, stage I to II Time since cancer diagnosis: mean 14.8 months Time beyond active treatment, mean (SD) months: • exercise group: time since mastectomy, 15.5 (5.9) months • control group: time since mastectomy, 17 (6.2) months Inclusion criteria: • histologically confirmed early stage (stages I, II) breast cancer • within 2 years of mastectomy • completion of chemotherapy, radiation therapy, or both

Cho 2006 *(Continued)*

Eligibility criterion related to interest or ability to exercise, or both:

- none reported

Exclusion criteria:

- evidence of recurrent or progressive cancer
- mental or systematic disease

Gender: female

Current age, mean (SD) years:

- exercise group: 48.7 (9.1) years
- control group: 49.6 (6.2) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level, n (%):

- exercise group: less than middle school, 3 (10.7%); High school, 10 (35.7%); more than college, 15 (53.6%)
- control group: less than middle school, 7 (26.0%); High school, 10 (37.0%); more than college, 10 (37.0%)

SES: not reported

Employment status: not reported

Comorbidities: not reported

Past exercise history: not reported

On hormone therapy, n (%):

- exercise group: 14/28 (50%)
- control group: 17/27 (63%)

Interventions	34 participants assigned to the exercise intervention, including: • exercise • psychology-based education • peer support group activity Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: to maximum HR of 40% to 60% Frequency: twice per week Duration of individual session: 90 minutes Duration of exercise program: 10 weeks Total number of exercise sessions: 20 Format: individual and group Facility: home and tertiary care hospital Professionally led: registered fitness instructor Adherence: not reported

Cho 2006 *(Continued)*

	Co-intervention: none Control group: 31 assigned to control group, consisting of • waiting list Contamination of control group: not reported
Outcomes	Outcomes: QoL and physiologic outcomes, including: • change in range of motion of the shoulder joint, assessed using ROM goniometer • change in psychosocial adjustment, assessed using 18-item, 4-point scale, 1 = never, 2 = no, 3 = yes, 4 = very much • change in QoL, using an instrument developed by Chae-Choe, with 27 items Outcomes were measured at baseline and 10 weeks: • exercise group: n = 34 at baseline, n = 28 at 10 weeks • control group: n = 31 at baseline, n = 27 at 10 weeks Adverse events: recurrence of cancer (n = 3)
Notes	Country: South Korea Funding: not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Stated that it is a quasi-randomized study but details not given
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	10 participants not included in analyses, 3 participants had an adverse event and 7 participants withdrew
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Cohen 2004

Study characteristics	
Methods	Study design: RCT Number randomized: 39; 20 to the exercise group and 19 to the control group Study start and stop dates: not reported Length of intervention: 7 weeks Length of follow-up: 1 week, 1 month and 3 months after the last session
Participants	Type cancer: lymphoma • exercise group: Stage 1, 22%; Stage II, 39%; Stage III, 17%; Stage IV, 22% • control group: Stage 1, 22%; Stage II, 33%; Stage III, 12%; Stage IV, 33% Time since cancer diagnosis: not reported Time beyond active treatment: not reported Inclusion criteria: • lymphoma • receiving chemotherapy or had received it within the past 12 month • ≥ 18 years old • able to read and speak English Eligibility criterion related to interest or ability to exercise, or both: • none reported Exclusion criteria: • major psychotic illnesses Gender: 12 female and 32 male Current age, mean, 51 years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: not reported SES: not reported Employment status: not reported Comorbidities: not reported Past exercise history, n: • exercise group: 4 • control group: 8 On hormone therapy, n: • exercise group: 1 • control group: 2
Interventions	19 participants assigned to the Tibetan yoga exercise intervention, including:

Cohen 2004 *(Continued)*

- controlled breathing and visualization
- mindfulness
- postures from the Tsa lung
- preliminary set of postures from the Trul khor (sngon 'gro).

The exercises are simple motions done with specific breathing patterns

Type exercise (aerobic/anaerobic): aerobic

Intensity of experimental exercise intervention: mild

Frequency: once per week, with recommendation to practice techniques at home at least daily

Duration of individual session: not reported

Duration of exercise program: 7 weeks

Total number of exercise sessions: 7

Format: group and individual

Facility: tertiary care hospital and home

Professionally led: Tibetan yoga instructor

Adherence: all participants attended at least 1 yoga session; 6 (32%) attended all 7 sessions; 5 (26%) attended 5 or 6 sessions; 6 (32%) attended 2
or 3 sessions; and 2 (10%) attended only 1 session

Co-intervention: none

Control group: 19 assigned to control group, consisting of

- wait list

Contamination of control group: not reported

Outcomes	Outcomes: QoL outcomes, including: • psychological distress, assessed using the Impact of Events Scale • anxiety, assessed using the Spielberger State Anxiety Inventory • depression, assessed using the CES-D • fatigue, assessed using the Brief Fatigue Inventory • sleep, assessed using the PSQI Outcomes were measured at baseline and 1 week, 1 month, and 3 months after the last yoga session: • exercise group: n = 20 at baseline, n = 19 at follow-up (time of measure not reported) • control group: n = 19 at baseline, n = 19 at follow-up (time of measure not reported) Adverse events: not reported
Notes	Country: US Funding: Bruce S. Gelb Foundation

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Group assignment was conducted sequentially using minimization"

Cohen 2004 *(Continued)*

Allocation concealment (selection bias)	Low risk	"The allocation process was concealed from all investigators because all the relevant information was entered into a computer program and group assignment was determined by the program"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	Although it was stated that only 1 study participant dropped out before the end of the study, data were presented only for 30 study participants, not 38 who completed the study
Selective reporting (reporting bias)	Low risk	There is no evidence of selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Courneya 2003a

Study characteristics	
Methods	Study design: cluster RCT, where clusters were psychotherapy classes Number randomized: 108; 60 (in 11 classes) to the exercise group and 48 (in 11 classes) to the control group Study start and stop dates: the group psychotherapy classes were conducted between September 1998 and April 2001 Length of intervention: 10 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, 40.6%; colon cancer, 9.4%; ovarian cancer, 5.2%; stomach cancer, 4.2%, melanoma, 4.2%; Hodgkin disease, 3.1%; NHL, 3.1%; brain cancer, 3.1%; lung cancer, 3.1%; other, 15.6%; missing, 8.3% Time since cancer diagnosis, mean (SD) months: • exercise group: 16.79 (18.45) months • control group: 15.71 (16.70) months Time beyond active treatment: not reported, although 43.5% of participants in exercise group and 45.2% of participants in control group were still receiving treatment Inclusion criteria: • diagnosis of cancer • voluntary participation in a group psychotherapy class offered at the cancer institute • ability to answer questions written in English Eligibility criterion related to interest or ability to exercise, or both:

Courneya 2003a *(Continued)*

- passing the rPAR-Q, a screening tool to determine the need to consult a physician before increasing exercise levels
- no contraindications to moderate-intensity exercise based on a submaximal fitness assessment were inclusionary

Exclusion criteria:

- none

Gender, %:

- exercise group: female, 84.4%
- control group: female, 86.7%

Current age, mean (SD) years:

- exercise group: 52.51 (10.21) years
- control group: 50.53 (10.08) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level, % completing university:

- exercise group: 57.1%
- control group: 60.5%

SES, % with family income > USD40,000:

- exercise group: 63.0%
- control group: 71.1%

Employment status, % currently employed:

- exercise group: 64.61%
- control group: 47.7%

Comorbidities: not reported

Past exercise history, mean (SD) minutes participants engaged in mild, moderate, or strenuous exercise:

- exercise group: 192.53 (227.43) minutes
- control group: 137.68 (117.76) minutes

On hormone therapy: not reported

Interventions	60 participants assigned to the personalized exercise intervention, including: • prescription for walking although participants were allowed to choose alternate mode of exercise (e.g. swimming, cycling) • group psychotherapy Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: goal was to achieve 65% to 75% of estimated HR maximum as soon as safely possible Frequency: 3 to 5 times per week Duration of individual sessions: 20 to 30 minutes Duration of exercise program: 10 weeks

Courneya 2003a *(Continued)*

	Total number of exercise sessions: variable, but maximum would be 50 sessions Format: individual Facility: home, with group psychotherapy classes offered in facility (Cross Cancer Institute) Not professionally led 48 participants assigned to the control group, including: • group psychotherapy Adherence: 51/60 participants completed the 10-week intervention; 43/51 (84.3%) achieved the minimum exercise prescription of 60 minutes of moderate to strenuous exercise per week and 16/51 (31.4%) achieved the optimum exercise prescription of 150 minutes of moderate to strenuous exercise per week. Total minutes of exercise, mean (SD) 196.65 (149.56) minutes Contamination of control group: mean (SD) minutes when participants in the control group participated in exercise 100.91 (104.24) minutes
Outcomes	No primary outcome was identified. QoL outcomes included: • QoL, assessed using the FACT-G and subscales for physical, functional, emotional, social/family, and spiritual wellbeing • satisfaction with life, measured using the SWLS • depression, assessed by the CES-D scale • anxiety, assessed by the STAI • fatigue, assessed using the 13-item FACT-F Outcomes were measured at baseline and 10 weeks: • exercise group: n = 60 at baseline, n = 51 at 10 weeks • control group: n = 48 at baseline, n = 45 at 10 weeks Subgroup analysis: several subgroup analyses were prespecified and conducted Adverse events: none reported
Notes	Country: Canada Funding: NIH, Canadian Institutes of Health Research, NCIC, CCS, CCS/NCIC Sociobehavioral Cancer Research Network

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Allocation sequence was generated using a random numbers table
Allocation concealment (selection bias)	High risk	Allocation was not completely concealed. It was concealed from the fitness appraiser but not from other study personnel
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions

Courneya 2003a *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	High risk	Although stated that analyses were conducted on an ITT basis, the treatment of missing data was not described. There was substantial attrition from the study in both study groups
Selective reporting (reporting bias)	Low risk	There is no evidence of selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Courneya 2003b

Study characteristics	
Methods	Study design: RCT Number randomized: 102; 69 to the exercise group and 33 to a waiting list control group Study start and stop dates: October 1998 to April 2001 Length of intervention: 16 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: colorectal cancer Time since cancer diagnosis: not reported Time beyond active treatment: surgery within the last 3 months Inclusion criteria: • recovery from surgery as indicated by attending physician • ability to understand English Eligibility criterion related to interest or ability to exercise, or both: • passed the revised Physical Activity Readiness Questionnaire • no contraindications to exercise as determined by a submaximal cardiorespiratory fitness test Gender: • exercise group: 54.8% male • control group: 64.5% male Current age, mean (SD) years: • exercise group: mean 59.92 (10.73) years • control group: 61.1 (9.93) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: • exercise group: 35% completed university • control group: 46.4% completed university SES:

Courneya 2003b *(Continued)*

- exercise group: 65.5% with income > USD40,000
- control group: 53.6% with income >USD40,000

Employment status:

- exercise group: 29.5% employed full time
- control group: 30.0% employed full time

Comorbidities: not reported

Past exercise history: mean (SD) number of minutes of exercise per week:

- exercise group: mild exercise, 121.45 (214.60) minutes; moderate exercise, 77.98 (137.01) minutes; strenuous exercise, 13.47 (69.36) minutes; total, 212.90 (248.04) minutes; % > 60 moderate/strenuous exercise, 40.3%, % > 150 moderate/strenuous, 27.4%
- control group: mild exercise, 164.03 (295.10); moderate exercise, 68.87 (97.57); strenuous exercise, 27.74 (57.43); total, 260.65 (323.77); % > 60 moderate/strenuous exercise, 41.9%, % > 150 moderate/strenuous, 32.3%

On hormone therapy: not reported

Interventions	69 participants assigned to the experimental exercise intervention, including: • home-based personalized exercise program that could be any activity designed to "improve functional wellbeing through cardiovascular and flexibility exercises", if none, walking prescribed Type exercise (aerobic/anaerobic): unclear Intensity of experimental exercise intervention: to 65% to 75% of HR Frequency: 3 to 5 times per week Duration of session: 20 to 30 minutes Duration of exercise program: 16 weeks Total number of exercise sessions: 48 to 80 Format: individual Facility: home based Not professionally led, but designed by professional Adherence: overall adherence, 75.8% Calculated as effect size (difference in variable between groups divided by SD of control group): • mild exercise = 0.20 • moderate exercise = 0.16 • strenuous exercise = 0.07 • moderate/strenuous = 0.15 • total exercise = 0.07 % > 60 moderate/strenuous, 75.8% % > 150 moderate/strenuous, 41.9% 33 participants assigned to control: • waiting list control Contamination of control group: overall, 51.6% Calculated as effect size (difference in variable between groups divided by SD of control group)

Courneya 2003b *(Continued)*

	% > 60 moderate/strenuous, 51.6% % > 150 moderate/strenuous, 32.3%
Outcomes	Primary outcome: QoL, measured at week 16 using: • FACT-C scale, includes subscales for physical, functional, emotional and social/family well-being, and colorectal subscale • FACT-G, excludes colorectal subscale • TOI score (sum of physical and FWB subscale and colorectal subscale) Secondary outcomes, all measured at 16 weeks included: • satisfaction with life, measured using the Satisfaction with Life scale (5 items rated on 7-point scale) • depression, measured using the CES-D scale • anxiety, measured using the STAI • cardiovascular fitness, measured using Modified Balke Treadmill Test • body composition, measured using Harpenden calipers • flexibility, measured using the sit and reach test Outcomes were measured at baseline and end of the intervention: • exercise group: n = 69 at baseline, n = 62 after intervention • control group: n = 33 at baseline, n = 31 after intervention Adverse events: none reported
Notes	Study country: Canada Funding source: NCIC and Alberta Heritage Foundation for Medical Research; Canadian Institutes of Health Research, CCS, Sociobehavioral Cancer Research Network

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Used a "random-numbers table"
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Fitness test conducted by certified fitness consultant blinded to the experimental group
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reasons for exclusions presented
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes

Courneya 2003b *(Continued)*

Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Courneya 2003c

Study characteristics	
Methods	Study design: RCT Number randomized: 53; 25 to the exercise group and 28 to the control group Study start and stop dates: recruitment from May 2001 to June 2001 Length of intervention: 15 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, n (%) Cancer stage, n (%): • exercise group: Stage I, 10 (42%); Stage IIa, 6 (25%); Stage IIb, 6 (25%); Stage IIIa, 2 (8%) • control group: Stage I, 11 (39%); Stage IIa, 11 (39%); Stage IIb, 5 (185); Stage IIIa, 1 (4%) Time since cancer diagnosis: not reported Time beyond active treatment: not reported Inclusion criteria: • histologically confirmed early-stage breast cancer • diagnosis between January 1999 and June 2000 • completed surgery, radiation therapy, chemotherapy, or a combination with or without current hormone therapy use • postmenopausal status • 50 to 69 years old Eligibility criteria related to interest or ability to exercise, or both: • contraindication to exercise on the basis of an exercise stress test Exclusion criteria: • evidence of recurrent or progressive disease • smoked within previous 12 months • non-English-speaking • not willing to travel to the exercise facility • known cardiac disease, uncontrolled hypertension, thyroid disease, diabetes, mental illness, infection, and immune or endocrine abnormality Gender: female Current age, mean (SD) years: • exercise: 59 (5) years • control: 58 (6) years Age at cancer diagnosis: not reported Ethnicity/race: not reported

Courneya 2003c *(Continued)*

	Education level, n (%) • exercise group: 7 (29%) completed university • control group: 16 (56%) completed university Household income > USD60,000, n (%): • exercise: 10 (44%) • control: 13 (48%) Employment status: employed full time, n (%) • exercise: 7 (29%) • control: 8 (29%) Comorbidities: none reported Past exercise history, mean (SD) minutes: • exercise: moderate, 62 (94) minutes; strenuous, 23 (56) minutes; moderate to strenuous, 85 (102) minutes; > 90 moderate to strenuous, 10 (42) minutes • control: moderate, 98 (126) minutes; strenuous, 26 (65) minutes; moderate to strenuous, 124 (146) minutes; > 90 moderate to strenuous, 12 (43) minutes On hormone therapy: exercise, 11 (46); control, 13 (46)
Interventions	25 participants assigned to the exercise intervention, including: • participants trained on recumbent or upright cycle ergometers Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: 70% to 75% maximal oxygen consumption in untrained subjects Frequency: 3 times per week Duration of sessions: 15 minutes for weeks 1 to 3, then increased by 5 minutes per week to 35 minutes at weeks 13 to 15. A 5-minute warm-up and cool-down period was included Duration of exercise program: 15 weeks Total number of exercise sessions: 45 Format: unclear Facility: facility Professionally led: sessions were supervised by exercise physiologists Adherence: the exercise group completed 98.4% (44.3 of 45) of the prescribed exercise sessions Co-intervention: none Control group: waiting list Contamination of control group: non–protocol-related exercise was < 15 minutes of moderate to strenuous exercise per week and was not different between groups
Outcomes	Primary outcomes included: • physical outcomes, including change in peak oxygen consumption • QoL outcomes, assessed using the FACT-B scale and the FACT-G scale Secondary outcomes, included:

Courneya 2003c *(Continued)*

- physiologic outcomes, including peak power output, oxygen consumption, power output at the ventilatory equivalent for oxygen and oxygen consumption and power output for the ventilatory equivalent for carbon dioxide
- QoL outcomes, including happiness, assessed using the Happiness Measure, self-esteem assessed using the RSE Scale, and fatigue assessed using FACT-F
- physical outcomes, including body weight, BMI, subcutaneous sum of skinfolds

Outcomes were measured at baseline and 15 weeks:

- exercise group: n = 25 at baseline, n = 24 at 15 weeks
- control group: n = 28 at baseline, n = 26 at 15 weeks

Adverse events:

- exercise group: lymphedema (n = 3), gynecologic complication (n = 1), and influenza (n = 1)
- control group: foot fracture (n = 1), bronchitis (n = 1)

Notes	Country: Canada Funding: NCIC, CCS, Canadian Institutes of Health Research, Izaak Walton Killiam Memorial Scholarship, Alberta Heritage Foundation for Medical Research studentship

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random-numbers table
Allocation concealment (selection bias)	Low risk	The allocation sequence and group assignments were generated by a research assistant and then enclosed in sequentially numbered and sealed envelopes. The contents of the envelopes were concealed from the project director, who assigned participants to groups
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not possible to blind study participants for self-report measures. Exercise physiologists were blinded for physical outcome measures
Incomplete outcome data (attrition bias) All outcomes	Low risk	1 study participant withdrew from the exercise group and 2 from the control group; they were not included in the physical outcome analyses. The QoL analyses included all but the 1 study participant who had withdrawn from the exercise group
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Courneya 2009

Study characteristics

Courneya 2009 *(Continued)*

Methods	Study design: RCT Number randomized: 122; 60 to an exercise group and 62 to the control group Study start and stop dates: 2005 to 2008 Length of intervention: 12 weeks Length of follow-up: 6 months
Participants	Type cancer: lymphoma Stage/type cancer, n (%): • exercise group: Stage I, 11 (18.3%); Stage II, 8 (13.3%); Stage III, 9 (15.0%); Stage IV, 15 (25.0%); NHL indolent, 25(41.7%); NHL aggressive, 24 (40.0%); HL, 11 (18.3%) • control group: Stage I, 7 (11.3%); Stage II, 15 (24.2%); Stage III, 8 (12.9%); Stage IV, 13 (21.0%); NHL indolent, 27(43.5%); NHL aggressive, 24 (38.7%); HL, 11 (17.7%) Time since cancer diagnosis, mean (SD) months since diagnosis: • exercise group: 25.3 (31.5) • control group: 33.0 (39.0) Time beyond active treatment: not reported, but some participants still being actively treated Inclusion criteria: • English speaking • ≥ 18 years • histologically confirmed HL or NHL • receiving chemotherapy or no treatment. Patients receiving chemotherapy may have started treatment before enrolment but needed to have at least 8 weeks of planned treatment remaining Eligibility criterion related to interest or ability to exercise, or both: • none Exclusion criteria: • uncontrolled hypertension • cardiac illness • residence more than 80 km from facility • not approved by their oncologist Gender, n (%): • exercise group: male, 37 (61.7%) • control group: male, 35 (56.5%) Current age, mean (range) years: • exercise group: 52.8 (18 to 77) years • control group: 53.5 (18 to 80) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level, completed university, n (%): • exercise group: 31 (51.7%) • control group: 32 (51.6%)

Courneya 2009 *(Continued)*

SES, income > USD60,000 per year n (%):

- exercise group: 34 (63.0%)
- control group: 39 (62.9%)

Employment status, employed, n (%):

- exercise group: 22 (36.7%)
- control group: 32 (51.6%)

Comorbidities, n (%):

- exercise group: arthritis, 24 (40.0%); hypercholesterolemia, 18 (30.0%); hypertension, 14 (23.3%)
- control group: arthritis, 14 (22.6%); hypercholesterolemia, 18 (29.0%); hypertension, 21 (33.9%)

Past exercise history, baseline exerciser, n (%):

- exercise group: 12 (20.0%)
- control group: 23 (37.1%)

On hormone therapy: not reported

Current chemotherapy, n (%):

- exercise group: 28 (46.7%)
- control group: 26 (41.9%)

Other characteristics, n (%):

- exercise group: overweight, 27 (45.0%); obese, 16 (26.7%); current smoker, 4 (6.7%)
- control group: overweight, 20 (32.3%); obese, 17 (27.4%); current smoker, 9 (14.5%)

Other characteristics, mean (SD):

- exercise: BMI 27.4 (4.5) kg/m^2; weight (kg), 81.8 (14.8) kg
- control: BMI 26.7 (5.4) kg/m^2; weight (kg), 78.5 (17.1) kg

Interventions	60 participants assigned to the exercise group, including • exercise on an upright or recumbent cycle ergometer (Life Fitness, Schiller Park, IL) for 12 weeks • one session per week of interval training above the ventilatory threshold in week 7 • one session of $VO_{2\ peak}$ interval training in week 9 Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: started at 60% of peak power output ($VO_{2\ peak}$) and increased by 5% each week to 75% by the fourth week Frequency: 3 times per week Duration of individual sessions: 15 to 20 minutes for first 4 weeks, increased by 5 minutes per week to 40 to 45 minutes in the ninth week Duration of exercise program: 12 weeks Total number of exercise sessions: 36 Format: group Facility: facility Professionally led by an exercise physiologist 62 participants assigned to the control group, including:

Courneya 2009 *(Continued)*

- request not to increase exercise above baseline levels
- offer of 4 weeks supervised exercise at end of the study

Adherence: attended a mean of 28/36 (77.8%) and a median of 33/36 (91.7%) supervised sessions. Duration and intensity were met during 27.8/28 (99.0%) and 25.4/28 (90.7%) supervised sessions, respectively:

- 45/60 (75%) participants attended ≥ 66% of sessions
- 3/60 (65%) participants attended ≥ 80% of sessions
- 21/60 (35%) participants attended 100% of sessions

Contamination of control group: the mean change in vigorous exercise from baseline: - 4 minutes:

- 49/62 (79%) participants reported no regular vigorous exercise during intervention
- 13/62 (21%) participants reported regular vigorous exercise during intervention

Outcomes	Primary outcome: patient-rated physical functioning, assessed using the TOI-An from the FACT-An scale Secondary QoL outcomes included: • total FACT-An • FACT-F • happiness, assessed by the Happiness scale • depression assessed by the CES-D • anxiety assessed by the SF STAI • lymphoma symptoms by the lymphoma scale of the FACT-Lym • general health by the single item on the MOS SF-12 Outcomes were measured at baseline, 12 weeks, and 6 months: • exercise group: n = 60 at baseline, n = 57 at 12 weeks, n = 55 at 6 months • control group: n = 62 at baseline, n = 60 at 12 weeks, n = 55 at 6 months Subgroup analyses: major disease type, current treatment status (on chemotherapy versus not), patient preference, age, sex, marital status, disease stage at entry, general health, BMI Adverse events: 3 adverse events related to exercise (back, hip, knee)
Notes	Country: Canada Funding: Lance Armstrong Foundation; Canada Research Chairs Program; Alberta Heritage Foundation for Medical Research; NCIC; and by CCS and the NCIC/CCS Sociobehavioral Cancer Research Network

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Computer-generated sequence"
Allocation concealment (selection bias)	Low risk	"The allocation sequence was generated independently and concealed in opaque envelopes from the study coordinator who assigned participants to groups."
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes

Courneya 2009 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	"Outcomes assessors were not always blinded to group assignment but were trained in standardizing testing procedures."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Although stated ITT analyses, missing data not accounted for. In exercise group, 3 participants did not complete QoL measures post intervention and 3 at 6 months In control group, 2 participants did not complete QoL measures postintervention, and 5 at 6 months
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Culos-Reed 2006

Study characteristics	
Methods	Study design: randomized controlled cross-over trial Number randomized: 38; 20 to the yoga group and 18 to the waiting list control group Study start and stop dates: not reported Length of intervention: 7 weeks Length of follow-up: to end of intervention
Participants	Type cancer: mostly breast cancer (85%) Time since cancer diagnosis, mean (SD) months: 55.95 (54.30) months Time beyond active treatment: > 3 months Inclusion criteria: • ≥ 18 years old • not currently undergoing active treatment • no additional health concerns • a minimum of 3 months post-treatment • recovery from surgery as indicated by attending physician Eligibility criterion related to interest or ability to exercise, or both: • Leisure Score Index used to assess baseline physical activity Exclusion criteria: • heart condition • hypertension or heart drugs • bone and joint problems • chest pain during activity or at rest • loss of balance or dizziness A medical examination was required for participation

Culos-Reed 2006 *(Continued)*

	Gender: 95% female Current age, mean (SD): 51.2 (10.3) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: not reported SES: not reported Employment status: not reported Comorbidities: not reported Past exercise history: assessed using LSI of the Godin Leisure-Time Activity Index, but not reported On hormone therapy: not reported
Interventions	20 participants assigned to the exercise group, including: • yoga Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: unclear, each individual worked at own exertion level Frequency: once per week Duration of individual sessions: 75 minutes Duration of exercise program: 7 weeks Total number of exercise sessions: 7 Format: group Facility: yoga studio Professionally led: instructor with Bachelor of Science degree in kinesiology and certified as a yoga instructor Adherence: not clearly reported, although it appears all completed the intervention 18 participants assigned to the control group, including: • usual daily activities during waiting list Contamination of control group: no yoga reported
Outcomes	Primary outcome: QoL, measured on all 38 participants before and after the exercise intervention, using: • Profile of Mood State, a 65-item scale that assesses 6 affective dimensions: Tension–anxiety, depression–dejection, anger–hostility, vigor–activity, fatigue inertia, and confusion–bewilderment. It measures state (versus trait) attributes • SOSI, which rates the frequency of stress-related symptoms on a 5-point scale ranging from never to frequently, during the past week. 10 subscale scores are derived from 95 individual items: (1) peripheral manifestations; (2) cardiopulmonary symptoms; (2a) symptoms of arousal, (2b) upper respiratory symptoms; (3) central neurologic symptoms; (4) gastrointestinal symptoms; (5) muscle tension; (6) habitual patterns (e.g. smoking, drinking, nail biting); (7) depression; (8) anxiety/fear; (9) emotional irritability; (10) cognitive disorganization • EORTC QLQ-C30, a 30-item questionnaire includes 5 functional domains of QoL: Physical function (5 items), emotional function (4 items), cognitive function (2 items), social function (2 items), and role

Culos-Reed 2006 *(Continued)*

	function (2 items). There are also several symptom scales: fatigue (3 items), pain (2 items), nausea and vomiting (2 items), and 1 item each for dyspnea, sleep disturbance, appetite, constipation, diarrhea, and financial difficulties. Finally, 2 items assess global QoL. 7 items are answered in a 'Yes–No' format, 21 items are evaluated on a 4-point Likert-type scale rating the presence of problems on a range from 'not at all' to 'very much'. Global 2 questions: 7-point scale with the anchors of 'very poor' (1) to 'excellent' (7). Item scores are added together to calculate the subscale scores Secondary outcomes: • LSI of the Godin Leisure-Time Activity Index was used to assess previous physical activity levels. The LSI contains 3 questions that assess the frequency of mild, moderate, and strenuous physical activity performed for at least 15-minute duration during free time in a typical week within the past month • physical parameters: height, weight, and girth; grip strength, measured with a dynamometer (to the nearest 1.0 kg); flexibility measured by sit and reach measurements (to the nearest 0.5 cm); Rockport Walking Test was used as a measure of functional capacity via distance traveled in 6 minutes • Adverse events: not reported Outcomes were measured at baseline and end of the intervention: • exercise group: n = 20 at baseline, n = 18 after the intervention • control group, n = 18 at baseline, n = 18 after the intervention Adverse events: none reported
Notes	Country: Canada Funding: Alberta Heritage Foundation for Medical Research Population Health Investigator Award; Canadian Institutes of Health Research New Investigator Award; University of Calgary Research Grant

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	Two participants in the yoga group not included in the analyses. No reason given for the exclusion
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	High risk	The small sample size and the lack of description of the recruitment and selection of participants could give rise to additional biases

Daley 2007a

Study characteristics	
Methods	Study design: RCT with 3 arms Number randomized: 108; 34 to an exercise-therapy group, 36 to an exercise-placebo group, and 38 to a control group Study start and stop dates: January 2003 to July 2005 Length of intervention: 8 weeks Length of follow-up: 24 weeks
Participants	Type cancer: breast cancer, stage not reported Time since cancer diagnosis: not reported Time beyond active treatment: 12 to 36 months Inclusion criteria: • not regularly active • treated for localized breast cancer 12 to 36 months previously • 18 to 65 years old Eligibility criterion related to interest or ability to exercise, or both: • high activity level • contraindication to exercise, as assessed using Physical Activity Readiness Questionnaire • must be willing to attend supervised exercise sessions 3 times per week for 8 weeks • must be an exercise pre-contemplator, contemplater, or preparer as defined by the TTM Exclusion criteria: • presence of metastases • inoperable or active locoregional disease as determined by clinician • physical or psychiatric impairment that would seriously influence physical mobility • suffering from nausea, anorexia, or other diseases affecting health Gender: female Current age: mean (SD) years • exercise-therapy group: 51.6 (8.8) years • exercise-placebo group: 50.6 (8.7) years • control group: 51.1 (8.6) years Age at cancer diagnosis: not reported Ethnicity/race: • exercise-therapy group: 34/34 (100%) white • exercise-placebo group: 35/36 (97.2%) white • control group: 37/38 (97.4%) white Education level: • exercise-therapy group: secondary and A levels, 17/34 (50.0%); degree, 5/34 (14.7%); other, 12/34 (35.3%) • exercise-placebo group: secondary and A levels, 12/35 (34.3%); degree, 13/35 (37.1%); other, 10/35 (28.5%)

Daley 2007a *(Continued)*

- control group: secondary and A levels, 18/33 (54.5%); degree, 6/33 (15.2%); other, 9/33 (27.2%)

SES: not reported

Employment status: employed

- exercise-therapy group: 26/34 (76.5%)
- exercise-placebo group: 25/36 (69.4%)
- control group: 21/34 (58.3%)

Comorbidities: experiencing lymphedema

- exercise-therapy group: 16/34 (47.0%)
- exercise-placebo group: 11/36 (30.6%)
- control group: 18/38 (47.3%)

Past exercise history: assessed, but not reported

On hormone therapy:

- exercise-therapy group: 25/34 (73.5%)
- exercise-placebo group: 25/36 (69.4%)
- control group: 29/38 (76.3%)

Interventions	34 participants assigned to the exercise therapy intervention, including: • -1-to-1 sessions with an exercise specialist • exercise education/guidance at each session Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: HR and RPE were assessed every 2 minutes during sessions. Exercise-therapy sessions involved moderate-intensity exercise (65% to 85% of age-adjusted HR maximum and RPE of 12 to 13) Frequency: 3 times per week Duration of session: 50 minutes Duration of exercise program: 8 weeks Total number of exercise sessions: 24 Format: 1-to-1 Facility: university Professionally led: exercise specialist Exercise-placebo group: 36 participants assigned to exercise-placebo group, including: • 24 1-to-1 50-minute sessions during 8 weeks with light-intensity body conditioning/stretching (e.g. flexibility and passive stretching) exercises during which HR was maintained below 40% HR reserve (typically was kept below 100 beats per minute) • conversations on topics of everyday life Control group: 38 participants were assigned to the control group, including: • no activity or education Adherence: attended at least 70% of sessions • exercise-therapy group: 77% • exercise-placebo group: 88.9%

Daley 2007a *(Continued)*

	Contamination of control group: these groups did not increase their activity level
Outcomes	Primary outcome: QoL outcomes, including: • FACT-G • FACT-B Secondary outcomes included QoL and physiologic outcomes, including: • fatigue, assessed using the Revised Piper Fatigue Scale • satisfaction with life • depression, assessed using the BDI-II • Physical Self-Perception Profile, including five 6-item subscales: perceived sports competence, attractiveness of body, physical conditioning competence, physical strength competence, and physical self-worth • physical activity and exercise behavior, assessed by asking participants how often they had participated in 1 or more physical activities for 20 to 30 minutes per session in the last 5 months and by completion of the stage of change for exercise ladder • aerobic fitness, assessed using submaximal 8-minute single-stage walking test performed on a treadmill • height • weight • body fat, using bioelectrical impedence analysis • muscle function, assessed using Bioidex isokinetic machine Outcomes were measured at baseline and 8 and 24 weeks: • exercise therapy group: n = 34 at baseline, n = 33 at week 8, n = 31 at week 24 • exercise-placebo group: n = 36 at baseline, n = 36 at week 8, n = 34 at week 24 • control group: n = 38 at baseline, n = 33 at week 8, n = 31 at week 24 Adverse events: none reported
Notes	Country: UK Funding: Cancer Research UK

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"performed using stratified random permuted blocks"
Allocation concealment (selection bias)	Low risk	Telephone randomization service provided by an independent trials unit
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	High risk	"Outcome assessors were not blinded to participants' group allocation"

Daley 2007a *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Data were analyzed on an ITT basis. "The trial statistician was blinded to group codes. Little's test was used to examine whether missing data were missing completely at random"
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Danhauer 2009

Study characteristics	
Methods	Study design: RCT Number randomized: 44; 22 to the exercise group and 22 to the control group Study start and stop dates: recruitment from August 2005 to October 2006 Length of intervention: 10 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, DCIS or stages I to IV Cancer stage, n (%): • exercise group: DCIS, 3 (13.6%); Stage I, 5 (22.7%); Stage II, 10 (45.5%); Stage III, 3 (13.6%); Stage IV, 1 (4.6%) • control group: DCIS, 5 (22.7%); Stage I, 9 (40.9%); Stage II, 3 (13.6%); Stage III, 2 (9.1%); Stage IV, 3 (13.6%) Time since cancer diagnosis, mean (SD) months: • e group: 24.4 (39.5) months • control group: 22.8 (35.6) months Time beyond active treatment: 2 to 24 months post primary treatment (surgery); 34% still in active treatment Inclusion criteria: • ≥ 18 years old • 2 to 24 months post primary treatment (surgery) following initial diagnosis • recurrence of breast cancer within the past 24 months (regardless of treatment status) • able to understand English Eligibility criterion related to interest or ability to exercise, or both: • physically able to attend restorative yoga classes Exclusion criteria: • medical contraindications as reported by physician Gender: women Current age, mean (SD) years:

Danhauer 2009 *(Continued)*

- exercise group: 54.3 (9.6) years
- control group: 57.2 (10.2) years

Age at cancer diagnosis: not reported

Ethnicity/race, n (%):

- exercise group: non-Hispanic white, 19 (86.4%); African-American, 2 (9.1%); Asian/Pacific Islander, 1 (4.6%)
- control group: non-Hispanic white, 20 (90.9%); African-American, 1 (4.6%); Asian/Pacific Islander, 1 (4.6%)

Education level, n (%):

- exercise group: high school diploma/GED, 0 (0.0%); some college or vocational school, 6 (27.3%); college graduate, 5 (22.7%); graduate study or degree, 1 (50.0%)
- control group: high school diploma/GED, 3 (13.6%); some college or vocational school, 8 (36.4%); college graduate, 3 (13.6%); graduate study or degree, 8 (36.4%)

SES, n (%):

- exercise group: < USD35,000, 6 (30.0%); USD35,000 to USD49,999, 3 (15.0%); USD50,000 to USD99,999, 8 (40.0%); ≥ USD100,000, 3 (15.0%)
- control group: < USD35,000, 5 (27.8%); USD35,000 to USD49,999, 1 (5.6%); USD50,000 to USD99,999, 5 (27.8%); ≥ USD100,000, 6 (33.3%)

Employment status: not reported

Comorbidities: not reported

Past exercise history, n (%):

- exercise group: never had done yoga, 20 (90.9%); no yoga experience in the past year, 20 (90.9%)
- control group: never had done yoga, 15 (68.2%); no yoga experience in the past year, 18 (81.8%)

On hormone therapy: not reported

Ongoing treatment, n (%):

- exercise group: receiving chemotherapy, 8 (36.4%); receiving radiation therapy, 6 (27.3%)
- control group: receiving chemotherapy, 3 (13.6%); receiving radiation therapy, 3 (13.6%)

Interventions	22 participants assigned to the exercise intervention, including: • restorative yoga which combined physical postures (*asanas*), breathing (*pranayama*), and deep relaxation (*savasana*). Yoga poses were modified based on participant needs. Poses included: mountain pose, arm and shoulder stretch, supported forward fold, seated sun salutation, and reclining twist with a bolster Type exercise (aerobic/anaerobic): aerobic and anaerobic Intensity of experimental exercise intervention: mild Frequency: once per week Duration of individual sessions: 75 minutes Duration of exercise program: 10 weeks Total number of exercise sessions: 10 Format: group Facility: Wake Forest University Health Sciences and local studio

Danhauer 2009 *(Continued)*

	Professionally led by yoga instructor with cancer-specific yoga training who was registered by the National Yoga Alliance 22 participants assigned to control group, including: • usual care • waiting list for yoga Adherence: 11 women attended ≥ 7 sessions; 6 women attended 3 to 6 sessions; and 5 women attended ≤ 2 sessions Contamination of control group: not reported
Outcomes	No primary outcomes were identified. Outcomes included: • physical health status (PCS and MCS), measured using the MOS SF-12, which is a 12-item self-report measure of perceived health and functioning • HRQoL, measured using FACT-B, which consists of the PWB, SWB, EWB, FWB, and breast cancer-specific concerns. Overall scores range from 0 to 144, where higher score indicates better HRQoL • fatigue, FACT-F scale, which is a 13-item 5-point Likert scale ranging from 0 (not at all) to 4 (very much so), with higher scores indicating lower fatigue levels • spirituality, measured using the FACT-Sp, which has 2 domains, sense of meaning/peace and role of faith, with responses ranging from 1 to 5 on a Likert scale ranging from 0 (not at all) to 4 (very much). Higher scores indicate higher levels of spirituality. Only the sense of meaning/peace subscale was included in this study • depression, measured using the CES-D, which is a 20-item self-report measure. Items are rated on a 4-point scale (0 = rarely or none of the time to 3 = most or all the time) and the total score ranges from 0 to 60. Higher scores indicate greater risk for depression • sleep dysfunction, measured using the PSQI, which is a 19-item self-report measure • positive and negative affect, measured using the PANAS, which is a 20-item measure. Items are scored on a 5-point scale with responses ranging from "very slightly or not at all" to "extremely" Outcomes were measured at baseline and 10 weeks (end of the intervention): • exercise group: n = 22 at baseline, n = 13 at 10 weeks • control group: n = 22 at baseline, n = 14 at 10 weeks Adverse events: cancer recurrence was reported for 4 women in the exercise group and 6 women in the control group. No adverse events were reported
Notes	Country: US Funding: Wake Forest University Comprehensive Cancer Center

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking

Danhauer 2009 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	Data were analyzed on an ITT basis. Participants who failed to return the study questionnaire were excluded from the analyses - 9 participants in the exercise group and 7 participants in the control group did not return the study questionnaire. One participant in the control group withdrew from the study
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Dimeo 2004

Study characteristics	
Methods	Study design: RCT Number randomized: 72; 34 to the exercise group and 35 to the control group Study start and stop dates: not reported Length of intervention: 3 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: lung cancer (n = 27), gastrointestinal cancer (n = 42) Cancer stage, n: • exercise group: Stage I, 10; Stage II, 13; Stage III, 8; Stage IV, 3 • control group: Stage I, 8; Stage II, 15; Stage III, 8; Stage IV, 4 Time since cancer diagnosis, mean (SD) days: • exercise group: 211 (24.5) days • control group: 174 (15.6) days Time beyond active treatment, mean (SD) days: • exercise group: 126 (153) days • control group: 134 (151) days Inclusion criteria: • 30 to 75 years old • ECOG score 0 to 2 • surgical intervention for a histologically confirmed lung or gastrointestinal tumor • understanding of written German Eligibility criterion related to interest or ability to exercise, or both: • none reported Exclusion criteria:

Dimeo 2004 *(Continued)*

- bone metastasis
- diabetes mellitus
- impaired left ventricular function
- coronary heart disease
- liver or kidney dysfunction
- psychiatric condition
- rheumatic disease
- hemoglobin concentration <10 g/dL
- ongoing chemotherapy, radiation therapy or immune therapy

Gender:

- exercise group: 26 male, 9 female
- control group: 25 male, 10 female

Current age, mean (SD) years

- exercise group: 55.1 (10) years
- control group: 60.0 (9.5) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level: not reported

SES: not reported

Employment status: not reported

Comorbidities: not reported

Past exercise history: not reported

On hormone therapy: not reported

Interventions	34 participants assigned to the exercise group, including: • stationary biking Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: 80% of the maximal HR in the stress test Frequency: 5 times per week Duration of session: 30 minutes Duration of exercise program: 3 weeks Total number of exercise sessions: 15 Format: group Facility: facility-based Professionally led: not clear, but supervised by an physician in the same room Adherence: not reported 35 participants assigned to the control group, including: • progressive relaxation training group (45 minutes 3 times per week for 3 weeks)

Dimeo 2004 *(Continued)*

	Contamination of control group: not reported
Outcomes	Outcome: QoL outcomes, using: • fatigue subsection of the EORTC QLQ-C30 Version 2, including 30 questions to evaluate emotional, cognitive, physical and social functioning (function scales), and severity of fatigue, nausea/vomiting, pain, dyspnea, insomnia, appetite loss, and constipation/diarrhea (symptom scales) Subgroups: differences between participants with lung and gastrointestinal tumors. No differences found, so combined data from both groups Outcomes were measured at baseline and 3 weeks (end of the intervention): • exercise group: n = 34 at baseline, n = 31 at 3 weeks • control group, n = 35 at baseline, n = 35 at 3 weeks Adverse events: 3 participants with thrombosis and infection in the exercise group
Notes	Country: Germany Setting: Laboratory Funding: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random number list
Allocation concealment (selection bias)	Low risk	"The randomisation sequence was concealed until assignment of interventions"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Low risk	3 patients in the exercise group were admitted to the hospital for the treatment of a concurrent disease (thrombosis, infection). Data for the 3 patients who did not complete the questionnaire after the intervention were evaluated using the "worst rank assumption"
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	High risk	Demographic information not reported

Dodd 2010

Study characteristics

Methods	Study design: RCT

Dodd 2010 *(Continued)*

	Number randomized: 119; 44 to an exercise group that began exercise during treatment (EE), 36 to an exercise group that began exercise after treatment (CE), and 39 to the control group Study start and stop dates: 1999 to 2005 Length of intervention: 4 to 6 months Length of follow-up: 1 year from baseline
Participants	Type cancer: breast, n = 112; ovarian, n = 6; colon, n = 1 Cancer stage, n (%): • EE group: Stage I, 13 (32.5%); Stage II, 19 (47.5%); Stage III, 8 (20.0%) • CE group: Stage I, 12 (35.3%); Stage II, 18 (52.9%); Stage III, 4 (11.8%) • control group: Stage I, 15 (40.5%); Stage II, 15 (40.5%); Stage III, 7 (18.9%) Time since cancer diagnosis: not reported Time beyond active treatment: unclear Inclusion criteria: • women • ≥ 18 years old • confirmed diagnosis of breast, colorectal, or ovarian cancer • able to read, write, and understand English • willing and able to provide written informed consent • KPS score of ≥ 60 Eligibility criterion related to interest or ability to exercise, or both: • absolute contraindications to exercise testing as established by the ACSM Exclusion criteria: • concurrent radiation therapy or bone marrow transplantation • uncontrolled hypertension or diabetes mellitus • pain intensity score of > 3 on a 0- to 10-point numeric rating scale • lytic bone lesion or orthopedic limitations • history of major depression or sleep disorders • diagnosis of AIDS-related malignancy or leukemia Gender: women Current age, mean (SD) years: • EE group: 49.4 (8.2) years • CE group: 50.4 (9.0) years • control group: 52.0 (10.8) years Age at cancer diagnosis: not reported Ethnicity/race, n (%): • EE group: white, 30 (68.2%); black, 4 (9.1%); Asian, 7 (15.9%); other, 3 (6.8%) • CE group: white, 27 (79.4%); black, 3 (8.8%); Asian, 3 (8.8%); other, 1 (2.9%) • control group: white, 31 (79.5%); black, 5 (12.8%); Asian, 2 (5.1%); other, 1 (2.6%) Education level: not reported SES, ≥ USD40,000, n (%):

Dodd 2010 *(Continued)*

- EE group: 35 (83.3%)
- CE group: 30 (90.9%)
- control group: 29 (76.3%)

Employment status, employed full or part time, n (%):

- EE group: 17 (41.5%)
- CE group: 19 (57.6%)
- control group: 14 (36.0%)

Comorbidities: not reported

Past exercise history, participation in regular exercise, n (%):

- EE group: 29 (67.4%)
- CE group: 24 (70.6%)
- control group: 22 (56.4%)

On hormone therapy: not reported

Interventions	80 participants assigned to the exercise intervention (44 in EE group and 36 in CE group), including: • individualized program adjusted to participant's fitness level and adjusted weekly to maintain the exercise prescription. The program consisted of a cardiovascular/aerobic exercise (e.g. walking, jogging, or cycling) Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: targeted to HR corresponding to 60% to 80% $VO_{2\ peak}$, and to achieve the Borg Scale of 12- to 14-point level ("somewhat hard"). Frequency: 3 to 5 times per week Duration of individual sessions: 20 to 30 minutes of continuous exercise Duration of exercise program: 4 to 6 months Total number of exercise sessions: not reported, but varied Format: individual Facility: home based Professionally supervised by exercise physiologist 39 participants assigned to control group, including: • usual care Adherence: the EE group reported an adherence rate of 73% at end of intervention and 75.7% at end of follow-up, and the CE group reported 86.7% adherence at end of intervention Contamination of control group: not reported
Outcomes	No primary outcome was identified. QoL outcomes included: • fatigue, measured by the Revised Piper Fatigue Scale • sleep dysfunction, measured by the General Sleep Disturbance Scale • depression, measured by the CES-D • pain, measured by the Worst Pain Intensity Scale Physical performance was measured using the KPS scale Outcomes were measured at baseline, 4 to 6 months (end of intervention) and 1 year:

Dodd 2010 *(Continued)*

- EE group: n = 44 at baseline, n = 39 at 4 to 6 months, n = 39 at 1 year
- CE group: n = 36 at baseline, n = 35 at 4 to 6 months, n = 35 at end of 1 year
- control group: n = 39 at baseline, n = 38 at 4 to 6 months, n = 38 at 1 year

Analyses were completed on 37 women in the EE group, 32 women in the CE group, and 37 women in the control group

Subgroup analysis: none reported

Adverse events:

- EE and CE groups: hip pain, sciatica (n = 16), arm discomfort (n = 4), knee discomfort (n = 10), ankle discomfort (n = 3), and foot discomfort (n = 8)
- asymptomatic ischemic electrocardiogram changes (i.e. ST-segment depression, n = 10), asymptomatic bigeminy (n = 6), and premature ventricular complexes (n = 9). Abnormal findings were sent and reviewed by participant's primary physician, oncologist, cardiologist, or a combination and 8 participants were discontinued from the study

Notes	Country: US Funding: National Cancer Institute; Clinical & Translational Science Institute, Clinical Research Center

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	Several participants in each of the study groups were excluded from the analyses. There was no ITT analysis and it is unclear how missing data were handled
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Donnelly 2011

Study characteristics

Methods	Study design: RCT

Donnelly 2011 *(Continued)*

	Number randomized: 33; 16 to the exercise group and 17 to the control group Study start and stop dates: recruitment form June 2008 to March 2009 Length of intervention: 12 weeks Length of follow-up: 6 months
Participants	Type cancer, gynecologic cancers (ovarian, endometrial, uterine, cervical, or mixed), n (%): • exercise group: ovarian, 6 (37.5%); endometrial, 6 (37.5%); uterine, 1 (6.3%); cervical, 2 (12.5%); mixed, 1 (6.3%) • control group: ovarian, 6 (35.3%); endometrial, 5 (29.4%); uterine, 3 (17.6%); cervical, 2 (11.8%); mixed, 1 (5.9%) Cancer stage, stage I-III, n (%): • exercise group: Stage I, 7 (43.8%); stage II, 7 (43.8%); stage III, 2 (12.5%) • control group: Stage I, 9 (52.9%); stage II, 3 (17.6%); stage III, 5 (29.4%) Time since cancer diagnosis, mean (SD) months: • exercise group: 8.7 (9.6) months • control group: 8.6 (8.9) months Time beyond active treatment: some women still receiving treatment Inclusion criteria: • women • ≥ 18 years old • diagnosis of gynecologic cancer (stage I to III) • completed surgery and either undergoing or completed anticancer treatment • within 3 years of diagnosis • report of mild to severe fatigue Eligibility criterion related to interest or ability to exercise, or both: • currently sedentary (i.e. vigorous physical activity < 20 minutes/week or moderate physical activity < 60 minutes/week for the past 6 months) was inclusionary Exclusion criteria: • current medical or psychiatric illness (i.e. unstable cardiovascular disease, uncontrolled hypertension, diabetes or respiratory disease, severe mental illness, cognitive dysfunction or orthopedic problems) • participation in other intervention trials • previous diagnosis of cancer • another fatigue-related comorbidity (fibromyalgia, chronic fatigue syndrome, multiple sclerosis, myalgic encephalopathy, lupus, or arthritis) Gender: women Current age, mean (SD) years: • exercise group: 53.5 (8.7) years • control group: 52.1 (11.8) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: not reported

Donnelly 2011 *(Continued)*

	SES: not reported Employment status, n (%): • exercise group: full/part-time, 1 (6.3%); sick leave, 9 (56.3%); housewife, 2 (12.5%); retired, 4 (25.0%) • control group: full/part-time, 3 (17.6%); sick leave, 8 (47.1%); housewife, 2 (11.8%); retired, 4 (23.5%) Comorbidities: not reported Past exercise history: not reported On hormone therapy: not reported
Interventions	16 participants assigned to the exercise intervention, including: • physical activity, including walking and strengthening exercises, implemented by an initial, individual face-to-face consultation with a physical therapist and physical activity consultations guidelines followed by weekly telephone calls for 10 weeks, a final face-to-face consultation at week 12 and 2 monthly follow-up calls Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: moderate Frequency: aim to meet physical activity guidelines (30 minutes of physical activity on at least 5 days a week) Duration of individual sessions: 30 minutes Duration of exercise program: 12 weeks Total number of exercise sessions: maximum of 60 Format: individual Facility: home based Professionally led with initial consultation with a professional physical therapist 17 participants assigned to control group, including: • usual care • telephone calls at same time and length as exercise group Adherence: 44% of all participants, or 58% of all individuals who remained medically unfit to take part Contamination of control group: unclear
Outcomes	Primary outcome: • fatigue, measured using the MFSI-SF and FACIT-F subscale Secondary outcomes: • QoL measured using the FACT-G scale • depression measured using the BDI-II • positive and negative affect measured using the PANAS • sleep dysfunction measured using the PSQI Outcomes were measured at baseline, 12 weeks (end of intervention) and 6 months' follow-up (9 months after baseline): • exercise group: n = 16 at baseline, n = 15 at 12 weeks, n = 12 at 6 months' follow-up • control group: n = 17 at baseline, n = 17 at 12 weeks, n = 17 at 6 months' follow-up

Donnelly 2011 *(Continued)*

	Subgroup analysis: none reported Adverse events: • exercise group: lung metastasis (n = 1), pulmonary embolism (n = 1), heart palpitations (n = 1) • control group: none reported
Notes	Country: UK Funding: Department of Employment and Learning, Northern Ireland

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers table was used to generate the allocation sequence
Allocation concealment (selection bias)	Low risk	Allocation was concealed in sequentially numbered opaque envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The outcome assessor was blinded to the study allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	The study used ITT analyses
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Fillion 2008

Study characteristics

Methods	Study design: RCT Number randomized: 94; 48 to the exercise group and 46 to the control group Study start and stop dates: not reported Length of intervention: 4 weeks Length of follow-up: 3 months
Participants	Type cancer: breast cancer, stages 0 to III Time since cancer diagnosis, mean (SD) days:

Fillion 2008 *(Continued)*

- exercise group: 256.7 (101.5) days
- control group: 256.8 (112.7) days

Time beyond active treatment: no more than 2 years

Inclusion criteria:

- women diagnosed with an initial nonmetastatic breast cancer
- completion of initial breast cancer treatment no longer than 2 years before enrolment
- received 1 series of adjuvant treatments of radiation therapy, or had received radiation therapy in combination with other adjuvant treatments (e.g. chemotherapy or hormonal therapy)
- understand and speak French
- live near the cancer center and available to take part in a series of 4 weekly sessions
- accept the randomization procedure

Eligibility criterion related to interest or ability to exercise, or both:

- pass revised Physical Activity Readiness Medical Examination
- obtain the authorization of supervising physician before performing the fitness assessment

Exclusion criteria:

- clinical levels of depression symptoms, as measured by HADS (score > 10)
- insomnia, as defined by the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
- any symptom of cancer recurrence
- known severe health problems other than cancer

Gender: female

Current age, mean (SD) years:

- exercise group: 53.09 (9.65) years
- control group: 51.84 (10.25) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level, n (%):

- exercise group: high school, 13 (29.5%); college graduate, 13 (29.5%); university graduate, 18 (40.9%)
- control group: high school, 15 (34.9%); college graduate, 13 (30.2%); university graduate, 15 (34.9%)

SES, n (%):

- exercise group: < USD15,000, 2 (5.0%); USD15,000 to USD29,999, 1 (2.5%); USD30,000 to USD44,999, 6 (15.0%); > USD49,999, 31 (77.5%)
- control group: < USD15,000, 3 (7.5%); USD15,000 to USD29,999, 6 (15.0%); USD30,000 to USD44,999, 5 (12.5%); > USD49,999, 26 (65.0%)

Employment status, n (%):

- exercise group: full-time, part-time, 8 (12.8%); absence due to illness, retired, unemployed, 36 (81.8%)
- control group: full-time, part-time, 16 (37.2%); absence due to illness, retired, unemployed, 27 (62.8%)

Comorbidities: not reported

Past exercise history: not reported

On hormone therapy, tamoxifen, nolvadex, zoladex, arimidex, n (%):

- exercise group: 29 (65.9%)

Fillion 2008 *(Continued)*

	• control group: 35 (81.4%)
Interventions	48 participants assigned to the exercise intervention, including: • 4 weekly group meetings of 2.5 hours and 1 short telephone booster session (5 to 15 minutes). 1 hour was devoted to the supervision of walking training by a kinesiologist or a trained research nurse • 1.5 hours to the psycho-educative, fatigue management sessions Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: unclear Frequency: 4 times per week Duration of exercise session: 1 hour Duration of program: 4 weeks Total number of exercise sessions: 16 Format: group Facility: facility and home Professionally led: kinesiologist led the exercise, and an oncology nurse led the psycho-educational component Adherence: 45 of 48 participants completed the full treatment Co-intervention: psycho-educative, fatigue management 46 participants assigned to the control group, including • normal activity Contamination of control group: not reported
Outcomes	Primary outcome: fatigue, measured with the General/Physical Fatigue subscale of the MFI Secondary outcomes: physical measures and QoL measures, including: • fitness, measured as submaximal oxygen consumption ($Vo_{2submax}$), was estimated from the Single-Stage Treadmill Walking Test • QoL, measured with the MOS SF-12 • energy level, measured using the Vigor subscale of the shortened Profile of Mood States • anxiety and depression, measured using the Profile of Mood States Outcomes were measured at baseline, 4 weeks, and 3 months: • exercise group: n = 48 at baseline, n = 45 at 4 weeks, n = 45 at 3 months • control group: n = 46 at baseline, n = 43 at 4 weeks, n = 43 at 3 months Adverse events, cancer recurrence, n: exercise group: 2 control group: 1
Notes	Country: Canada Funding: BFonds de Recherche en Sante du Quebec, Investigator Award
Risk of bias	

Fillion 2008 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	sequence of randomization was "computer generated"
Allocation concealment (selection bias)	Low risk	"sealed envelopes, which were concealed to both kinesiologist and patient"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	4 participants from the exercise group were not included in the analyses (withdrew, n = 1; cancer recurrence, n = 2; metastatic breast cancer diagnosis, n = 1); 3 participants from the control group were not included in the analyses (withdrew, n = 2; cancer recurrence, n = 10)
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Heim 2007

Study characteristics	
Methods	Study design: quasi-RCT Number randomized: 63; 32 assigned to the exercise group and 31 to the control group Study start and stop dates: not reported Length of intervention: unclear Length of follow-up: 3 months' postrehabilitation
Participants	Type cancer: breast cancer Time since cancer diagnosis: not reported Time beyond active treatment: not reported, but at least 6 weeks since surgery or chemotherapy Inclusion criteria: • score of 4 or more on a linear analog scale evaluating fatigue ranging in value from 0 to 10 Eligibility criterion related to interest or ability to exercise, or both: not reported Exclusion criteria: • psychiatric condition • < 6 weeks since surgery or chemotherapy

Heim 2007 *(Continued)*

	Gender: female Current age, n (%): • exercise group: 31 to 50 years, 14 (44%); 51 to 70 years, 18 (56%) • control group: 31 to 50 years, 18 (58%); 51 to 70 years, 13 (42%) Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: not reported SES: not reported Employment status, n (%): • exercise group: working, professional life, 19 (59%) • control group: working, professional life, 23 (74%) Comorbidities: not reported Past exercise history (before disease), n (%): • exercise group: no sports, 8 (25%); < 1 hour per week, 11 (34%); 1 to 2 hours per week, 11 (34%); ≥ 3 hours per week, 2 (6%) • control group: no sports, 8 (26%); < 1 hour per week, 10 (32%); 1 to 2 hours per week, 10 (32%); ≥ 3 hours per week, 3 (10%) On hormone therapy: none
Interventions	32 participants assigned to the exercise intervention, including: • educational program, physical therapy, group exercise, and psycho-oncologic interventions • brochure with instructions for 9 muscle strength and 9 stretching exercises for all large muscle groups, demonstrated by instructor • instructions for aerobic exercises (walking program), coordination, and relaxation Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: not reported Frequency: instructions were to complete strength training 3 times per week and aerobic exercise for 30 minutes twice per week Duration of exercise session: not reported Duration of exercise program: not reported Total number of exercise sessions: not reported Format: individual Facility: inpatient rehabilitation center, but exercises could also be completed at home Professionally led: initial instruction and printed brochures, but no further instruction Adherence: assessed as percentage (where adherence to program was equal to 100%), adherence to: • muscle strength was 26% at end of rehabilitation and 37% at 3 months after rehabilitation • stretching was 30% at end of rehabilitation and 42% at 3 months after rehabilitation • aerobic exercises were 163% at end of rehabilitation and 192% at 3 months after rehabilitation 31 participants assigned to control group, including:

Heim 2007 *(Continued)*

	• educational program, physical therapy, group exercise, and psycho-oncological interventions Contamination of control group: not reported, although this group received group exercises
Outcomes	Outcomes: QoL outcomes, including: • FACIT • HADS • MFI • questionnaire on physical activity and motivation to perform exercises and sport • aerobic capacity, using Harvard step test • muscular strength, using Digimax Multifunktionstest Outcomes were measured at baseline, after rehabilitation and at 3 months, 59 participants with complete data: • exercise group: number at baseline not reported, n = 32 at 3 months • control group: number at baseline not reported, n = 31 at 3 months • total number of participants at 6 months = 59, not reported by group assignment Adverse events: not reported
Notes	Country: Germany Funding: German Fatigue Society

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"according to their admission to hospital; depending on the alternating weeks they were allocated to the intervention group or the control group."
Allocation concealment (selection bias)	High risk	Because of alternation, the investigators were aware of the next treatment assignment
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	There were complete data packets for 59 participants, but no information on missing patients. Also, "More patients in the control group (15) than in the training group (12) did not continue the study"
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Unclear risk	Poorly described study

Herrero 2006

Study characteristics	
Methods	Study design: RCT Number randomized: 20, 10 to the exercise group and 10 to the control group Study start and stop dates: recruitment was from November 2003 to April 2004 Length of intervention: 8 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: Stage I or II ductal breast cancer, stage at diagnosis not reported Time since cancer diagnosis: not reported Time beyond active treatment, mean (SD) months: • exercise group: 36 (13) months • control group: 35 (12) months Inclusion criteria: • postmenopausal women surviving breast cancer • 2 to 5 years post-treatment • 40 to 60 years old • previous anticancer treatment consisting of surgery with axillary lymphadenectomy and both post-surgery radiation therapy and chemotherapy Eligibility criterion related to interest or ability to exercise, or both: • physical activity level: walking less than a total of 30 to 60 minutes 3 days per week • performing no strenuous exercise such as running, cycling, swimming, or resistance training Exclusion criteria: • cardiac disease (NYHA II or greater) • uncontrolled hypertension (blood pressure > 160/90 mmHg) • uncontrolled pain, or any other condition that contraindicated exercise training in cancer patients or survivors, for example increased risk of bone fractures • severe anemia (< 8 g/dL) • platelet count lower than 50 x 10^9/µL • lymphedema Gender: female Current age, mean (SD) years: • exercise group: 50 (5) years • control group: 51 (10) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level: not reported SES: not reported Employment status: not reported Comorbidities: not reported

Herrero 2006 *(Continued)*

	Past exercise history: limited by exclusion criteria On hormone therapy: not reported Body mass and BMI, mean (SD): • exercise group: body mass, 66.7 (10.5) kg; BMI, 24.0 (3.2) kg/m^2 • control group: body mass, 67.7 (8.9) kg; BMI, 25.1 (3.5) kg/m^2

Interventions	10 participants assigned to an exercise group, including:
	
	• 10-minute warm-up and cool-down period, consisting of:
	  • cycle-ergometer pedaling at very light workloads
	  • stretching exercises for all major muscle groups
	• 70-minute core portion of the training session
	  * resistance training with 11 exercises engaging the major muscle groups (chest press, shoulder press, leg extension, leg curl, leg press, leg calf rise, abdominal crunch, low back extension, arm curl, arm extension, and lateral pull-down)
	  * exercises performed through the full range of motion normally associated with correct technique for each exercise
	  * stretching of muscles involved in an exercise performed at the end of each set of resistance exercise
	• blood total creatine kinase levels were measured every week to ensure that the training program did not induce excessive muscle damage, that is levels < 167 International units
	
	Type exercise (aerobic/anaerobic): aerobic and anaerobic
	
	Intensity of experimental exercise intervention: not reported
	
	Frequency: 3 times per week
	
	Duration of sessions: 90 minutes
	
	Duration of program: 8 weeks
	
	Total number of exercise sessions: 24
	
	Format: not reported
	
	Facility: community fitness club (Miranda de Ebro, Spain)
	
	Professionally led: supervised by experienced investigator
	
	Adherence: mean (SD) % = 91.1% (6.9%);
	
	Control group: 10 participants assigned to:
	
	• usual activities with no moderate to heavy exercise
	
	Contamination of control group: not reported

Outcomes	Primary outcome: physical and QoL outcomes, including: • cardiorespiratory test to measure peak oxygen uptake (peak VO_2) • dynamic strength endurance test, maximum number of repetitions for chest and leg press exercise at 30% to 35% and 100% to 110% of body mass • sit-stand test, frequency count per time • EORTC QLQ-C30 questionnaire, a 30-item questionnaire on physical, role, social, emotional, cognitive and functioning, and a global scale of QoL (maximum score of 100). The physical and global scores were assessed • hematocrit and hemoglobin level Secondary outcomes, included:

Herrero 2006 *(Continued)*

	• body composition, assessed indirectly through changes in body mass and subcutaneous skinfolds. Skinfold measurements were made at 3 sites (triceps, abdominal, and supra iliac) to estimate percentage of body fat • total muscle mass (kg), estimated from anthropometrical data following the prediction equation using multislice magnetic resonance imaging Outcomes were measured at baseline and 8 weeks (end of the intervention): • exercise group: n = 10 at baseline, n = 8 at 8 weeks • control group: n = 10 at baseline, n = 8 at 8 weeks Adverse events: none reported
Notes	Country: Spain Funding: Universidad Europea de Madrid

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Low risk	"The treatment allocation system was set up so that the researcher who was in charge of enrolling participants did not know in advance which treatment the next person would get"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"Research assistants (exercise physiologists) with no knowledge of group assignments were designated to measure the outcome variables"
Incomplete outcome data (attrition bias) All outcomes	High risk	2 participants in each group withdrew. No information provided on withdrawals
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Knols 2011

Study characteristics

Methods	Study design: RCT Number randomized: 131; 64 to the exercise group and 67 to the control group Study start and stop dates: enrolment from January 2005 to November 2008 Length of intervention: 12 weeks

Knols 2011 *(Continued)*

	Length of follow-up: 3 months from end of the intervention
Participants	Type cancer, n (%): • exercise group: leukemia (AML), 19 (29.7%); chronic lymphocytic leukemia, 5 (7.8%); leukemia (ALL), 0 (0%); HL, 5 (7.8%); NHL, 11 (17.2%); multiple myeloma, 17 (26.6%); osteomyelofibrosis, 3 (4.7%); amyloidosis, 1 (1.6%); testicular cancer, 3 (4.7%) • control group: leukemia (AML), 12 (17.9%), chronic lymphocytic leukemia, 9 (13.4%); leukemia (ALL), 2 (3%); HL, 9 (13.4%); NHL, 14 (20.9%); multiple myeloma, 20 (29.9%); osteomyelofibrosis, 1 (1.5%); amyloidosis, 0 (0%); testicular cancer, 0 (0%) Time since cancer diagnosis: not reported Time between HSCT and study, mean (SD) days: • exercise group: 81 (36) days • control group: 78 (35) days Inclusion criteria: • male or females • ≥ 18 years old • basic fluency in the German language • 3 weeks up to 6 months after autologous or allogenic HSCT Eligibility criterion related to interest or ability to exercise, or both: • none reported Exclusion criteria: • GVHD (except for grade I not requiring treatment) • painful joints • unstable osteolyses • chronic pain • lesions in the central or peripheral nervous system • uncontrolled cardiovascular disease, thyroid disease, or diabetes. Gender, n (%): • exercise group: male, 38 (59.4%); female, 26 (40.6%) • control group: male, 39 (58.2%); female, 28 (41.8%) Current age, mean (SD) years: • exercise group: 46.7 (13.7) years • control group: 56.6 (12.0) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level, n (%): • exercise group: secondary school, 8 (12.5%); vocational, 26 (40.6%); higher professional, 12 (18.8%); college/university, 18 (28.1%) • control group: secondary school, 10 (14.9%); vocational, 30 (44.8%); higher professional, 9 (13.4%); college/university, 18 (26.9%) SES: not reported Employment status: not reported

Knols 2011 *(Continued)*

	Comorbidities: not reported Past exercise history, n (%): • exercise group: almost completely inactive, 25 (39.1%); minimum 20 minutes walking/cycling per day, 39 (60.9%). • control group: almost completely inactive, 19 (28.4%); minimum 20 minute walking/cycling per day, 48 (71.6%) On hormone therapy: not reported Other, BMI, mean (SD, range): • exercise group: 22.9 (4.3, 15 to 38) • control group: 23.9 (4.0, 14 to 34)
Interventions	64 participants assigned to the exercise intervention, including a supervised physical exercise program with: • endurance exercises, including ergometer cycling, starting with a 10-minute warm-up and maintenance of aerobic performance for at least 20 minutes • progressive resistance training using dumbbells and a standard strength program including squats, step-ups and -downs, barbell rotations and upright rowing. The program could be extended with chest press, triceps extension, biceps curl, modified curl ups, and calf raises Type exercise (aerobic/anaerobic): aerobic and anaerobic Intensity of experimental exercise intervention: 20 minutes at a predefined individual HR (from 50% to 60%, increasing up to 70% to 80% of the estimated maximum HR) Frequency: twice per week Duration of individual sessions: 10 minutes warm-up, 20 minutes maintenance plus time for resistance training Duration of exercise program: 12 weeks Total number of exercise sessions: 24 Format: unclear if individual or group Facility: facility base in a physical therapy practice or fitness center Professionally led by physiotherapist or a physical trainer 67 participants assigned to the control group, including: • usual care Adherence: average participation in the physical exercise program was 85% (range = 21% to 100%), representing about 20.5 of 24 training sessions. 22 patients attended all (100%) of the sessions; 37.5% attended > 80% of the sessions; 17.2% attended ≥ 66% of the sessions; and 10.9% attended < 66% of the sessions Contamination of control group: 7.5% of the control group patients reported a minimum of 20 physical exercise sessions
Outcomes	Primary outcome included QoL and physical outcomes, including: • physical functioning, measured using the physical function subscale of the EORTC QLQ-C30 • physical function measures, including knee extension, grip strength, walking speed, and functional exercise capacity Secondary QoL outcomes included:

Knols 2011 *(Continued)*

- fatigue, measured using the German language version of the fatigue subscale of FACT-An
- fatigue, measured using the QLQ-C30 subscale
- role function, measured using the QLQ-C30 subscale
- cognitive function, measured using the QLQ-C30 subscale
- social function, measured using the QLQ-C30 subscale
- pain, measured using the QLQ-C30 subscale
- insomnia, measured using the QLQ-C30 subscale

Secondary physical function outcomes included:

- body composition
- quantified walking activity
- physical activity, measured using the International Physical Activity Questionnaire (IPAQ) short form, telephone-version

Outcomes were measured at baseline, 12 weeks, and 3 months:

- exercise group: n = 64 at baseline, n = 57 at 12 weeks, n = 51 at 3 months
- control group: n = 62 at baseline, n = 57 at 12 weeks, n = 54 at 3 months

Subgroup analysis: several subgroup analyses were prespecified and conducted

Adverse events: none reported

Notes	Country: Switzerland Funding: Zurcher Kresliga (Zurich) and the Eidenossiche Sportkommission (Magglingen)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Sequence generated using a minimization procedure
Allocation concealment (selection bias)	Low risk	Results of the randomisation were "...stored in opaque envelopes. The allocation sequence and contents of the envelopes were concealed by study personnel"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, participants and study personnel could not be masked or blinded to the allocation to the intervention
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Independent assessors of physical outcomes were blinded to group assignments and carried out the assessments
Incomplete outcome data (attrition bias) All outcomes	Low risk	In the exercise group, 7 participants lost by end of intervention and additional 6 from end of intervention to end of the follow-up. In control group, 10 participants lost by end of 12 weeks and additional 3 at end of follow-up. Investigators included all study participants in the ITT analyses
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

McNeely 2008a

Study characteristics	
Methods	Study design: RCT Number randomized: 52; 27 to the exercise group and 25 to the control group Study start and stop dates: recruitment from October 1, 2005 to October 31, 2006 Length of intervention: 12 weeks Length of follow-up: to end of the intervention

Participants	Type cancer: HNC
	
	Cancer stage, n (%):
	
	• exercise group: Stage I, 2 (7%); Stage II, 3 (11%); Stage III, 9 (33%); Stage IV, 12 (44%)
	• control group: Stage I, 1 (4%); Stage II, 3 (12%); Stage III, 3 (12%); Stage IV, 18 (72%)
	
	Time since cancer diagnosis: not reported
	
	Time beyond active treatment, median (range) months:
	
	• exercise group: 12 (2 to 120) months
	• control group: 17 (2 to 180) months
	
	Inclusion criteria:
	
	• surgical treatment, including radical neck dissection, modified radical neck dissection, and other variants of selective neck dissection
	• KPS score of at least 60%
	• no evidence of residual cancer in the neck and no distant metastasis
	• completion of adjuvant HNC treatment
	• symptoms of shoulder dysfunction attributed to spinal accessory nerve damage, with ≥ 3 of the following signs:
	  * atrophy of the upper trapezius muscle
	  * shoulder droop
	  * scapular misalignment
	  * winging of the scapula with elevation of the arm
	  * limitation in shoulder abduction range of motion
	
	Eligibility criterion related to interest or ability to exercise, or both: none reported
	
	Exclusion criteria:
	
	• history of shoulder or neck pathology unrelated to cancer treatment
	• comorbid medical illness or psychiatric illness that would prevent completion of treatment or interfere with follow-up
	
	Gender, n (%):
	
	• exercise group: men, 20 (74%); women, 7 (26%)
	• control group: men, 17 (68%); women, 8 (32%)
	
	Current age, mean (range) years:
	
	• exercise group: 53 (32 to 76) years
	• control group: 57 (43 to 76) years

McNeely 2008a *(Continued)*

	Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level, completed university, n (%): • exercise group: completed university, 12 (44%) • control group: completed university, 14 (56%) Income, > USD80,000/year: • exercise group: > USD80,000/year, 9 (33%) • control group: > USD80,000/year, 12 (48%) SES, on disability, n (%): • exercise group: on disability, 11 (41%) • control group: on disability, 9 (36%) Employment status: not reported Past exercise history, report currently exercising, n (%): • exercise group: report currently exercising, 4 (15%) • control group: report currently exercising, 4 (16%) On hormone therapy: not reported

Interventions	27 participants assigned to the exercise intervention, including:
	
	• PRET
	  * supervised active and passive range of motion/stretching exercises, postural exercises, and basic strengthening exercises with light weights (1 to 5 kg) and elastic resistance bands
	  * strengthening exercises tailored for each individual to provide progressive overload of the following muscle groups: rhomboids/middle trapezius; levator scapula/upper trapezius; biceps; and triceps, deltoid, and pectoralis major, consisting of 2 sets of 10 to 15 repetitions of 5 to 8 exercise
	
	Type exercise (aerobic/anaerobic): aerobic and anaerobic
	
	Intensity of experimental exercise intervention: PRET = starting at 25% to 30% of their 1-RM strength and slowly progressing to 60% to 70% of their 1-RM strength by the end of the intervention period
	
	Frequency: minimum of 2 supervised sessions per week (with the option of a third session at the center or at home) for the 12-week intervention period
	
	Duration of session: not reported
	
	Duration of program: 12 weeks
	
	Total number of exercise sessions: 24 to 36 sessions
	
	Format: unclear
	
	Facility: facility twice a week and home or facility once per week
	
	Professionally led: physical therapist with experience working with HNC survivors provided intervention for both groups
	
	25 participants assigned to the control group, including:
	
	• supervised active and passive ROM/stretching exercises, postural exercises, and basic strengthening exercises with light weights (1 to 5 kg) and elastic resistance bands. Strengthening exercises focused on the following muscle groups: rhomboids/middle trapezius; levator scapula/upper trapezius; biceps; and triceps, deltoid, and pectoralis major

McNeely 2008a *(Continued)*

	Adherence: follow-up assessment for the primary outcome was 92% • exercise group: 95% • control group: 87% Contamination of control group: unclear whether the control group engaged in any exercise
Outcomes	Primary outcome: patient-rated shoulder pain and disability • change in patient-rated shoulder pain, assessed using the SPADI, based on a score from 0 to 100 with higher scores indicating more pain/disability • change in shoulder disability, assessed using the Neck Dissection Impairment index, which provides a measure of treatment-specific QoL and includes 10 questions, scored from 1(a lot) to 5 (not at all). Higher scores reflecting a greater impact on QoL Secondary outcomes: QoL and fatigue outcomes, including: • FACT-G • FACT-An subscale • fatigue subscale Outcomes were measured at baseline and 12 weeks (end of the intervention): • exercise group: n = 27 at baseline; at 12 weeks, n = 25 for self-reported outcomes and n = 24 for strength and range of motion. 27 participants included in analysis • control group: n = 25 at baseline; at 12 weeks, n = 23 for self-reported outcomes and n = 22 for strength and range of motion. 25 participants included in analysis Adverse events: • exercise group: colon cancer, n = 1; Soft tissue injury as a result of exercise participation, n = 1; Hospitalization for acute cholecystitis, followed by stroke, n = 1 • control group: recurrence of HNC, n = 2
Notes	Country: Canada Funding: Research Award from the Physiotherapy Foundation of Canada; Full Time Health; Research Studentship from the Alberta Heritage Foundation for Medical Research; Canada Research Chairs Program; Research Team Grant from the NCIC with funds from the CCS and the Sociobehavioural Cancer Research Network

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"An independent researcher generated the allocation sequence by using a computer-generated code"
Allocation concealment (selection bias)	Low risk	"The allocation sequence and contents of the envelopes were enclosed in sequentially numbered an sealed (opaque) envelopes"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions

McNeely 2008a *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	There was significant attrition from the due to adverse effect. The authors conducted ITT analyses by using baseline-observation-carried-forward analyses
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Mehnert 2011

Study characteristics	
Methods	Study design: RCT Number randomized: 63; 35 to the exercise group and 28 to the control group Study start and stop dates: not reported Length of intervention: 10 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: non-metastatic breast cancer, stages I to III Stage of cancer, n (%): • exercise group: Stage I, 17 (56.7%); Stage IIA, 8 (26.7%); Stage IIB, 3 (10.0%); Stage IIIA, 1 (3.3%); Stage IIIB, 1 (3.3%) • control group: Stage I, 13 (46.4%); Stage IIA, 7 (25.0%); Stage IIB, 5 (17.9%); Stage IIIA, 3 (10.7%); Stage IIIB, 0 (0%) Time since cancer diagnosis: not reported Time beyond active treatment: at least 4 weeks following chemotherapy, radiation therapy, or both Inclusion criteria: • 18 to 65 years old • primary nonmetastatic breast cancer • minimum 4 weeks after completion of chemotherapy, radiation therapy, or both Eligibility criterion related to interest or ability to exercise, or both: • any disorder that could interfere with ability to perform the physical exercise program Exclusion criteria: • severe acute or chronic illness other than cancer (e.g. disorders of the musculoskeletal system) Gender: female Current age, mean (SD) years: • exercise group: 53.03 (7.40) years • control group: 50.64 (9.44) years Age at cancer diagnosis: not reported

Mehnert 2011 *(Continued)*

	Ethnicity/race: not reported Education level, n (%): • exercise group: elementary school, 2 (6.9%); junior high school, 17 (58.6%); high school, 2 (6.9%); university degree, 8 (27.6%) • control group: elementary school, 4 (14.3%); junior high school, 11 (39.3%); high school, 6 (21.4%); university degree, 7 (25.0%) SES: not reported Employment status, n (%): • exercise group: employed, 17 (60.7%); retired, 5 (17.8%); housewife, 4 (14.3%); unemployed, 1 (3.6%); unable to work/sick leave, 1 (3.6%) • control group: employed, 14 (50.0%); retired, 3 (10.7%); housewife, 4 (14.3%); unemployed, 6 (21.4%); unable to work/sick leave, 1 (3.6%) Comorbidities: not reported Past exercise history: • exercise group: 19/30 (63.3%) engaged in regular sport and fitness activities for a mean (SD) of 2.3 (1.4) training hours per week • control group: 10/28 (35.7%) engaged in regular sport and fitness activities for a mean (SD) of 1.7 (0.9) training hours per week On hormone therapy: not reported
Interventions	35 participants assigned to a structured physical training program developed to promote muscular strength and exposure, including: • gymnastics, movement games, and relaxation • moderate walking and jogging conducted outside Although 35 women were assigned to the exercise intervention, 5 women refused to participate prior to the baseline assessment Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: maximum of 60% VO_2 Frequency: twice per week Duration of individual sessions: 90 minutes Duration of exercise program: 10 weeks Total number of exercise sessions: 20 Format: group Facility: indoor sports facility and outdoors Professionally led by trained member of study staff (qualified physical therapist or sport therapist) 28 participants assigned to control group, including: • usual care Adherence: not reported Contamination of control group: not reported
Outcomes	No primary outcome was identified. QoL outcomes included:

Mehnert 2011 *(Continued)*

- anxiety, measured using HADS
- depression, measured using HADS
- cancer-specific HRQoL, measured using the EORTC QLQ-C30 subscale
- generic HRQoL, measured using MOS SF-36
- psychological symptoms, measured using SCL-90-R
- body image, measured using a German version of the BIQ

Outcomes were measured at baseline and 10 weeks:

Different numbers of participants had data at the 2 time points as follows:

Anxiety and depression:

- exercise group: n = 30 at baseline, n = 30 at 10 weeks
- control group: n = 28 at baseline, n = 28 at 10 weeks

Individual body Image:

- exercise group: n = 27 at baseline, n = 27 at 10 weeks
- control group: n = 27 at baseline, n = 27 at 10 weeks

Social body image:

- exercise group: n = 30 at baseline, n = 27 at 10 weeks
- control group: n = 27 at baseline, n = 27 at 10 weeks

Numbers of individuals with data for cancer-specific HRQoL, generic HRQoL, and psychological symptoms were not reported

Subgroup analysis: none reported

Adverse events: no cancer recurrences or adverse events reported

Notes	Country: Germany Funding: Friedrich and Louise Homann Foundation, Hamburg, Germany

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	It is unclear how the allocation sequence was generated
Allocation concealment (selection bias)	Low risk	The randomization was adequately concealed through external randomization
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	It is unclear how missing data were handled. 5 randomized participants were reported to have "cancelled" participation in the exercise group

Mehnert 2011 *(Continued)*

Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	High risk	The small sample size and participation in physical exercise by women in the study groups can contribute to bias

Milne 2008a

Study characteristics	
Methods	Study design: cross-over RCT. Only information for first period included here Number randomized: 58; 29 to the immediate exercise group and 29 to the delayed exercise control group Study start and stop dates: recruitment between January 2005 and March 2005 Length of intervention: 12 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer Cancer stage, n (%): • exercise group: Stage I, 6 (20.7%); Stage IIa, 14 (48.3%); Stage IIb, 9 (31.0%); Stage IIIa, 0 (0%) • control group: Stage I, 9 (31.0%); Stage IIa, 11 (37.9%); Stage IIb, 7 (24.1%); Stage IIIa, 2 (6.9%) Time since cancer diagnosis: within 24 months Time beyond active treatment, mean (SD) months: • exercise group: 12.6 (4.62) months • control group: 13.4 (3.4) months Inclusion criteria: • ≥18 years old • English speaking • Stage I to II breast cancer • within 24 months of cancer diagnosis • completion of all treatment except hormone therapy Eligibility criterion related to interest or ability to exercise, or both. Participants were excluded if: • previous engagement in a formal exercise programs for 6 months prior to participation in the study • failed the rPAR-Q Exclusion criteria: • evidence of recurrent disease Gender: female Current age, mean (SD) years: • exercise group: 55.2 (8.4) years • control group: 55.1 (8.0) years Age at cancer diagnosis: not reported

Milne 2008a *(Continued)*

	Ethnicity/race: not reported Education level, university education, n (%): • exercise group: had university education, 11 (37.9%) • control group: had university education, 15 (51.7%) SES: not reported Employment status, part- or full-time employment, n (%): • exercise group: 16 (55.2%) • control group: 16 (55.2%) Comorbidities: not reported Past exercise history: not reported On hormone therapy: 74.1% on hormone therapy
Interventions	29 participants assigned to the immediate exercise group, including: • combined aerobic (cycle and rowing ergometers, mini-trampoline, and step-up blocks) • resistance training (12 different exercises, including chest press, chest extension, biceps curls, triceps extension, leg extension, leg curls, hip abduction and adduction, back extension, abdominal crunches, standing flys, and leg press) • stretching Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: • aerobic: the cardiovascular component was conducted for 20 minutes and ended with a 5-minute cool-down period • resistance: for each exercise, participants performed 2 sets of 10 to 15 repetitions of light weights and progressed to a heavier weight once the current weight and repetitions could be achieved and with good form • stretching: 5 minutes of stretching at the beginning and end of each session Frequency: 3 times per week Duration of sessions: 1 hour Duration of program: 12 weeks Total number of exercise sessions: 36 Format: individual and group Facility: Health and Rehabilitation Program Clinic at The University of Western Australia Professionally led: supervised by exercise physiologists Adherence: average attendance was 60.4% (21.7 of 36 sessions) with a median of 23 (63.9%) and a range of 11 to 36 29 participants assigned to the control group, including: • delayed exercise group, asked not participate in exercise during weeks 1 to 12 • telephone calls at weeks 3, 6, 9, and 12 weeks Contamination of control group: not reported
Outcomes	Primary outcome: QoL outcomes, including:

Milne 2008a *(Continued)*

	• QoL measured using FACT-B scale. Scale description and score range not provided Secondary outcomes, included: • fatigue, measured using the SCFC. Description and score range not provided • Social Physique Anxiety, measured using the SPAS-7 • physical fitness, assessed by SFT Outcomes were measured at baseline, and weeks 6, 12, 18, and 24, except for SFT, which was measured at baseline and 12 weeks only: • exercise group: n = 29 at baseline, n = 29 at 6 weeks, n = 29 at 12 weeks, n = 28 at 18 weeks, n = 28 at 24 weeks • control group: n = 29 at baseline, n = 29 at 6 weeks, n = 29 at 12 weeks, n = 28 at 18 weeks, n = 28 at 24 weeks Adverse events: none reported
Notes	Country: Australia Funding: CCS and the NCIC/CCS Sociobehavioral Cancer Research Network

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomization
Allocation concealment (selection bias)	Low risk	"Group assignments were concealed from the project director who recruited participants to the trial"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Low risk	All study participants completed week 12 assessments and were included in the analysis
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	High risk	Low adherence

Moadel 2007

Study characteristics

Methods	Study design: RCT Number randomized: 164; 108 to a yoga exercise group and to 56 to the control group

Moadel 2007 *(Continued)*

	Study start and stop dates: 2001 to 2005 Length of intervention: 12 weeks Length of follow-up: 1, 3, and 6 months
Participants	Type cancer: breast cancer, stage of disease: • exercise group: Stage I, 42%; Stage II, 36%; Stage III, 17%; Stage IV, 5% • control group: Stage I, 50%; Stage II, 38%; Stage III, 12%; Stage IV, 0% Time since cancer diagnosis, mean (SD, range) years: • exercise group: 1.15 (1.14, 0.06 to 4.06) years • control group: 0.98 (1.13, 0.03 to 4.70) years Time beyond active treatment: receiving chemotherapy,%: • exercise group: at baseline, 30%; at 3 months, 36% • control group: at baseline, 23%; at 3 months, 27% Randomization was stratified by treatment status Inclusion criteria: • ≥ 18 years old • new/recurrent breast cancer (Stages I to III) diagnosis within previous 5 years • high performance status (ECOG performance status < 3) • ability to speak English or Spanish Eligibility criterion related to interest or ability to exercise, or both: • not actively practicing yoga Exclusion criteria: none reported Gender: female Current age, mean (SD, range) years: • exercise group: 55.11 (10.07, 32 to 75) years • control group: 54.23 (9.81, 28 to 71) years Age at cancer diagnosis: not reported Ethnicity/race: • exercise group: African-American, 42%; Hispanic, 30%; non-Hispanic white, 22%; other, 6% • control group: African-American, 43%; Hispanic, 34%; non-Hispanic white, 23%; other, 0% Education level: • exercise group: high school, 69%; college/graduate, 31% • control group: high school, 89%; college/graduate, 11% SES: not reported Employment status: not reported Comorbidities: not reported Past exercise history: not reported On hormone therapy:

Moadel 2007 *(Continued)*

- exercise group: at baseline, 24%; at 3 months, 36%
- control group: at baseline, 41%; at 3 months, 50%

Interventions	108 participants assigned to exercise group, consisting of yoga with each session including: • physical stretches and poses • breathing exercises • meditation All exercises were done in a seated or reclined position. Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: mild Frequency: once per week, but participants were allowed to attend more than 1 session per week and asked to practice yoga at home Duration of sessions: 90 minutes Duration of program: 12 weeks Total number of exercise sessions: 12 sessions Facility: facility and home Professionally led: not reported 56 participants assigned to control group, including: • waiting list Adherence: 26 (31%) participants did not attend any classes, but 8 reported practicing yoga at home at least a few times per week. The mean number of classes attended by active class participants was 7.00 (SD, 3.80) classes. Of 59 participants reporting data, 61% practiced yoga at home at least a few times per week Contamination of control group: not reported

Outcomes	No primary outcome was identified. QoL outcomes included:
	• global QoL, measured using FACT-G and subscales of:
	  * PWB
	  * FWB
	  * EWB
	  * SWB
	• fatigue, assessed using FACIT-F
	• spiritual well-being, assessed using FACIT-Sp
	• mood, assessed using sub-scales of POMS
	Outcomes were measured at baseline and 12 weeks:
	• exercise group: n = 84 at baseline, n = 84 at 12 weeks
	• control group: n = 44 at baseline, n = 44 at 12 weeks
	Subgroup analysis: by treatment status
	Adverse events: none reported

Notes	Country: US Funding: National Cancer Institute, Langeloth Foundation

Moadel 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Generation of allocation sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Investigators stated that they used an "intention-to-treat approach" but it is unclear how the 24 drop-outs in the exercise arm and the 12 drop-outs in the control arm were handled
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Mustian 2004

Study characteristics	
Methods	Study design: RCT Number randomized: 31; 17 to a Tai Chi Chuan exercise group and to 14 to the control group Study start and stop dates: not reported Length of intervention: 12 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, Stage 0 to IIIb Time since cancer diagnosis: not reported Time beyond active treatment: between 1 week and 30 months Inclusion criteria: • female • breast cancer, stage 0 to lllb • post treatment Eligibility criterion related to interest or ability to exercise, or both: • physician's clearance for fitness testing and exercise

Mustian 2004 *(Continued)*

	• engaging in moderate to vigorous physical activity more than once per week • physical limitations prohibiting exercise Exclusion criteria: • clinical mental illness requiring psychotropic drugs, or by self report • presence of catheters or drains Gender: female Current age, mean (SD, range): 52 (9, 33 to 78) years Age at cancer diagnosis: not reported Ethnicity/race: 90% Caucasian Education level: 90% some college SES: > USD40,000 household income, 62% Employment status: employed outside the home, 65% Comorbidities: not reported Past exercise history: not reported On hormone therapy: 56% had received hormonal therapy but specific hormones not reported
Interventions	17 participants assigned to the exercise intervention, including Tai Chi Chuan, comprised of: • warm-up exercises and basic Chi Kung for 10 minutes • Tai Chi Chuan for 40 minutes • 15-move short-form of Yang-style Tai Chi Chuan • regulatory breathing, imagery and meditation for 10 minutes Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: moderate Frequency: 3 times per week Duration of sessions: 60 minutes Duration of program: 12 weeks Total number of exercise sessions: 36 Format: group Facility: facility Professionally led: ACSM-certified health and fitness instructor certified in Tai Chi Control group: 14 participants assigned to the control group, including: • psychosocial support therapy Adherence: • exercise group: 72% exercise rate with 100% compliance • control group: 67% attendance rate with 100% compliance Contamination of control group: 10%
Outcomes	Primary outcome: QoL outcomes, including:

Mustian 2004 *(Continued)*

	• FACIT-F, 28-question survey, scale from 0 to 4 • self esteem assessed by RSE: scoring 1 - strongly agree, 5 - strongly disagree Secondary outcomes: physical outcomes, including: • aerobic capacity, estimated using a 6-minute walk test protocol • muscular strength, evaluated using a handgrip dynamometer to assess the maximal voluntary grip strength • flexibility, assessed using goniometer measurements Outcomes were measured at baseline, 6 weeks, and 12 weeks (immediate postintervention): • exercise group: n = 17 at baseline, n = 11 at 6 weeks, n = 11 at 12 weeks • control group, n = 14 at baseline, n = 10 at 6 weeks, n = 10 at 12 weeks Adverse events: no cancer recurrence reported; cognitive deficits reported as reason for treatment termination
Notes	Country: US Funding: Susan Stout Exercise Science Research Fund, Sally Schindel Cone Women's and Gender Studies Research Fund

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	6 participants in the exercise group and 4 in the control group withdrew and were not included in the analyses
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Oh 2008

Study characteristics

Methods	Study design: RCT

Oh 2008 *(Continued)*

	Number randomized: 30; 15 to the exercise group and 15 to the control group Study start and stop dates: recruitment took place from July 2006 to August 2006 Length of intervention: 8 weeks Length of follow-up: to end of the intervention
Participants	Type cancer, n: • exercise group: breast, 6; ovary, 4; lymphoma, 1; lung, 1; colon, 0; others, 3 • control group: breast, 6; ovary, 2; lymphoma, 1; lung, 1; colon, 3; others, 2 Time since cancer diagnosis: not reported Time beyond active treatment: some patient still undergoing chemotherapy; randomization stratified by whether still being treated or completed therapy Inclusion criteria: • confirmed diagnosis of cancer at any stage • ≥ 18 years old • ECOG performance status of 0 to 3 • expected survival length of > 12 months • ability to complete all study questionnaires and sign the consent form Eligibility criterion related to interest or ability to exercise, or both: • medical contraindication for exercise (e.g. significant orthopedic problem or cardiovascular disease) • already practicing Qigong Exclusion criteria: • diagnosis of other major medical or psychiatric disorder • history of epilepsy, brain metastasis, delirium, or dementia Gender, n:· • exercise group: male, 3; female, 12 • control group: male, 3; female, 12 Age group, n: • exercise group: 36 to 45 years, 2; 46 to 55 years, 4; 56 to 65 years, 3; 66 to 75 years, 6 • control group: 36 to 45 years, 2; 46 to 55 years, 3; 56 to 65 years, 9; 66 to 75 years, 1 Age at cancer diagnosis: not reported Ethnicity, n: • exercise group: Caucasian, 11; Asian, 3; Indigenous Australian, 1 • control group: Caucasian, 14; Asian, 0; Indigenous Australian, 1 Education level, n: • exercise group: primary, 1; secondary, 5; tertiary, 9 • control group: primary, 1; secondary, 4; tertiary, 10 SES: not reported Employment status: not reported Comorbidities: not reported

Oh 2008 *(Continued)*

	Past exercise history: limited by eligibility criteria On hormone therapy: not reported
Interventions	15 participants assigned to exercise group, consisting of medical qigong, with each session including: • 15 minutes of general discussion • 30 minutes of gentle stretching and body movement in standing postures • 15 minutes of movement in seated posture, and • 30 minutes of breathing exercise Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: mild Frequency: once or twice a week for 8 weeks and recommendation to practice at home daily Duration of sessions: 90 minutes, 1 hour for home sessions Duration of program: 8 weeks Total number of exercise sessions: maximum of 16 facility-based and 56 home-based sessions Facility: facility Professionally led: experienced medical qigong instructor who was a Chinese medicine practitioner 15 participants assigned to control group, including: • usual care Adherence: not reported Contamination of control group: not reported

Outcomes	The primary outcomes, QoL and symptom experience, included:
	
	• global QoL, measured using EOTRC QLQ-C30, and subscales of:
	  * physical function
	  * role function
	  * emotional function
	  * cognition function
	  * social function
	  * fatigue
	  * nausea
	  * pain
	  * dyspnea
	  * insomnia
	  * appetite
	  * constipation
	  * diarrhea
	  * perceived financial impact of the disease
	
	Physiologic outcomes included:
	
	• c-reactive protein
	
	Outcomes were measured at baseline and 8 weeks:
	
	• exercise group: n = 15 at baseline, n = 8 at 8 weeks
	• control group: n = 15 at baseline, n = 10 at 8 weeks

Oh 2008 *(Continued)*

	Subgroup analysis: none reported Adverse events: none reported
Notes	Country: Australia Funding: University of Sydney Cancer Research Fund

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomization was done by a computer program"
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	ITT analysis not completed. Of 15 randomized participants in each treatment group, 7 withdrew from the exercise group and 5 from the control group
Selective reporting (reporting bias)	Unclear risk	There is no evidence of selective reporting of outcomes
Other bias	Unclear risk	Small sample size can put study at risk of bias

Oh 2010

Study characteristics

Methods	Study design: RCT Number randomized: 162; 79 to the exercise group and 83 to the control group Study start and stop dates: first recruitment phase was between July 2006 and August 2007 and the second recruitment phase was from August 2007 to May 2008 Length of intervention: 10 weeks Length of follow-up: to end of the intervention
Participants	Type cancer, n (%): • exercise group: breast, 26 (37.7%); lung, 6 (8.7%); prostate, 8 (11.6%); colorectal/bowel, 8 (11.6%); others, 23 (33.3%) • control group: breast, 21 (30.9%); lung, 3 (4.4%); prostate, 4 (5.9%); colorectal/bowel, 8 (11.8%); others, 32 (47.1%)

Oh 2010 *(Continued)*

	Time since cancer diagnosis: not reported Time beyond active treatment: 36 (47.4%) of patients in the intervention group still undergoing cancer treatment and 34 (45.9%) in the control group; randomization stratified by whether still being treated or completed therapy Inclusion criteria: • confirmed diagnosis of malignancy at any stage • ≥ 18 years old • expected survival length of > 12 months Eligibility criterion related to interest or ability to exercise, or both: • medical contraindication for exercise (e.g. significant orthopedic problem or cardiovascular disease) • already practicing Qigong Exclusion criteria: • diagnosis of other major medical or psychiatric disorder • history of epilepsy, brain metastasis, delirium, or dementia Gender, n (%):· • exercise group: male, 31 (39.2%); female, 48 (60.8%) • control group: male, 38 (45.8); female, 45 (54.2%) Age, mean (SD) years: • exercise group: 60.1 (11.7) years • control group: 59.9 (11.3) years Age at cancer diagnosis: not reported Ethnicity, n (%): • exercise group: Caucasian, 57 (77.0%); Asian, 10 (13.5%); Indigenous Australian, 1 (1.4%); other, 6 (8.1%) • control group: Caucasian, 49 (64.5%); Asian, 17 (22.4%); Indigenous Australian, 1 (1.3%); other, 9 (11.8%) Education level, n (%): • exercise group: primary, 1 (1.3%); secondary, 35 (45.5%); undergraduate, 19 (24.7%); postgraduate, 22 (28.6%) • control group: primary, 7 (9.2%); secondary, 34 (44.7%); undergraduate, 19 (25.0%); postgraduate, 16 (21.1%) SES: not reported Employment status: not reported Comorbidities: not reported Past exercise history: limited by eligibility criteria On hormone therapy: not reported
Interventions	79 participants assigned to exercise group, consisting of medical Qigong, with each session including: • 15 minutes of general discussion • 30 minutes of gentle stretching and body movement in standing postures • 15 minutes of movement in seated posture, and

Oh 2010 *(Continued)*

	• 30 minutes of meditation and including breathing exercises Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: mild Frequency: twice per week for 10 weeks and recommendation to practice at home daily Duration of sessions: 90 minutes for supervised sessions, 30 minutes for home sessions Duration of program: 10 weeks Total number of exercise sessions: maximum of 20 facility-based and 70 home-based sessions Facility: facility Professionally led: experienced medical Qigong instructor who was a Chinese medicine practitioner 83 participants assigned to control group, including: • usual care Adherence: not reported Contamination of control group: not reported

Outcomes	The primary outcome of QoL included:
	
	• QoL, measured using the FACT-G, and subscales of:
	  * PWB
	  * SWB
	  * emotional well-being
	  * FWB
	
	Secondary outcomes includes:
	
	• fatigue, measured using the FACT-F scale
	• mood, measured using the POMS and subscales of:
	  * tension and anxiety
	  * depression
	  * anger and hostility
	  * lack of vigor
	  * fatigue
	  * confusion
	
	Physiologic outcomes included:
	
	• C-reactive protein
	
	Outcomes were measured at baseline and 10 weeks:
	
	• exercise group: n = 79 at baseline, n = 54 at 10 weeks
	• control group: n = 83 at baseline, n = 54 at 10 weeks
	
	Subset: cognitive function outcomes were reported for a subset of patients enrolled after October 2007, including:
	
	• EORTC QLQ-C30 cognitive subscale
	• FACT-Cog subscales of:
	  * perceived cognitive impairment
	  * perceived cognitive abilities
	  * impact of cognitive impairments on QoL

Oh 2010 *(Continued)*

	For this group, outcomes were measured at baseline and 10 weeks • exercise group: n = 37 at baseline, n = 23 at 10 weeks • control group: n = 44 at baseline, n = 31 at 10 weeks Subgroup analysis: none reported Adverse events: none reported
Notes	Country: Australia Funding: University of Sydney Cancer Research Fund

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomization, by computer..."
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Low risk	ITT analysis completed. There were 25 drop-outs in the exercise group and 29 drop-outs in the control group, and missing values were "dealt with by multiple imputation..."
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Ohira 2006

Study characteristics

Methods	Study design: RCT Number randomized: 86; 43 to the exercise group and 43 to the delayed exercise control group Study start and stop dates: October 2001 to June 2002 Length of intervention: 6 months Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer

Ohira 2006 *(Continued)*

Cancer stage, n (%):

- exercise group: DCIS, 7 (18%); Stage I, 16 (43%); Stage II, 13 (34%); Stage III, 2 (5%)
- control Group: DCIS, 5 (12%); Stage I, 16 (39%); Stage II, 18 (44%); Stage III, 2 (5%)

Time since cancer diagnosis, mean (range) years:

- exercise group: mean 1.73 (0.58 to 3.59) years
- control group: mean 2.02 (0.44 to 11.42) years

Time beyond active treatment, mean (range) years:

- exercise group: mean 1.21 (0.28 to 2.84) years
- control group: mean 1.09 (0.25 to 3.12) years

Inclusion criteria:

- completed all treatment except hormonal therapy for breast cancer
- body weight stable within 10% over the past year
- nonsmokers for at least the past 2 years

Eligibility criterion related to interest or ability to exercise, or both:

- sedentary to moderately physically active (no more than 3 sessions per week of no more than moderate-intensity activity, no weight training history)

Exclusion criteria:

- medical conditions prohibiting participation in a weight training program
- morbidly obese (BMI > 40 kg/m^2)
- hypertensive (systolic blood pressure > 160 mmHg, diastolic blood pressure > 99 mm Hg, or both)
- currently on a weight loss plan or planning to start a weight loss plan during the period of the study
- planning to move away from the area or be away from area for > 3 weeks during study
- not pregnant or lactating, or planning to become pregnant during the study period

Gender: female

Current age, mean (SD) years:

- exercise group: 53.3 (8.7) years
- control group: 52.8 (7.6) years

Age at cancer diagnosis: not reported

Ethnicity/race, n (%) Caucasian:

- exercise group: 39 (98%) Caucasian
- control group: 41 (100%) Caucasian

Education level, n (%):

- exercise group: some college or vocational training, 8 (20%); college degree, 22 (55%); graduate or professional degree, 10 (25%)
- control group: some college or vocational training, 7 (17%); college degree, 22 (54%); graduate or professional degree, 12 (29%)

SES: not reported

Employment status: not reported

Comorbidities: not reported

Past exercise history: limited by eligibility criteria

Ohira 2006 *(Continued)*

	On hormone therapy, n (%): • exercise group: tamoxifen, 30 (77%); anastozole 3 (8%); other 0 (0%) • control group: tamoxifen 27 (66%); anastozole 5 (12%); other, 1 (2%) Other: postmenopausal, n (%): • exercise group: postmenopausal, 34 (85%) • control group: postmenopausal, 32 (78%)
Interventions	43 participants assigned to the exercise intervention, including: • 9 common weight-training exercises using variable resistance machines and free weights (for muscles of the chest, back, shoulders, and arms, buttocks, hips, and thighs) • stretching exercises to perform before and after each weight-training session Type exercise (aerobic/anaerobic): aerobic and anaerobic Intensity of experimental exercise intervention: not reported Frequency: twice per week Duration of sessions: not reported Duration of program: 6 months Total number of exercise sessions: maximum of 52 Format: started as supervised group (4 in a group); after 13 weeks, participants encouraged to work out with buddy(ies). Log sheets checked weekly by fitness trainer and if no data were recorded in 1 week, the fitness trainer called the participant Facility: gym facility Professionally led: certified fitness professional 43 participants assigned to control group, including: • waiting list Adherence: baseline to 6 months, 1 participant attended < 80% of the sessions. From months 7 to 12, 14 exercise group participants attended < 70% of sessions Contamination of control group: not reported
Outcomes	Primary outcome: QoL and physiologic outcomes, including: • CARES-SF, which includes 59 items and 5 subscales for physical, psychosocial, medical interaction, marital, sexual, and other miscellaneous subscales. Items are assessed using a 5-point Likert scale (0 = 'not at all', 1 = 'a little', 2 = 'a fair amount', 3 = 'much', 4 = 'very much') that queries the applicability of the problem/statement to the patient within the last month. Items of the CARES-SF are combined into a global summary score. Both the global summary score and individual subscale scores range from 0 to 100 and lower scores indicate fewer problems • weight • height • dual energy X-ray absorptiometry (for body composition), including skin pinch meter/scale • standard blood panel • upper and lower body strength assessed by 1-RM, using the same machines trained on for 9 exercises • depressive symptoms, measured with the CES-D, a 20-item questionnaire scored on a standard 4-point scale (0 to 3) for each item, with a potential range of 0 to 60 Outcomes were measured at baseline and 6 months:

Ohira 2006 *(Continued)*

- exercise group: n = 43 at baseline, n = 39 at 6 months
- control group: n = 43 at baseline, n = 40 at 6 months

Subgroup analyses: post hoc subgroup analyses, including postmenopausal status, baseline levels of sport and leisure physical activity, baseline level of energy intake (kilocalories), and 6-month changes in physical activity and energy intake

Adverse events:

- cancer recurrence: n = 4 total, 2 recurrences each in the exercise and control groups
- some limited musculoskeletal issues which were self resolving

Notes	Country: US Setting: Recreation center Funding: Susan G. Komen Foundation and grants to the UMN GCRC from the NIH

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Random number table"
Allocation concealment (selection bias)	Low risk	"The randomization procedure used prevented investigators from influencing treatment allocation"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, masking or blinding of study participants was not possible; however, it is unclear whether the lack of masking could influence the outcomes
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"Measurement staff remained blinded until the end of the study, with the exception of the strength testing staff..."
Incomplete outcome data (attrition bias) All outcomes	High risk	4 participants lost to follow-up in the exercise group, 2 due to recurrences and 2 due to withdrawals; 3 participants were lost to follow-up in the control group, 2 due to recurrences and 1 due to withdrawal
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Payne 2008

Study characteristics

Methods	Study design: RCT Number randomized: 20; 10 to the exercise group and 10 to the control group Study start and stop dates: 9-month period but dates not reported Length of intervention: 14 weeks

Payne 2008 *(Continued)*

	Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer Time since cancer diagnosis: not reported Time beyond active treatment: not reported Inclusion criteria: • postmenopausal women • diagnosis of breast cancer • receiving hormonal therapy with tamoxifen, anastrozole, or letrozole during the period of recruitment and study enrolment • ≥ 55 years old • complaints of fatigue • speak English • KPS score of ≥ 80 Eligibility criterion related to interest or ability to exercise, or both: • no neuromuscular deficits Exclusion criteria: • documented history of neurologic deficits or mental illness (e.g. psychotic deficits) within the past year Gender: female Current age, mean (SD) years: 64.7 (6.3) years Age at cancer diagnosis: not reported Ethnicity/race, n (%): 18 (90%) Caucasian and 2 (10%) African-American Education level, n (%): 11th or 12th grade education, 5 (25%); some college level education, 7 (35%); college degree or higher level of education, 8 (40%) SES, household income, n (%): ≤ USD20,000, 4 (20%); USD20,001 to USD40,000, 5 (25%); USD40,001 to USD60,000, 4 (20%); > USD60,000, 6 (30%); refused to answer 1 (5%) Employment status, n (%): employed, 3 (15%); homemaker, 1 (5%); retired, 11 (55%); other employment status, 2 (10%); did not provide employment information, 3 (15%) Comorbidities: not reported Past exercise history: not reported On hormone therapy: all used tamoxifen, anastrozole, or letrozole
Interventions	10 participants assigned to the exercise intervention, including: • home based walking activity Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: moderate Frequency: 4 times per week Duration of individual sessions: 20 minutes Duration of exercise program: 14 weeks

Payne 2008 *(Continued)*

	Total number of exercise sessions: 56 sessions Format: individual Facility: home Not professionally led, but explained by study coordinator 10 participants assigned to the control group, including: • usual care Adherence: 9 out of the 10 women completed the study, adherence data on number of sessions completed are not specified Contamination of control group: not reported
Outcomes	No primary outcome was identified. QoL outcomes included: • fatigue, measured using the Piper Revised Fatigue Scale • sleep disturbance, measured using the PSQI • depressive symptoms, measured using the CES-D Physiologic outcomes included: • blood chemistry, including cortisol, serotonin, interleukin-6, bilirubin markers Outcomes were measured at baseline, 12 weeks, and 14 weeks: • exercise group: n = 10 at baseline, n = 10 at 12 weeks, n = 9 at 14 weeks • control group: n = 10 at baseline, n = 9 at 12 weeks, n = 9 at 14 weeks Subgroup analysis: none reported Adverse events: none reported
Notes	Country: US Funding: NIH/National Institute of Nursing Research

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Generation of allocation sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	There was no description on how missing data were handled. Participants in each group withdrew from the study

Payne 2008 *(Continued)*

Selective reporting (reporting bias)	Low risk	There is no evidence of selective reporting of outcomes
Other bias	High risk	Small sample size and low recruitment rate can put study at risk of bias

Penttinen 2011

Study characteristics	
Methods	Study design: RCT Number randomized: 573; 302 to the exercise group and 271 to the control group Study start and stop dates: enrolment between September 2005 and September 2007 Length of intervention: 12 months Length of follow-up: 6 and 12 months after baseline
Participants	Type cancer: breast cancer Time since cancer diagnosis: not reported Time beyond active treatment, median (range) weeks: • time since surgery: 33 (27 to 40) weeks • time since last chemotherapy: 12 (5 to 17) weeks • time since last radiation therapy: 4 (-2 to 10) weeks Inclusion criteria: • histologically confirmed newly diagnosed invasive breast cancer (T1-4 NO-3 MO) • pre- and postmenopausal • treated with adjuvant chemotherapy or radiation therapy within 4 month • started endocrine therapy (antiestrogens, aromatase inhibitors, luteinizing hormone-releasing hormone agonists, or a combination) no more than 4 months earlier • 35 to 68 years old • signed informed consent prior to beginning protocol specific procedures Eligibility criterion related to interest or ability to exercise, or both: • incapable of training, (e.g. severe cardiac disease, osteoporosis, severe knee arthrosis, ligament or cartilage injuries at lower extremities) • other serious illness or medical condition, which could be contraindication for exercise Exclusion criteria: • male gender • prior malignancy except basal cell carcinoma or in situ carcinoma • hematogenous metastases (M1) • systemic adjuvant therapy • postmenopausal women with antiestrogens as the only adjuvant treatment (with or without radiation therapy) • pregnancy or recent lactation (< 1 year) • severe cardiac disease (NYHA class III or greater), myocardial infarction within 12 months, uncontrolled hypertension • verified osteoporosis (proximal femur or lumbar spine T-score < -2.5 or fracture without trauma)

Penttinen 2011 *(Continued)*

	• concomitant medications affecting calcium and bone metabolism such as bisphosphonates, calcitonin, parathyroid hormone, selective estrogen receptor modulators, oral corticosteroids (over 6 months), anticonvulsants (phenytoin, carbamatsebin), and prolonged heparin therapy • other diseases affecting calcium and bone metabolism such as hyperthyroidism, newly diagnosed hypothyroidism, primary hyperparathyroidism, renal failure, chronic hepatic diseases, organ transplant • residency more than 1 hour from the exercise center • competitive athlete • only treated with radiation therapy Gender: female Current age, mean (range) years: 52.4 (35 to 68) years Age at cancer diagnosis: not reported Ethnicity/race: not reported Education level, mean (SD) years: 13.9 (3.4) years SES: not reported Employment status, n (%): self-employed, 20 (3.7%); upper-level employees, 101 (18.8%); lower-level employees, 227 (42.3%); manual workers, 45 (8.4%); students, 5 (0.9%); pensioner, housewives, 98 (18.2%); unemployed, 16 (3.0%); missing, 25 (4.7%) Comorbidities, n (%): hypertension, 108 (20.1%); cardiovascular diseases, 10 (1.9%); diabetes, 12 (2.2%); psychiatric disease, 48 (8.9%) Past exercise history, n (%): inactive, 92 (17.1%); exercising with moderate intensity, 282 (52.5%); vigorous exercising, 114 (21.2%); missing, 49 (9.1%) On hormone therapy, n (%): 445 (82.9%) BMI, n (%): < 25 (normal weight), 231 (43.0%); 25 to 30 (overweight), 204 (38.0%); > 30 (obese), 102 (19.0%)
Interventions	302 participants assigned to a 2-component supervised exercise training intervention, with each component performed in alternate weeks. The components included: • 1 supervised session of vigorous step aerobics, comprised of 150 to 180 jumps and leaps in diverging directions, progressing from 10-cm high benches to 15-cm benches after 4 months and 20-cm benches after 8 months. Music was set as 118 beats per minute • circuit training started with 100 steps and hops per session, progressing 150 to 180 steps and hops per session, with more demanding jumps in the later phase. The session started with a 20-second training period followed by 60-second rest and progressed to a 40:60 second training/rest ratio and then a 30:60 second ratio with more demanding jumps such as heel drops, star jumps, and skate jumps In addition, there were 2 to 3 similar home training sessions Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: vigorous Frequency: once per week and 2 to 3 similar home sessions Duration of individual sessions: 60 minutes Duration of exercise program: 12 months Total number of exercise sessions: 52 supervised sessions and 104 to 156 home training sessions Format: group (at facility) and individual (at home) Facility: facility and home

Penttinen 2011 *(Continued)*

	Professionally led by experienced physical therapists who all had also received a similar training for the study exercises 271 participants assigned to the control group, including: • usual care Adherence: 24 premenopausal trainees attended a median of 30/52 (58%) supervised training sessions • 6/124 (5%) did not attend any training • 23/124 (18%) attended < once a month • 78/124 (63%) attended at least every second week (i.e. > 25 times) Based on 109 returned training diaries, premenopausal participants completed home training on average 2.8 times weekly for a total time of 2.9 hours. The median total number of training sessions (supervised and home training sessions together) was 3.3 times per week (interquartile range 2.4 to 4.6) Postmenopausal trainees attended a median of 33/52 (63%) training sessions: • 2/138 (< 2%) did not attend any session • 27/138 (20%) attended sessions < once a month • 96/138 (70%) attended at least every second week Based on 122 returned training diaries, postmenopausal participants completed home training 3.2 times (107%) weekly for a total time of 3.5 hours. The median total number of training sessions was 4.3 times per week (interquartile range 2.3 to 5.4) Contamination of control group: not reported
Outcomes	No primary outcome was identified. QoL outcomes included: • QoL, measured using the EORTC QLQ-C30 • fatigue, measured using the FACIT-F scale • depression, measured using the BDI Physical outcomes included: • physical performance • body composition • bone density QoL outcomes were measured at baseline, but there is no report on whether these outcomes were measured at follow-up. Physiologic outcomes were measured at baseline, and 6 and 12 months. These outcomes were measured as follows: • exercise group: n = 302 at baseline, n = 262 for physiologic outcomes at 6 and 12 months • control group: n = 271 at baseline, n = 236 for physiologic outcomes at 6 and 12 months Subgroup analysis: none reported Adverse events: none reported
Notes	Country: Finland Funding: Finnish Cancer Institute; Finnish Cancer Foundation; Academy of Finland; Social Insurance Institution of Finland; Finnish Ministry of Education; Finska Läkaresällskapet; Special government grant for health science research; Helander Foundation; Gyllenberg Foundation; Paulo Foundation; Kurt and

Penttinen 2011 *(Continued)*

	Doris Palander Foundation; Finnish Cultural Foundation and Medical Fund of the Pirkanmaa Hospital District; Finnish Astra-Zeneca sponsored step benches for the study; Finnish Breast Cancer group

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	There was no ITT analysis and it is unclear how missing data were handled. Analyses were completed on 262/302 women in the intervention group and in 236/271 women in the control group
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Pinto 2003

Study characteristics

Methods	Study design: RCT Number randomized: 24, 12 to the exercise group and 12 to the control group Study start and stop dates: not reported Length of intervention: 12 weeks Length of follow-up: within 1 week of completing the intervention
Participants	Type cancer: breast cancer, Stage 0 to II Cancer stage, n (%): • Stage 0, 2 (9%); Stage I, 18 (78%); Stage II, 3 (13%) Time since cancer diagnosis, mean (SD) days: 452.6 (234.9) days Time beyond active treatment, mean (SD) days: 323.5 (212.3) days Inclusion criteria: • diagnosed with Stage 0 to II breast cancer over the past 3 years

Pinto 2003 *(Continued)*

- post-surgery, had completed chemotherapy, and/or radiation
- informed consent
- physician sign-off
- agree to be assigned randomly to either of the 2 study groups

Eligibility criterion related to interest or ability to exercise, or both:

- sedentary women (exercised less than 3 times per week for 20 minutes per session)

Exclusion criteria:

- current medical or psychiatric illness making compliance with protocol difficult or dangerous (e.g. coronary artery disease, hypertension, diabetes)
- orthopedic problems or neuropathies that would limit exercise training
- medications likely to alter training responses (e.g. beta-blockers) or affect distress outcomes (e.g. antidepressants)

Gender: female

Current age: mean (SD) years, 52.5 (6.8) years

Age at cancer diagnosis: not reported

Ethnicity/race: 100% Caucasian

Education level: high school, 30%; college degree, 30%

SES: 42% reported an annual household income of > USD50,000

Employment status: 16 (67%) working full-time

Other:

- mean (SD) weight was 150.0 (28.2) lbs
- mean BMI (SD) was 26.8 (4.1)
- 17 (71%) reported being postmenopausal at baseline

Past exercise history: not reported

On hormone therapy, n (%): 14 (61%)

Interventions	12 participants assigned to the exercise group, including: • 10 minutes of warm-up (cardiovascular and flexibility), 10 minutes of cool-down (cardiovascular and flexibility), and 30 minutes of cardiovascular activity in the individual's target HR zone • strength (weight) training during final month Type exercise (aerobic/anaerobic): aerobic first 2 months; both aerobic and anaerobic during final month Intensity of experimental exercise intervention: 60% to 70% of peak HR by the end of the 12-week intervention Frequency: 3 times per week Duration of session: 50 minutes Duration of program: 12 weeks Total number of exercise sessions: 36 Format: group and home-based exercise program Facility: both facility and home

Pinto 2003 *(Continued)*

	Professionally led: fitness instructor for class, supervised and upgraded program by exercise physiologist once per week Control group: 12 participants assigned to the control group, including • waiting list, with no stress tests completed Adherence: 3 participants withdrew within the first 3 weeks Contamination of control group: not reported
Outcomes	Outcome: QoL and physiologic outcomes, including: • POMS, a 65-item measure of 6 mood states including anger, tension, depression, vigor, fatigue, and confusion over the past week, and a summary score (total mood disturbance). Response options ranged between 0 = not at all to 4 = extremely • PANAS, which includes 20 items with each requiring a response to 'how you are feeling at the moment?' scored on a 1 to 5 Likert scale (1 = very slightly, 5 = extremely) • BES, a 35-item scale assessing a subject's evaluation of sexual attractiveness, physical condition, and weight concerns. Response options ranged between 1 = have strong negative feelings to 5 = have strong positive feelings • BMI • weight • physical performance on stress test Outcomes were measured at baseline and within 1 week of end of the 12 week training: • exercise group: n = 12 at baseline, n = 9 after intervention • control group: n = 12 at baseline, n = 6 at 12 weeks for QoL outcomes; physiologic outcomes were not measured at follow-up Adverse events: none reported
Notes	Country: US Funding: National Institute of Mental Health, NIH

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	3 participants withdrew in the exercise group, and 6 participants withdrew in the control group

Pinto 2003 *(Continued)*

Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Pinto 2005

Study characteristics	
Methods	Study design: RCT Number randomized: 86; 43 to the exercise group and 43 to the control group Study start and stop dates: 1998 to 2003 Length of intervention: 12 weeks Length of follow-up: 12 weeks, 6 months, and 9 months
Participants	Type cancer: breast cancer Cancer stage, Stage 0 to II, n (5): • exercise group: Stage 0, 8 (18.6%); Stage I, 17 (39.5%); Stage II, 18 (41.9%) • control group: Stage 0, 6 (14.0%); Stage I, 15 (34.9%); Stage II, 22 (51.2%) Time since cancer diagnosis, mean (SD) years: • exercise group: 1.74 (1.49) years • control group: 1.93 (1.37) years) Time beyond active treatment: not reported Inclusion criteria: • ≥ 18 years old • diagnosed with Stage 0 to II breast cancer over the last 5 years • completed surgery, chemotherapy, radiation therapy, or a combination • ambulatory (able to walk 1 mile without assistive devices) • willing to be randomized Eligibility criterion related to interest or ability to exercise, or both: • currently sedentary (exercised < once a week for 20 minutes at vigorous intensity or < twice a week for 30 minutes at moderate intensity for the past 6 months) Exclusion criteria: • prior history of cancer, except for nonmelanoma skin cancer • medical or current psychiatric illness making compliance with the study protocol difficult or dangerous • cardiovascular disease, diabetes, or orthopedic problems that would limit exercise training Gender: female Current age, mean (SD) years: • exercise group: 53.42 (9.08) years

Pinto 2005 *(Continued)*

- control group: 52.86 (10.38) years

Age at cancer diagnosis: not reported

Ethnicity/race, n (%):

- exercise group: white, 42 (97.7%); African-American, 0 (0%); Native American, 0 (0%); Asian/Pacific Islander, 0 (0%)
- control group: white, 40 (93.0%); African-American, 1 (2.3%); Native American, 1 (2.3%); Asian/Pacific Islander, 1 (2.3%)

Education level, n (%):

- exercise group: high school diploma, 7 (16.3%); vocational/trade school, 1 (2.3%); some college, 7 (16.3%); Associates degree, 6 (14.0%); Bachelors degree, 11 (25.6%); graduate school, 11 (25.6%)
- control group: high school diploma, 8 (18.6%); vocational/trade school, 0 (0%); some college, 17 (39.5%); Associates degree, 7 (16.3%); Bachelors degree, 5 (11.6%); graduate school, 6 (14.0%)

SES, household income, n (%):

- exercise group: ≤ USD29,999, 0 (0%); USD30,000 to USD49,999, 11 (25.6%); ≥ USD50,000, 27 (62.8%)
- control group: ≤ USD29,999, 12 (28.0%); USD30,000 to USD49,999, 5 (11.7%); ≥ USD50,000, 24 (55.8%)

Employment status, n (%):

- exercise group: employed full time, 23 (53.5%); employed part time, 12 (27.9%; retired/homemaker/medical leave, 8 (18.6%)
- control group: employed full time, 24 (55.8%), employed part time, 4 (9.3%), retired/homemaker/medical leave, 15 (34.9%)

Other, BMI, mean (SD) kg/m^2:

- exercise group: 27.5 (5.04) kg/m^2
- control group: 28.56 (5.50) kg/m^2

Comorbidities: none reported

Past exercise history: not reported

On hormone therapy, n (%):

- exercise group: 21 (48.8%)
- control group: 32 (74.4%)

Interventions	43 participants assigned to the exercise group, including: • 12-week, home-based physical activity program, including brisk walking, biking, swimming, or use of home exercise equipment • in-person instruction on exercising • weekly physical activity counseling via telephone, including individually based reinforcement, problem-solving, and monitoring participation • weekly tip sheets on physical activity • cancer survivorship tip sheet Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: 55% to 65% of maximum HR Frequency: 10 minutes each on 2 days per week, gradually increased to 30 minutes per day on 5 days per week Duration of sessions: 10 to 30 minutes

Pinto 2005 *(Continued)*

	Duration of program: 12 weeks Total number of exercise sessions: 24 to 60 Format: individual Facility: home-based Professionally led: unclear; individual instruction provided but not reported by whom and regular telephone calls made, but not reported who made the calls Adherence: participants wore a pedometer, but adherence not reported Control group: 43 participants assigned to control group, including: • no change in current level of physical activity for 12 weeks. • phone calls from research staff • cancer survivor tip worksheet Contamination of control group: not reported
Outcomes	Primary outcome: QoL outcomes, including: • POMS, a 65-item questionnaire, measures a variety of mood states including anger, tension/anxiety, depression, vigor, fatigue, confusion, and total mood disturbance: vigor and total mood score used as primary outcomes in this study. Response options are presented on a scale of 0 to 4 (0 = not at all, 4 = extremely) • level of fatigue, assessed by asking participants to place a vertical mark on a 10-cm linear analog scale. The scale was scored by measuring the distance in millimeters from the left anchor (i.e., "0") to the vertical mark. Higher scores represent greater fatigue Secondary outcomes: QoL and physiologic outcomes, including: • BES, a 35-item scale assessing a subject's evaluation of sexual attractiveness, weight concerns, and physical condition with 3 subscales, where higher scores indicate higher esteem • Physical Activity Recall, assessed by standardized self-report interview • Rockport 1-mile walk test, with time to walk 1 mile measured All outcomes measured at baseline and end of intervention: • exercise group: n = 43 at baseline, n = 39 after intervention • control group: n = 43 at baseline, n = 43 after intervention Subgroups: baseline demographics, hormone therapy, type partner controlled for in the analyses. Employment and education dichotomized outcomes Adverse events: not reported
Notes	Country: US Funding: National Cancer Institute grant

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described

Pinto 2005 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	4 participants withdrew from the exercise group, and 2 participants withdrew from the control group before the 6-month assessment. Another 2 participants in the control group withdrew before the 9 month assessment
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Rogers 2009

Study characteristics	
Methods	Study design: RCT Number randomized: 41; 21 to the exercise group and 20 to the control group Study start and stop dates: recruitment from April 2006 to May 2007 Length of intervention: 12 weeks Length of follow-up: 3 months after end of the intervention
Participants	Type cancer: breast Cancer stage, Stage I to III, n (%): • exercise group: Stage I, 6 (29%); Stage II, 11 (52%); Stage III, 4 (19%) • control group: Stage I, 6 (30%); Stage II, 10 (50%); Stage III, 4 (20%) Time since cancer diagnosis: not reported Time beyond active treatment: • exercise group, mean (SD) months: months since surgery [n = 21(100%)], 35 (38) months; months since chemotherapy [n = 17 (81%)], 36 (39) months; months since radiation [n = 18 (86%)], 35 (41) months • control group, mean (SD) months: months since surgery [n = 20 (100%)], 34 (30) months; months since chemotherapy [n = 17 (85%)], 30 (31) months; months since radiation [n = 6 (80%)], 30 (31) months Inclusion criteria: • female • 18 to 70 years old • history of stage I, II, or IIIA breast cancer • English speaking • currently taking an aromatase inhibitor or estrogen receptor modulator • medical clearance provided by physician • at least 8 weeks postsurgery

Rogers 2009 *(Continued)*

Eligibility criterion related to interest or ability to exercise, or both:

- engaging in 60 or more minutes of vigorous physical activity or 150 or more minutes of moderate plus vigorous activity per week during the past month based on self-report

Exclusion criteria:

- dementia or organic brain syndrome
- medical, psychological, or social characteristic that would interfere with the ability to fully participate in program activities and assessments
- contraindication to participate in a regular physical activity program (e.g. unstable angina, debilitating arthritis pain)
- inability to ambulate
- plans to relocate outside the study area during the study period
- breast cancer recurrence or metastasis.

Gender: female

Current age, mean (SD) years:

- exercise group: 52 (15) years
- control group: 54 (8) years

Age at cancer diagnosis: not reported

Ethnicity/race, n (%):

- exercise group: white, 19 (90%); other, 2 (10%)
- control group: white, 19 (95%); other, 1 (5%)

Education level, mean (SD) years:

- exercise group: 15 (2) years
- control group: 15 (2) years

SES household income, n (%):

- exercise group: < USD10,000, 1 (5%); USD10,000 to USD35,000, 7 (33%); USD35,000 to USD50,000, 3 (14%); > USD50,000, 10 (48%)
- control group: < USD10,000, 1 (5%); USD10,000 to USD35,000, 0 (0%); USD35,000 to USD50,000, 5 (25%); > USD50,000, 14 (70%)

Employment status: not reported

Comorbidities:

- exercise group: comorbidity score on a scale from 0 to 5: 2 (1.4)
- control group: comorbidity score on a scale from 0 to 5: 2 (1.6)

Past exercise history: not reported

On hormone therapy:

- exercise group: months on hormonal therapy, mean (SD) months, 15 (15) months; estrogen receptor modulator, n (%), 7 (33%); aromatase inhibitor, n (%), 14 (67%)
- control group: months on hormonal therapy, mean (SD) months, 22 (18) months; estrogen receptor modulator, n (%), 4 (20%); aromatase inhibitor, n (%), 16 (80%)

Interventions	21 participants assigned to the exercise group, including: • 6 discussion group sessions with a clinical psychologist at baseline, and weeks 1, 2, 4, 6, and 8 • 6 supervised exercise programs (walking), 3 per week during weeks 1 and 2, 2 per week during weeks 3 and 4, and 1 per week during weeks 5 and 6

Rogers 2009 *(Continued)*

- 40 home-based exercise (walking), 2 per week during weeks 3 and 4, 3 per week during weeks 5 and 6, 5 per week during weeks 7 through 12
- 3 individual update counseling sessions with an exercise specialist during week 8, 10, and 12

Type exercise (aerobic/anaerobic): aerobic

Intensity of experimental exercise intervention: transition from baseline to week 12 to 150 minutes of moderate-intensity activity

Frequency: gradually increased from 3 times a week to 5 times a week

Duration of sessions: not reported

Duration of program: 12 weeks

Total number of exercise sessions: 52

Format: individual exercise; group peer support

Facility: facility and home

Professionally led: exercise specialists certified (or certification-eligible) by the ACSM

Adherence: participants completed 100% (252/252) of the individual exercise sessions, 95% (60/63) of the individual update sessions, and 98% (123/126) of the group session for overall 99% (435/441) adherence. 6% (4/63) of update sessions were completed by telephone

20 participants were assigned to the control group, including:

- usual care, including written materials about physical activity available through the American Cancer Society

Contamination of control group: not reported

Outcomes	Outcomes: QoL outcomes and physiologic outcomes, including • FACT-B, including subscales of physical functioning, SWB, EWB, FWB, and additional concerns • FACT-G, the sum of the physical functioning, SWB, EWB, and FWB • FACT-F, a 13-item instrument • FACT-Cog, a 42-item instrument • FACT-ES, a 19-item instrument • sleep dysfunction, assessed using the PSQI • joint pain, stiffness, and physical function, using a 5-point Likert scale version (1 = none to 5 = extreme) of the 24-item WOMAC • objective activity monitoring, measured using a GT1M accelerometer • self-reported leisure time physical activity, assessed using the Godin Leisure-Time Exercise Questionnaire • stage of motivational readiness for physical activity, classified as precomtemplation, contemplation, preparation, action, and maintenance • fitness, assessed using a submaximal treadmill test and Naughton protocol to estimate oxygen consumption at 85% of predicated maximal HR • muscle strength, assessed using back/leg extensor dynamometers (Takei Back-A model #Tkk5002 - i.e. best of 3 attempts) and handgrip dynamometer (Lafeyette Model No. 78010) • BMI • waist to hip ratio, using a nonstretching tape measure to measure the waist and hip circumferences over undergarments with 3 measurements averaged • percent body fat and bone mineral density, assessed by dual energy X-ray absorptiometry • caloric intake, assessed with a 3-day diet record (i.e. 1 weekend and 2 weekdays) and analyzed with Diet Analysis Plus software, version 7.0.1 (Thomson) • perceived health, assessed using a 5-point Likert scale

Rogers 2009 *(Continued)*

	Outcomes were measured at baseline, 12 weeks, and 3 months after intervention (6 months after randomization) • exercise group: n = 21 at baseline, n = 20 at 12 weeks, n = 19 at 6 months • control group: n = 20 at baseline, n = 19 at 12 weeks, n = 17 at 6 months Adverse events: none reported
Notes	Country: US Funding: Southern Illinois University School of Medicine Excellence in Academic Medicine Award, Brooks Medical Research Fund, Memorial Medical Center Foundation and Regional Cancer Center

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Computer-generated"
Allocation concealment (selection bias)	Unclear risk	"Sealed envelopes"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	The investigators stated that they conducted an ITT analyses, but 2 participants withdrew from the exercise group and 3 from the control group. The authors also reported that the rate of missing data for the FACT-ES and the FACT-Cog exceeded the pre-specified amount for imputation of values and they analyzed
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Segar 1998

Study characteristics

Methods	Study design: quasi-randomized partial cross-over controlled trial. We include only the first treatment period Number randomized: 30; 10 to an exercise group, 10 to an exercise and behavioral modification group, and 10 to a control group. The exercise and exercise and behavioral modification groups were combined for all analyses Study start and stop dates: not reported Length of intervention: 10 weeks

Segar 1998 *(Continued)*

	Length of follow-up: 12 weeks
Participants	Type cancer: breast cancer Time since cancer diagnosis: not reported Time beyond active treatment, mean (SD, range) months: 41.8 (24.9, 1 to 99) months Time since surgery, mean (SD, range) months: • exercise groups: 43.7 (26.2, 1 to 99) months • control group: 38.1 (23.2, 5 to 73) months Inclusion criteria: • any type of breast cancer surgery • 30 to 65 years old Eligibility criterion related to interest or ability to exercise, or both: • not currently participating in exercise • no contraindications to exercise • written release from the physician Exclusion criteria: • cardiovascular or pulmonary disease • known physical disabilities Gender: female Current age, mean (SD, range) years: • exercise groups: 47.5 (7.1, 35 to 62) years • control group: 51.8 (8.1, 40 to 64) years Age at cancer diagnosis: not reported Ethnicity/race: • exercise groups: white, 70%; black, 12%; Asian, 12%; and Native American, 6% • control group: white, 88%; black, 12% Education level: • exercise groups: graduate training, 38%; college graduate, 31%; some college, 25%; and high school graduate, 6% • control group: graduate training, 12%; college graduate, 38%; some college, 25%; and high school graduate, 25% SES: not reported Employment status: • exercise groups: employed outside the home, 56%; homemaker, 25%; unemployed, 0; retired, 6%; other, 13% • control group: employed outside the home, 25%; homemaker, 12%; unemployed, 25%; retired, 13%; other, 25% Comorbidities: not reported Past exercise history: not reported

Segar 1998 *(Continued)*

	On hormone therapy: not reported
Interventions	10 participants assigned to the exercise group, including • request to exercise a minimum of 30 minutes on 4 days per week, with type of exercise (stationary bike, stair climbers, and hydraulic resistance exercise equipment) as chosen by participant Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: ≥ 60% of age-predicted maximum HR (mild to moderate) Frequency: 4 days per week Duration of session: 30 to 40 minutes per session Duration of program: 10 weeks Total number of exercise sessions: 40 Format: unclear Facility: exercise facility was at the university or that of the participants' preference Professionally led: not professionally led 10 participants assigned to the exercise and behavioral modification group, including: • exercise as described for the exercise group • behavioral modification by self awarded rewards (activity, food, treats, or movies) to serve as reinforcements Adherence: overall compliance, assessed from self-reported exercise logs averaged 1363 (SD = 577) minutes over the 10 weeks, where 100% compliance is equivalent to 1,200 minutes. Average compliance for participants reaching at least 89% compliance was 1532 (SD = 103) minutes (mean compliance of 130%) with a range form 89 to 250% 10 participants initially assigned to the control group, including: • instructions to maintain sedentary lifestyle Contamination of control group: unclear
Outcomes	Outcomes: QoL outcomes, including: • change in depressive symptoms, measured using the 21-item BDI questionnaire, with scale score ranging between 0 and 63. A higher score indicates greater depressive symptoms • change in anxiety symptoms, measured using the STAI, 20 items, 1 = not at all, 4 = very much so • change in self esteem, measured using the RSE Inventory, which is a unidimensional 64-item questionnaire with 10 scales that reflect self-evaluation of self-esteem Outcomes were measured at baseline and at 10 weeks: • exercise groups: n = 16 at baseline, n = 16 at 10 weeks • control group: n = 8 at baseline, n = 8 at 10 weeks Adverse events: none reported
Notes	Country: US Funding: Michigan Initiative for Women's Health Grant
Risk of bias	

Segar 1998 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"Subjects were rotated sequentially into two treatment conditions and one control group"
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	4 participants were excluded from the exercise group and 2 participants were excluded from the control group. Exclusion from the analyses occurred because of attrition or noncompliance with the study protocol
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Speck 2010

Study characteristics	
Methods	Study design: RCT Number randomized: 295; 148 (71 with lymphedema and 77 without lymphedema) to the exercise group and 147 (70 with lymphedema and 77 without lymphedema) to the control group Study start and stop dates: recruitment took place between October 2005 and February 2007 Length of intervention: 12 months Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer Cancer Stage, for women with lymphedema, n (%): • exercise group: DCIS, 0 ()%); Stage I, 13 (24%); Stage II, 20 (37%); Stage III, 4 (7%); unknown, 17 (32%) • control group: DCIS 6 (10%); Stage I, 8 (14%); Stage II, 14 (24%); Stage III, 11 (19%); unknown, 19 (33%) Cancer Stage, for women without lymphedema, n (%): • exercise group: DCIS 1 (2%); Stage I, 22 (37%); Stage II, 19 (32%); Stage III, 6 (10%); unknown, 11 (19%) • control group: DCIS 1 (2%); Stage I, 22 (35%); Stage II, 17 (27%); Stage III, 2 (3%); unknown, 21 (33%) Time since cancer diagnosis, for women with lymphedema, mean (SD) months: • exercise group: 78 (42) months • control group: 89 (45) months

Speck 2010 *(Continued)*

Time since cancer diagnosis, for women without lymphedema, mean (SD) months:

- exercise group: 37 (14) months
- control group: 42 (15) months

Time beyond active treatment: not reported

Inclusion criteria:

- female
- history of unilateral nonmetastatic breast cancer
- BMI < 50 kg/m^2
- currently cancer free

Additional inclusion criteria, for women with lymphedema:

- 1 to 15 years post-diagnosis
- at least 1 lymph node removed
- presence of lymphedema

Additional inclusion criteria, for women without lymphedema:

- 1 to 5 years post-diagnosis
- at least 2 lymph nodes removed

Eligibility criteria related to interest or ability to exercise, or both:

- medical condition limiting participation in an exercise program
- weight lifting in the year prior to study entry
- weight stable and not actively trying to lose weight (by self-report)

Exclusion criteria:

- affected limb changes of > 5% between repeated baseline measurements
- medical conditions or medications that would negatively affect the ability to assess BMI
- history of bilateral lymph node dissection
- plans for surgery or pregnancy
- plans to move from the area or be away for ≥ 1 month during the study period

Additional exclusion criteria, for women with lymphedema:

- intensive therapy within the past 3 months
- recorded 10% change in volume or circumference of the affected arm in the last 3 months for ≥ 7 days
- more than 1 lymphedema-related infection requiring antibiotics (cellulitis) within the past 3 months
- change in Activities of Daily Living

Gender: female

Current age, for women with lymphedema, mean (SD) years:

- exercise group: 56 (9) years
- control group: 58 (9) years

Current age, for women without lymphedema, mean (SD) years:

- exercise group: 55(7) years
- control group: 57 (8) years

Age at cancer diagnosis: not reported

Ethnicity/race, for women with lymphedema, n (%):

Speck 2010 *(Continued)*

- exercise group: white, 31 (57%); black, 23 (43%); other, 0 (0%)
- control group: white, 34 (59%); black, 22 (38%); other, 2 (3%)

Ethnicity/race, for women without lymphedema, n (%):

- exercise group: white, 42 (71%); black, 13 (22%); other, 4 (7%)
- control group: white, 52 (83%); black, 11 (17%); other, 0 (0%)

Education level, for women with lymphedema, n (%):

- exercise group: high school or less, 8 (15%); some college, 17 (31%); college or more, 29 (54%)
- control group: high school or less, 12 (21%); some college, 20 (34%); college or more, 26 (45%)

Education level, for women without lymphedema, n (%):

- exercise group: high school or less, 4 (7%); some college, 20 (34%); college or more, 35 (59%)
- control group: high school or less, 9 (14%); some college, 18 (29%); college or more, 36 (57%)

SES: not reported

Employment status, for women with lymphedema, n (%):

- exercise group: professional, 15 (28%); clerical/service, 7 (13%); unemployed, 2 (4%); other/unknown, 19 (35%); retired, 11 (20%)
- control group: professional, 17 (29%); clerical/service, 8 (14%); unemployed, 1 (2%); other/unknown, 10 (17%); retired, 22 (38%)

Employment status, for women without lymphedema, n (%):

- exercise group: professional, 21 (36%); clerical/service, 12 (20%); unemployed, 2 (3%); other/unknown, 15 (25%); retired, 9 (15%)
- control group: professional, 25 (40%); clerical/service, 9 (14%); unemployed, 2 (3%); other/unknown, 14 (22%); retired, 13 (21%)

Comorbidities: n reported

Past exercise history: not reported

On hormone therapy: not reported

Interventions	148 participants (71 with lymphedema and 77 without lymphedema) assigned to the exercise intervention, consisting of progressive strength (weight) training including: • 10 minutes of cardiovascular exercise warm-up • brief range of motion stretching of the major muscle groups to be worked during strength training • 5 to 15 minutes of exercises intended to strengthen spinal stabilization muscles and deep abdominal muscles and increase awareness of body-mind connection • stretching at the end of each session for injury prevention purposes, during which each stretch was held for at least 30 seconds • weights were increased throughout the intervention period Type exercise (aerobic/anaerobic): aerobic and anaerobic Intensity of experimental exercise intervention: not reported Frequency: twice per week Duration of individual sessions: 90 minutes Duration of exercise program: 52 weeks Total number of exercise sessions: 104 sessions

Speck 2010 *(Continued)*

	Format: group Facility: community fitness center, usually a Young Mens Christian Association facility Professionally led and supervised for the first 3 months by fitness professionals employed by the fitness center and certificated by a National Committee for Certifying Agencies accredited organization, such as the ACSM 147 participants (70 with lymphedema and 77 without lymphedema) assigned to the control group, including: • waiting list control • requested not to change current level of exercise Adherence: for women with lymphedema: median attendance to weight lifting sessions was 88%; for women without lymphedema: median attendance to weight lifting sessions was 79% Contamination of control group: not reported
Outcomes	Primary outcome: • limb volume assessed by water displacement Secondary QoL measures included: • body image, measured using the Body Image and Relationship Scale QoL measures assessed using the Health and Attitudes Survey, including: • MOS SF-36 • PSQI • relationship and body image • fatigue • Coopersmith self-esteem • life orientation • visual analog QoL scale • medical outcome and support • temporal SWLS • depression survey Secondary physical measures included: • lymphedema, assessed by circumference measure and onset/flare-up moderators • upper body function testing, assessed using the 9 Hole Peg Test of Finger Dexterity • body composition, measured using BMI, body fat, bone density • muscle strength testing • bioelectrical spectroscopy • pain, assessed using a visual analog scale All outcomes were measured at baseline and 1 year. Some outcomes were also measured at 3 and 6 months. For women with lymphedema, outcomes were measured as follows: • exercise group: n = 71 at baseline, n = 54 at 1 year • control group: n = 70 at baseline, n = 58 at 1 year For women without lymphedema, outcomes were measured as follows: • exercise group: n = 77 at baseline, n = 59 at 1 year • control group: n = 77 at baseline, n = 63 at 1 year

Speck 2010 *(Continued)*

	Subgroup analysis: a large number of subgroup analyses were prespecified Adverse events: none reported
Notes	Country: US Funding: NIH/National Cancer Institute and the Public Health Services Research Grant

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Allocation sequence was a computer generated minimization scheme
Allocation concealment (selection bias)	Low risk	"...de-identified data for ... variables were entered after completion of all baseline measures, the study coordinator then called participants to reveal the outcome of randomization and to schedule groups for the supervised groups"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to mask the participants; however, it is unclear whether the outcome was influenced by a lack of blinding
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Study personnel were not masked or blinded about the allocation, but the outcome assessors measuring physical outcomes were blinded to allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	There is no evidence that missing data were adequately and appropriate addressed. Large numbers of study participants withdrew; among those with lymphedema, 17 women in the exercise group and 12 women in control group withdrew; where, among women without lymphedema, 18 women in the exercise group and 16 women in control group withdrew
Selective reporting (reporting bias)	High risk	The report presents on only some of the outcomes measured in the trial
Other bias	Low risk	No other biases were apparent

Tang 2010

Study characteristics

Methods	Study design: RCT Number randomized: 72; 37 to the exercise group and 35 to the control group Study start and stop dates: not reported Length of intervention: 8 weeks Length of follow-up: 1 month and 2 months
Participants	Type cancer, n (%): • exercise group: breast, 23 (63.9%); gastrointestinal, 6 (16.7%); nasopharyngeal, 4 (11.1%); lung, 0 (0%); other, 3 (8.3%)

Tang 2010 *(Continued)*

- control group: breast, 16 (45.7%); gastrointestinal, 5 (14.3%); nasopharyngeal, 3 (8.6%); lung, 4 (11.4%); other, 7 (20%)

Time since cancer diagnosis, mean (SD) years:

- exercise group: 3.56 (3.92) years
- control group: 4.13 (4.06) years

Time beyond active treatment, mean: unclear whether treatment was concluded. Author reports cancer treatment, n (%):

- exercise group: 7 (19.4%)
- control group: 14 (40.0%)

Inclusion criteria:

- ≥ 18 years old
- diagnosed with cancer
- complaint of sleep disturbance with a PSQI score > 5
- approved for participation by their oncologist
- able to communicate in Mandarin or Taiwanese

Eligibility criterion related to interest or ability to exercise, or both:

- neuromuscular deficits that would contraindicate a walking exercise intervention
- have not regularly undertaken more than 1 session of moderate intensity exercise each week over the past 6 months

Exclusion criteria:

- uncontrolled hypertension, cardiac, or psychiatric illness
- blood pressure > 140/90 mmHg

Gender, n (%):

- exercise group: male, 5 (13.9%); female, 31 (86.1%)
- control group: male, 12 (34.3%); female, 23 (65.7%)

Current age, mean (SD) years:

- exercise group: 47.36 (10.14) years
- control group: 56.37 (12.43) years

Age at cancer diagnosis: not reported

Ethnicity/race: not reported

Education level, mean (SD) years:

- exercise group: 9.97 (3.67) years
- control group: 8.26 (4.66) years

SES: not reported

Employment status, n (%):

- exercise group: working, 13 (36.1%); not working, 23 (63.9%)
- control group: working, 10 (28.6%); not working, 25 (71.4%)

Comorbidities: not reported

Past exercise history: not reported

Tang 2010 *(Continued)*

	On hormone therapy: not reported
Interventions	37 participants assigned to a walking exercise intervention, including: • instructions to walk briskly (at a pace that was faster than normal), starting with a 5-minute warm-up (walking slowly) and finishing with a 5-minute cool-down after completing the 30-minute walk • exercise booklet - written material for home use focusing on safety and proper technique Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: rating of perceived exertion between 11 to 13, with a rating of 6 = resting and 20 = very, very hard Frequency: 3 times per week Duration of individual sessions: 30 minutes plus 5 minutes warm-up and 5 minutes cool-down Duration of exercise program: 8 weeks Total number of exercise sessions: 24 Format: individual Facility: home Not professionally led. 35 participants assigned to the control group, including: • instructions to maintain current lifestyle for 8 weeks • instructions to record in a diary provided by the researchers any exercise taken beyond what they normally do • invited to begin their own walking program following study completion at 8 weeks Adherence: 32/36 (89%) reached an adherence rate of at least 50%. Mean (SD) number of complete exercise sessions 20.03 (6.60) Contamination of control group: not reported
Outcomes	Primary outcome • sleep quality, assessed using the PSQI Secondary outcomes included: • QoL, measured using the physical and mental component subscales of the MOS SF-36 Outcomes were measured at baseline, 1 month, and 2 months: • exercise group: n = 37 at baseline, n = 35 at 1 month, n = 36 at 2 months • control group: n = 35 at baseline, n = 35 at 1 month, n = 35 at 2 months Subgroup analysis: none specified Adverse events: no adverse events reported
Notes	Country: Taiwan Funding: not reported

Risk of bias

Bias	Authors' judgement	Support for judgement

Tang 2010 *(Continued)*

Random sequence generation (selection bias)	Low risk	Allocation sequence was generated using a table of random numbers
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Low risk	The study was analyzed on an ITT basis. Missing observations, including those incurred by participant drop-outs, were imputed by the "last observation carried forward" method. The disproportionate attrition from the intervention group places the study at a high risk of bias
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Targ 2002

Study characteristics	
Methods	Study design: RCT Number randomized: 181; 93 to the exercise group and 88 to the control group Study start and stop dates: not reported Length of intervention: 12 weeks Length of follow-up: to end of the intervention
Participants	Type cancer: breast cancer, Stage, n (%) • exercise group: Stage I, 31 (39%); Stage II, 35 (43%); Stage III, 5 (6%); Stage IV, 5 (6%); missing, 3 (4%) • control group: Stage I, 16 (32%); Stage II, 26 (52%); Stage III, 7 (14%); Stage IV, 2 (4%); missing, 5 (10%) Time since cancer diagnosis: within 18 months of diagnosis Time beyond active treatment: not reported, but some women were on chemotherapy, n (%): • exercise group: 42 (54%) • control group: 24 (48%) Inclusion criteria: • 26 to 78 years old • within 18 months of initial diagnosis of primary breast cancer or metastatic breast cancer Eligibility criterion related to interest or ability to exercise, or both: none reported

Targ 2002 *(Continued)*

Exclusion criteria: none reported

Gender: female

Current age, mean (SD) years:

- exercise group: 49 (8.6) years
- control group: 47 (8.8) years

Age at cancer diagnosis: not reported

Ethnicity/race, n (%):

- exercise group: Asian, 8 (11%); Hispanic, 1 (1%); African-American, 4 (5%); Caucasian, 62 (83%)
- control group: Asian, 2 (4%); Hispanic, 1 (2%); African-American, 4 (8%); Caucasian, 41 (85%)

Education level, n (%):

- exercise group: less than 8th grade, 1 (1%); some high school, 1 (1%); high school graduate, 0 (0%); some college, 13 (16%); college graduate, 22 (28%); postdoctorate study, 42 (53%)
- control group: less than 8th grade, 0 (0%); some high school, 0 (0%); high school graduate, 0 (0%); some college, 7 (14%); college graduate, 9 (18%); postdoctorate study, 34 (68%)

SES, income, n (%):

- exercise group: < USD15,000, 6 (8%); USD15,000 to USD29,000, 9 (12%); USD30,000 to USD44,000, 10 (13%); USD45,000 to USD49,000, 17 (22%); > USD50,000, 36 (46%)
- control group: < USD15,000, 2 (4%); USD15,000 to USD29,000, 5 (10%); USD30,000 to USD44,000, 9 (18%); USD45,000 to USD49,000, 5 (10%); > USD50,000, 28 (57%)

Employment status: not reported

Comorbidities: not reported

Past exercise history, number of days spent exercising and minutes of exercise, mean (SD) days and minutes:

- exercise group: 4.16 (1.82) days and 49.58 (23.09) minutes
- control group: 4.22 (1.62) days and 51.5 (25.64) minutes

On hormone therapy, n (%):

- exercise group: 28 (53)
- control group: 11 (48)

Postmenopausal status, n (%):

- exercise group: 17 (29)
- control group: 4 (14)

Interventions	93 participants assigned to an intensive lifestyle change and group support program that included: • weekly Health Series discussion group, followed by a 90-minute dance/movement program • weekly session consisting of silent meditation and guided imagery Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: mild to moderate Frequency: once per week Duration of individual sessions: 90 minutes Duration of exercise program: 12 weeks

Targ 2002 *(Continued)*

	Total number of exercise sessions: 12 Format: group Facility: facility Professionally led: nurse 88 participants assigned to the control group, including: • unstructured psycho-educational support group Adherence: 6 women did not attend any session, but no other adherence was noted Contamination of control group: not reported

Outcomes	Outcomes: QoL outcomes, including:
	
	• change in overall QoL, measured using FACIT and subscales
	  * PWB
	  * SWB
	  * EWB
	  * FWB
	  * additional concerns
	• change in mood as measured by the POMS and subscales
	  * anxiety
	  * depression
	  * anger
	  * vigor
	  * fatigue
	  * confusion
	• change in spiritual function, measured using the FACIT-Sp and the Principle of Living Survey
	
	Outcomes were measured at baseline and at 12 weeks:
	
	• exercise groups: n = 93 at baseline, n = 79 at 12 weeks
	• control group: n = 88 at baseline, n = 88 at 12 weeks
	
	Adverse events: none reported

Notes	Country: US Funding: United States Department of Defense Material Command

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The generation of the random sequence was not described
Allocation concealment (selection bias)	Unclear risk	Whether the treatment assignment was concealed from study personnel and participants was not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking

Targ 2002 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Although it was stated that an ITT analysis was performed, there were 7 women who dropped out in the intervention group and 24 in the control group and an additional 27 who did not attend any session and were not included in the analyses
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

Thorsen 2005

Study characteristics	
Methods	Study design: RCT Number randomized: 139; 69 to the exercise group and 70 to the control group Study start and stop dates: not reported Length of intervention: 14 weeks Length of follow-up: to end of the intervention
Participants	Type cancer, n: lymphoma, 25; breast, 42; gynecologic, 24; testicular, 20 • exercise group, n (%): breast, 21 (36%); gynecologic, 13 (22%); lymphoma, 14 (24%); testicular, 11 (19%) • control group, n (%): breast, 21 (40%); gynecologic, 11 (21%); lymphoma, 11 (21%); testicular, 9 (17%) Time since cancer diagnosis: not reported Time beyond active treatment, mean (SD) days: 28 (9) days Inclusion criteria: • primary treatment including chemotherapy for malignant lymphomas and breast, gynecologic, or testicular cancer • discontinuation of treatment approximately 1 month before baseline evaluation • 18 to 50 years old Eligibility criterion related to interest or ability to exercise, or both: • none reported Exclusion criteria: • major physical or mental comorbidity • evidence of disease at the time of intervention • geographical obstacles for repeated physical tests Gender, n (%): • exercise group: male, 19 (32%); female, 40 (68%)

Thorsen 2005 *(Continued)*

	• control group: male, 17 (33%); female, 35 (67%) Current age, mean (SD) years: • exercise group: 39.0 (8.4) years • control group: 39.1 (8.6) years Age at cancer diagnosis: 39 years Ethnicity/race: not reported Education level: not reported SES: not reported Employment status: not reported Other, BMI, mean (SD) kg/m^2: • exercise group: 25.4 (.7) • control group: 24.5 (3.6) Comorbidities: not reported Past exercise history: not reported On hormone therapy: not reported
Interventions	69 participants assigned to the exercise group, including: • home supervised exercise program, including walking, cycling, jogging, and ball games Type exercise (aerobic/anaerobic): aerobic Intensity of experimental exercise intervention: 13 to 15 on Borg perceived exertion scale (where the score range from 6 to 20, and a value of 6 means no exertion at all and 20 means maximal exertion) and 60% to 70% of maximum HR if using a HR monitor Frequency: twice a week Duration of sessions: 30 minutes Duration of program: 14 weeks Total number of exercise sessions: 28 Format: individual Facility: home Professionally led: exercise instructor Adherence: 30 (51%) of participants reported being "physically active" during the treatment period Co-intervention: none 70 participants were assigned to the control group, including: • instructed to "be as physically active as they would have been if they were not informed about this study" Contamination of control group: 24 (46%) of participants reported being "physically active" during the treatment period
Outcomes	Primary outcome: physical outcome, measured by:

Thorsen 2005 *(Continued)*

- change in cardiorespiratory fitness (maximum oxygen uptake), assessed by using the Åstrand-Rhyming indirect bicycle ergometer test

Secondary outcomes: QoL outcomes, including:

- mental distress, as assessed by the HADS, a self-rating scale developed to screen for levels of anxiety and depression, particularly in somatically ill patients. HADS consists of 14 items, 7 items for anxiety and 7 items for depression. Each item is scored from 0 (not present) to 3 (maximally present). Total HADS score ranges from 0 to 21 for each subscale, where 0 means no anxiety or depression symptoms and 21 defines the maximum of mental distress. If less than 3 responses were missing on each subscale, they were inputted by the individual mean scale value
- EORTC QLQ-C30, which comprise functional and symptom scales. The dimensions of physical function, emotional function, fatigue, and global QoL were considered in this trial. These scales were transferred to a 0 to 100 scale, which was calculated by using the scoring manual provided by the QLQ-C30. For the physical function, emotional function, and global QoL scales, a higher score indicates better level of functioning, whereas increasing values on the fatigue scale indicate more symptoms

Outcomes were measured at baseline and 14 weeks:

- exercise group: n = 69 at baseline, n = 59 at 14 weeks
- control group: n = 70 baseline, n = 52 at 14 weeks

Adverse events: none reported

Notes	Country: Norway Funding: Norwegian Foundation for Health and Rehabilitation and The Norwegian Cancer Society

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Computerized random assignment"
Allocation concealment (selection bias)	Low risk	"The Norwegian Radium Hospital was responsible for computerized random assignment"
Blinding of participants and personnel (performance bias) All outcomes	High risk	Owing to the nature of the intervention, it was not possible to blind the participants; however, it is unclear whether the outcome was influenced by a lack of masking
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study personnel and outcome assessors were not masked or blinded to the study interventions
Incomplete outcome data (attrition bias) All outcomes	High risk	Although the authors stated that the ITT analysis was used, 10 participants in the exercise group withdrew and 18 participants in the control group withdrew
Selective reporting (reporting bias)	Low risk	There appears to be no selective reporting of outcomes
Other bias	Low risk	The trial appears to be free of other problems that could put it at a high risk of bias

1-RM: 1-repetition maximum; ACSM: American College of Sports Medicine; AIDS: acquired immune deficiency syndrome; ALL: acute lymphoblastic leukemia; AML: acute myeloid leukemia; BDI: Beck Depression Inventory; BES: Body Esteem Scale; BIQ: Body Image

Questionnaire; BMI: body mass index; CARES-SF: Cancer Rehabilitation Evaluation System Short Form; CCS: Canadian Cancer Society; CES-D: Centers for Epidemiologic Studies - Depression scale; CI: confidence interval; DCIS: ductal carcinoma in situ; ECOG: Eastern Cooperative Oncology Group; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core Module; EWB: emotion well-being; FACIT-F: Functional Assessment in Chronic Illness Therapy - Fatigue; FACIT-Sp: Functional Assessment of Chronic Illness Therapy - Spiritual; FACT-An: Functional Assessment of Cancer Therapy - Anemia; FACT-B: Functional Assessment of Cancer Therapy - Breast; FACT-C: Functional Assessment of Cancer Therapy - Colorectal; FACT-Cog: Functional Assessment of Cancer Therapy - Cognitive Function; FACT-ES: Functional Assessment of Cancer Therapy - Endocrine Symptoms; FACT-F: Functional Assessment of Cancer Therapy - Fatigue; FACT-G: Functional Assessment of Cancer Therapy - General; FACT-Lym: Functional Assessment of Cancer Therapy - Lymphoma; FACT-Sp: Functional Assessment of Cancer Therapy - Spirituality; FWB: functional well-being; GED: General Education Diploma; GVHD: graft-versus-host disease; HADS: Hospital Anxiety and Depression Scale; HL: Hodgkin lymphoma; HNC: head and neck cancer; HR: heart rate; HRQoL: health-related quality of life; HSCT: hematopoietic stem cell transplant; ITT: intention to treat; KPS: Karnofsky Performance Status; LASA: Linear Analog Self-Assessment; LSI: Leisure Score Index; MAC: Mental Adjustment to Cancer; MFI: Multidimensional Fatigue Inventory; MOS SF-12: Medical Outcomes 12-Item Short Form Health Survey; MOS SF-36: Medical Outcomes 36-Item Short Form Health Survey; MCS: mental component status; MFSI-SF: Multidimensional Fatigue Symptom Inventory Short Form; NCIC: National Cancer Institute of Canada; NHL: non-Hodgkin lymphoma; NIH: National Institutes of Health; NYHA: New York Heart Association; PANAS: Positive and Negative Affect Schedule; PCS: physical component status; POMS: Profile of Moods Scale; PRET: progressive resistance exercise training; PSQI: Pittsburgh Sleep Quality Index; PWB: physical well-being; QoL: quality of life; RCT: randomized controlled trial; ROM: range of motion; rPAR-Q: revised Physical Activity Readiness Questionnaire; RPE: ratings of perceived exertion; RSE: Rosenberg Self-Esteem; SCFC: Schwartz Cancer Fatigue Scale; SCL-90-R: Symptom Checklist-90 Revised; SD: standard deviation; SES: socioeconomic status; SFT: Submaximal Fitness Test; SOSI: Symptoms of Stress Inventory; SPADI: Shoulder Pain and Disability Index; SPAS-7: Social Physique Anxiety Scale-7 items; STAI: State-Trait Anxiety Index; SWB: social/family well-being; SWLS: Satisfaction with Life Scale; TOI: Trial Outcome Index; TOI-An: Trial Outcome Index - Anemia; TTM: Trans Theoretical Model; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Basen-Engquist 2006	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Beurskens 2007	This study was excluded as the exercise was aimed toward reduction in shoulder dysfunction rather than for improvement in whole body function or QoL
Blanchard 2001	This study was excluded as it was not an RCT or a CTT
Bloom 2008	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Braam 2010	This study was excluded as it did not exclude people below the age of 18 years
Carmack Taylor 2004	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Carmack Taylor 2006	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Carmack Taylor 2007	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Carson 2009	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Cheema 2006	This study was excluded as it was not an RCT or a CCT and it did not compare an exercise with no exercise, another intervention, or usual care
Cheung 2003	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care

Study	Reason for exclusion
Cinar 2008	This study was excluded as the exercise was aimed toward reduction in lymphedema and improvement in shoulder mobility rather than for improvement in whole body function or QoL
Courneya 2004a	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Courneya 2004b	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Courneya 2004c	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Courneya 2005	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Courneya 2008	This study was excluded as the majority (> 90%) of participants were undergoing chemotherapy
Daley 2007	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Demark-Wahnefried 2003	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Demark-Wahnefried 2003a	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Demark-Wahnefried 2006	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Demark-Wahnefried 2007	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Dincer 2007	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Dong 2006	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Duijts 2009	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Elkin 1998	This study was excluded as it was not an RCT or a CCT, it did not exclude people below the age of 18 years; it did not compare an exercise with no exercise, another intervention, or usual care; and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Elliott 2006	This study was excluded as it did not include cancer survivors and it did not compare an exercise with no exercise, another intervention, or usual care
Emslie 2007	This study was excluded as it was not a randomized controlled trial or a controlled clinical trial; did not compare an exercise with no exercise, another intervention, or usual care; and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Fairey 2005	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome

Study	Reason for exclusion
Fairey 2005a	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Filocamo 2005	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Gordon 2005	This study was excluded as it was not an RCT or a CCT
Griffith 2009	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Hayes 2004	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer, and it did not exclude people below the age of 18 years
Houborg 2006	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Hughes 2004	This study was excluded as it was not an RCT or a CCT and it did not compare an exercise with no exercise, another intervention, or usual care and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Hughes 2008	This study was excluded as it was not an RCT or a CCT and it did not compare an exercise with no exercise, another intervention, or usual care
Jarden 2009	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Jones 2004a	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Jones 2008	This study was excluded as it was not a randomized controlled trial or a controlled clinical trial, and it did not compare an exercise with no exercise, another intervention, or usual care
Kilbreath 2006	This study was excluded as the exercise was aimed toward reduction in shoulder dysfunction rather than for improvement in whole body function or QoL
Kim 2010	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Kim 2011	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Kolden 2002	This study was excluded as it was not an RCT or a CCT, and it did not compare an exercise with no exercise, another intervention, or usual care
Korstjens 2008	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Lazowski 1999	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Ligibel 2008	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Livingston 2011	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care

Study	Reason for exclusion
Mansky 2006	This study was excluded as it was not an RCT or a CCT; it did not compare an exercise with no exercise, another intervention, or usual care; and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Mathewson-Chapman 1997	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care, and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Matthews 2007	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care, and it did not measure overall HRQoL or an HRQoL domain as a study outcome
May 2008	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
McClure 2010	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
McKenzie 2003	This study was excluded as the exercise was aimed toward reduction in lymphedema rather than for improvement in whole body function or QoL
McNeely 2004	This study was excluded as the exercise was aimed toward reduction in shoulder dysfunction rather than for improvement in whole body function or QoL
Midtgaard 2006	This study was excluded as it was not an RCT or a CCT; it did not exclude people who were undergoing active cancer treatment for either the primary or recurrent cancer; it did not compare an exercise with no exercise, another intervention, or usual care; and the exercise intervention was initiated before completion of active treatment
Milne 2008	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Morey 2009	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Mustian 2006	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Mutrie 2007	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Nikander 2007	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Osei-Tutu 2005	This study was excluded as the participants were not cancer survivors or undergoing active treatment for cancer
Pinto 2009	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Poorkiani 2010	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Rabin 2006	This study was excluded as it was not an RCT or a CCT and it did not compare an exercise with no exercise, another intervention, or usual care
Sandel 2005	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care

Study	Reason for exclusion
Schneider 2007	This study was excluded as it was not an RCT or a CCT, and it did not compare an exercise with no exercise, another intervention, or usual care
Segal 2001	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Segal 2003	This study was excluded as all the participants were undergoing active cancer treatment for either the primary or recurrent cancer
Snyder 2009	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Sprod 2005	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Twiss 2009	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Vallance 2007	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care, and it did not measure overall HRQoL or an HRQoL domain as a study outcome
Vallance 2007a	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Vallance 2008	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Vallance 2008a	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care and it did not measure overall HRQoL or an HRQoL domain as a study outcome
van Weert 2005	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
van Weert 2010	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
von Gruenigen 2009	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care
Wall 2000	This study was excluded as it did not measure overall HRQoL or an HRQoL domain as a study outcome
Yeh 2011	This study was excluded as it was not an RCT or a CCT, and it did not compare an exercise with no exercise, another intervention, or usual care
Zhang 2006	This study was excluded as it did not compare an exercise with no exercise, another intervention, or usual care

CCT: controlled clinical trial; HRQoL: health-related quality of life; QoL: quality of life; RCT: randomized controlled trial.

Characteristics of studies awaiting classification *[ordered by study ID]*

Utz-Billing 2010

Methods	Randomized controlled trial

Utz-Billing 2010 *(Continued)*

Participants	Women who have had breast cancer surgery (breast-sparing therapy or mastectomy)
Interventions	Yoga classes compared with waiting list controls
Outcomes	QoL assessed by EORTC QLQ-C23, physical function assessed by the FACT-B, and disabilities of the upper limbs
Notes	Published abstract

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core Module; FACT-B: Functional Assessment of Cancer Therapy - Breast; QoL: quality of life;

Characteristics of ongoing studies *[ordered by study ID]*

Devonish 2007

Study name	Physical activity for cancer survivors
Methods	RCT
Participants	Mixed cancer survivors
Interventions	A 16-week home-based physical activity program, which was supplemented with biweekly group sessions compared with waiting list control
Outcomes	Fitness training, physical activity levels, QoL assessed using the EORTC QLQ-C30, mood state assessed using the POMS scale
Starting date	2007
Contact information	Julia A. Devonish University of Calgary
Notes	Published abstract

Galvao 2009

Study name	RADAR
Methods	RCT
Participants	Cohort undergoing or previously treated for prostate cancer involving androgen deprivation therapy
Interventions	Supervised resistance/aerobic exercise compared with standard physical activity recommendation
Outcomes	Outcomes include aerobic walking capacity, anthropometric measures (abdominal obesity), various blood markers, self-reported physical activity, HRQoL assessed using the EORTC QLQ-C30, falls self-efficacy assessed using the activities-specific balance, psychological distress assessed using the BSI, nutrition, lower body physical function. In addition, at 1 of the study sites, additional outcomes assessed include body composition, muscle strength, balance and risk of falling
Starting date	2009

Galvao 2009 *(Continued)*

Contact information	Daniel A. Galvao, Vario Health Institute, School of Exercise, Biomedical and Health Sciences, Edith Cowan University, Joondalup, WA, Australia
Notes	Published protocol. Trial Registration: ACTRN 12609000729224

Hayes 2011

Study name	Exercise for Health
Methods	RCT
Participants	Postsurgical women with a first diagnosis of invasive breast cancer
Interventions	Patient-centered aerobic and strength-based exercise program, delivered either face-to-face or by telephone compared with usual care
Outcomes	The primary outcome is HRQoL assessed using the FACT-B questionnaire
Starting date	October 2006. Recruitment ended June 2008
Contact information	Sandra Hayes, School of Public Health, Queensland University of Technology
Notes	Published protocol and baseline characteristics. Registered at ANZCTR (ACTRN12609000809235)

Jones 2010

Study name	Exercise Intensity Trial (EXCITE)
Methods	RCT
Participants	Postmenopausal with histologically confirmed breast cancer following completion of primary therapy
Interventions	Moderate intensity aerobic training performed on a treadmill, moderate to high intensity aerobic training performed on a motorized treadmill, or attention control comprised of stretching
Outcomes	The primary outcome is $VO_{2\ peak}$ and the secondary outcomes include physiologic determinants of $VO_{2\ peak}$, pulmonary function, cardiovascular O_2 delivery, brachial artery endothelial function, QoL, fatigue, and depression
Starting date	August 2010
Contact information	Lee W. Jones, Duke University Medical Center, Durham NC USA
Notes	Published protocol. Registered at ClinicalTrials.gov (NCT0118367)

Jones 2010a

Study name	Lung Cancer Exercise Training Study (LUNGEVITY)

Jones 2010a *(Continued)*

Methods	RCT
Participants	Postoperative small cell lung cancer patients who had surgery between 6 and 36 months earlier
Interventions	Aerobic training on either treadmill or cycle ergometer, resistance training using stationary weight machines, combination training, or an attention control group, comprised of progressive stretching and social interaction
Outcomes	The primary outcome is $VO_{2\ peak}$ and secondary outcomes include QoL, fatigue, dyspnea, and depression
Starting date	January 2010
Contact information	Lee W. Jones, Duke University Medical Center, Durham NC USA
Notes	Published protocol. Registered at ClinicalTrials.gov (NCT01068210)

Kampshoff 2010

Study name	Resistance and Endurance exercise After Chemotherapy (REACT)
Methods	RCT
Participants	Patients with histologically confirmed primary breast, colon or ovarian cancer, or lymphomas with no indication of recurrent or progressive disease, who have completed adjuvant chemotherapy with curative intention
Interventions	High-intensity resistance training, low to moderate intensity resistance training, or behavioral motivational counseling
Outcomes	The primary outcome is cardiorespiratory fitness, muscle strength, and fatigue. Secondary outcomes are QoL, body composition, bone mineral density, neuropathy, objective and self-reported physical activity, mood and sleep disturbance, functioning in daily life, return to work, cost from a social perspective, adverse events, compliance, and satisfaction with the intervention
Starting date	March 2010
Contact information	Caroline S Kampshof EMGO Institute for Health and Care Research, Department of Public and Occupational Health, VU University Medical Center, Amsterdam, The Netherlands
Notes	Published protocol. Registered at the Netherlands Trial Register NTR2153

Persoon 2010

Study name	Exercise Intervention after Stem Cell Transplantation (EXIST)
Methods	Pilot study followed by RCT
Participants	Patients with multiple myeloma or non-Hodgkin lymphoma
Interventions	18-week high-intensity resistance and interval training physical exercise program and counseling compared with usual care

Persoon 2010 *(Continued)*

Outcomes	Primary outcome is cardiorespiratory fitness, muscle strength, and fatigue. Secondary outcomes are body composition, bone mineral density, QoL, neuropathy, objective and self reported physical activity level, mood disturbance, functioning in daily life, return to work, and cost from a social perspective
Starting date	March 2010
Contact information	Marie Jose Kersten, Department of Hematology, Academic Medical Center, University of Amsterdam, The Netherlands
Notes	Published protocol. Registered at the Netherlands Trial Register (NTR2341)

Saxton 2006

Study name	Effect of a lifestyle intervention in women recovering from breast cancer treatment
Methods	RCT
Participants	Women who have undergone appropriate treatment for operable breast cancer and are no longer undergoing chemotherapy or radiation therapy
Interventions	lifestyle intervention (incorporating dietary energy restriction in conjunction with aerobic exercise training compared with usual care
Outcomes	Primary outcome measures include body weight and body composition. Secondary outcomes include psychological stress assessed using the Perceived Stress Scale, depression assessed using the Beck Depression Inventory, cardiorespiratory fitness, QoL assessed using the FACT-G and FACT-B, physical activity behavior, and biomarkers associated with disease recurrence and physiologic health status
Starting date	2006
Contact information	John M Saxton Center for Sport and Exercise Science, Sheffield Hallam University, Sheffield, UK
Notes	Published protocol. Trial Registration: ISRCTN08045231

Sekse 2011

Study name	Rehabilitation of women following treatment for gynaecological cancer
Methods	RCT
Participants	Women who have completed curative treatment for gynecologic cancer
Interventions	Group physical training versus education and counseling versus control
Outcomes	Outcomes include QoL (global and health-related), coping, fatigue, sexuality, anxiety and depression
Starting date	2011
Contact information	Ragnhild Johanne Tveit Sekse, Haukeland University Hospital, Bergen, Norway

Sekse 2011 *(Continued)*

Notes	Published abstract

Spence 2007

Study name	Supervised exercise rehabilitation program for colorectal cancer survivors
Methods	RCT
Participants	Patients with colorectal cancer who have had surgery and are due to complete adjuvant chemotherapy
Interventions	ImPACT program (I'm Physically Active after Cancer Treatment), an aerobic exercise session versus usual care
Outcomes	The primary outcome is fatigue and secondary outcomes include cardiorespiratory fitness, biomarkers of health, QoL, program acceptance, adherence and compliance, and safety
Starting date	January 2006
Contact information	Rosalind R. Spence, School of Human Movement Studies, The University of Queensland, Brisbane, Australia
Notes	Published protocol. Registered at the Australian New Zealand Trial Register (ACTRN 012606000395538)

Vardy 2010

Study name	CHALLENGE
Methods	RCT
Participants	Patients with resected stage II or III colon cancer who have completed adjuvant therapy from Australia and Canada
Interventions	36 month physical activity program versus standard of care
Outcomes	Outcomes include fatigue, QoL, depression, anxiety, sleep, body composition, exercise behavior and fitness, as well as an economic evaluation of the program
Starting date	2010
Contact information	J. Vardy, University of Sydney, NSW, Australia
Notes	Published abstract

Walsh 2010

Study name	Prescribed Exercise After CHemotherapy (PEACH)
Methods	RCT

Walsh 2010 *(Continued)*

Participants	Patients with diagnosis of solid tumor and completion of adjuvant chemotherapy or radiation therapy with curative intent within the preceding 2 to 4 months
Interventions	Exercise training using treadmills, cycle ergometers, and rowing machine or stepper or other aerobic exercise versus usual care
Outcomes	The primary outcome is physical fitness and secondary outcomes include QoL, current activity level, cancer-related fatigue, and satisfaction
Starting date	January 2010
Contact information	Julie M. Walsh, Discipline of Physiotherapy, Trinity Centre for Higher Sciences, St. James's Hospital, Dublin, Ireland
Notes	Published protocol. Registered at ClinicalTrials.gov (NCT01030887)

BSI: Brief Symptom Inventory; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core Module; FACT-B: Functional Assessment of Cancer Therapy - Breast; FACT-G: Functional Assessment of Cancer Therapy - General; HRQoL: health-relate quality of life; QoL: quality of life; POMS: Profile of Mood State; RCT: randomized controlled trial

DATA AND ANALYSES

Comparison 1. Health-related quality of life

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Overall quality of life change score	15		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1.1 Up to 12 weeks' follow-up	11	826	Std. Mean Difference (IV, Random, 95% CI)	0.48 [0.16, 0.81]
1.1.2 More than 12 weeks' to less than 6 months' follow-up	3	181	Std. Mean Difference (IV, Random, 95% CI)	0.14 [-0.38, 0.66]
1.1.3 6 months' follow-up	2	115	Std. Mean Difference (IV, Random, 95% CI)	0.46 [0.09, 0.84]
1.2 Overall quality of life values	20		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.2.1 Up to 12 weeks' follow-up	16	760	Std. Mean Difference (IV, Random, 95% CI)	0.49 [0.24, 0.74]
1.2.2 More than 12 weeks' to less than 6 months' follow-up	5	353	Std. Mean Difference (IV, Random, 95% CI)	0.11 [-0.10, 0.32]
1.2.3 6 months' follow-up	2	115	Std. Mean Difference (IV, Random, 95% CI)	0.25 [-0.12, 0.62]
1.3 FACT An change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only

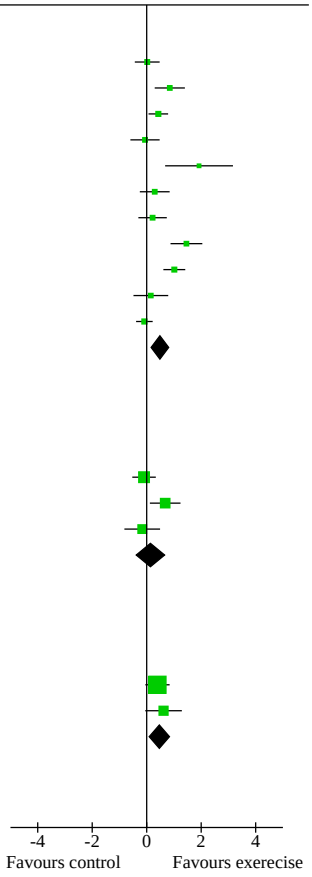
Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1.3.1 Up to 12 weeks' follow-up	2	183	Mean Difference (IV, Random, 95% CI)	7.10 [1.50, 12.71]
1.4 FACT-An follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.4.1 Up to 12 weeks' follow-up	2	174	Mean Difference (IV, Random, 95% CI)	4.25 [-3.28, 11.78]
1.5 FACT-B change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.5.1 Up to 12 weeks' follow-up	2	96	Mean Difference (IV, Random, 95% CI)	9.29 [-3.73, 22.30]
1.5.2 More than 12 weeks' to less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	8.80 [2.34, 15.26]
1.5.3 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	8.20 [-0.29, 16.69]
1.6 FACT-B follow-up values	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.6.1 Up to 12 weeks' follow-up	4	188	Mean Difference (IV, Random, 95% CI)	9.82 [-0.76, 20.40]
1.6.2 More than 12 weeks' to less than 6 months' follow-up	2	112	Mean Difference (IV, Random, 95% CI)	4.83 [-1.71, 11.36]
1.6.3 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	1.90 [-11.33, 15.13]
1.7 FACT-C change	1	93	Mean Difference (IV, Random, 95% CI)	-1.30 [-7.19, 4.59]
1.7.1 More than 12 weeks' to less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-1.30 [-7.19, 4.59]
1.8 FACT-C follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.8.1 Up to 12 weeks' follow-up	1	17	Mean Difference (IV, Random, 95% CI)	14.00 [2.59, 25.41]
1.8.2 More than 12 weeks' to less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-2.40 [-10.19, 5.39]
1.9 FACT-G change	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.9.1 Up to 12 weeks' follow-up	3	198	Mean Difference (IV, Random, 95% CI)	4.94 [-0.08, 9.95]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1.9.2 More than 12 weeks' to less than 6 months' follow-up	2	145	Mean Difference (IV, Random, 95% CI)	2.10 [-3.98, 8.17]
1.9.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	3.93 [0.45, 7.40]
1.10 FACT-G follow-up values	9		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.10.1 Up to 12 weeks' follow-up	6	318	Mean Difference (IV, Random, 95% CI)	5.90 [-0.36, 12.16]
1.10.2 More than 12 weeks' to less than 6 months' follow-up	3	207	Mean Difference (IV, Random, 95% CI)	1.79 [-1.93, 5.50]
1.10.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	3.73 [-2.08, 9.53]
1.11 FACIT change	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
1.11.1 Up to 12 weeks' follow-up	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
1.12 FACIT follow-up values	1	21	Mean Difference (IV, Random, 95% CI)	-2.60 [-21.19, 15.99]
1.12.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	-2.60 [-21.19, 15.99]
1.13 QLQ-C30 change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.13.1 Up to 12 weeks' follow-up	2	73	Mean Difference (IV, Random, 95% CI)	15.66 [-7.78, 39.09]
1.13.2 More than 12 weeks' less than 6 months' follow-up	1	101	Mean Difference (IV, Random, 95% CI)	-1.60 [-8.15, 4.95]
1.14 QLQ-C30 follow-up values	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.14.1 Up to 12 weeks' follow-up	3	109	Mean Difference (IV, Fixed, 95% CI)	13.80 [7.17, 20.43]
1.14.2 More than 12 weeks' less than 6 months' follow-up	1	101	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-7.13, 6.13]
1.15 CARES follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.15.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-3.20 [-7.19, 0.79]
1.16 CARES change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1.16.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-1.70 [-3.58, 0.18]
1.17 Chae and Cho change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.17.1 Up to 12 weeks' follow-up	1	55	Mean Difference (IV, Random, 95% CI)	1.00 [0.39, 1.61]
1.18 Chae and Cho follow-up values	1	55	Mean Difference (IV, Random, 95% CI)	0.70 [0.01, 1.39]
1.18.1 Up to 12 weeks' follow-up	1	55	Mean Difference (IV, Random, 95% CI)	0.70 [0.01, 1.39]
1.19 QoL Index follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.19.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	3.30 [-5.35, 11.95]

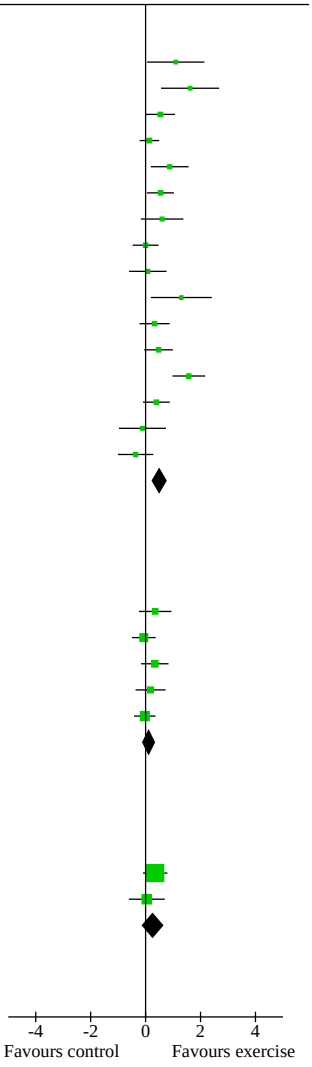
Analysis 1.1. Comparison 1: Health-related quality of life, Outcome 1: Overall quality of life change score

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Up to 12 weeks' follow-up								
Cadmus 2009	0.3	5.7	37	0.2	4.4	37	9.9%	0.02 [-0.44 , 0.48]
Cho 2006	0.9	1.3	28	-0.1	1	27	9.0%	0.85 [0.29 , 1.40]
Courneya 2009	10.6	22.1	60	1.1	22.1	62	10.7%	0.43 [0.07 , 0.79]
Dimeo 2004	17	51	62	20	36	17	9.2%	-0.06 [-0.60 , 0.48]
Herrero 2006	17.7	8.3	8	-10.4	17.7	8	4.3%	1.92 [0.68 , 3.17]
McNeely 2008a	4.4	10.6	27	1.7	6.9	25	9.1%	0.29 [-0.25 , 0.84]
Mehnert 2011	9.72	21.03	30	5.55	16.5	27	9.3%	0.22 [-0.31 , 0.74]
Milne 2008a	12.6	12.6	29	-3	8	29	8.8%	1.46 [0.87 , 2.04]
Oh 2010	8.86	9	54	-0.13	8.6	54	10.3%	1.01 [0.61 , 1.42]
Rogers 2009	4.5	8.4	20	2.9	12	18	8.3%	0.15 [-0.49 , 0.79]
Targ 2002	4.98	16.1	79	6.62	21.7	88	11.1%	-0.08 [-0.39 , 0.22]
Subtotal (95% CI)			**434**			**392**	**100.0%**	**0.48 [0.16 , 0.81]**
Heterogeneity: Tau² = 0.22; Chi² = 46.35, df = 10 (P < 0.00001); I² = 78%								
Test for overall effect: Z = 2.95 (P = 0.003)								
1.1.2 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003b	0.9	12.2	62	2	9	31	38.7%	-0.10 [-0.53 , 0.33]
Courneya 2003c	5.7	7.4	24	0.6	7.4	28	32.6%	0.68 [0.12 , 1.24]
Thorsen 2005	8.8	9.4	18	10.4	9.7	18	28.7%	-0.16 [-0.82 , 0.49]
Subtotal (95% CI)			**104**			**77**	**100.0%**	**0.14 [-0.38 , 0.66]**
Heterogeneity: Tau² = 0.13; Chi² = 5.49, df = 2 (P = 0.06); I² = 64%								
Test for overall effect: Z = 0.52 (P = 0.61)								
1.1.3 6 months' follow-up								
Ohira 2006	2.3	4.5	39	0.6	4	40	69.4%	0.40 [-0.05 , 0.84]
Rogers 2009	3.4	11	19	-3.5	10.8	17	30.6%	0.62 [-0.05 , 1.29]
Subtotal (95% CI)			**58**			**57**	**100.0%**	**0.46 [0.09 , 0.84]**
Heterogeneity: Tau² = 0.00; Chi² = 0.29, df = 1 (P = 0.59); I² = 0%								
Test for overall effect: Z = 2.45 (P = 0.01)								
Test for subgroup differences: Chi² = 1.34, df = 2 (P = 0.51), I² = 0%								



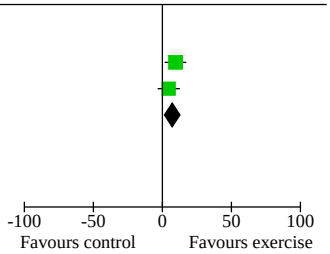
Analysis 1.2. Comparison 1: Health-related quality of life, Outcome 2: Overall quality of life values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.2.1 Up to 12 weeks' follow-up								
Bourke 2011	120	11	8	106	13	9	3.8%	1.10 [0.06 , 2.14]
Burnham 2002	95.1	4.4	14	86.6	6.2	7	3.7%	1.62 [0.56 , 2.68]
Cho 2006	7	1.3	28	6.3	1.3	27	7.2%	0.53 [-0.01 , 1.07]
Courneya 2009	151.4	21.7	60	148.1	25.7	62	8.9%	0.14 [-0.22 , 0.49]
Culos-Reed 2006	78.2	20.24	18	62.05	15.46	18	6.0%	0.88 [0.19 , 1.56]
Daley 2007a	90.85	13.47	33	83.15	14.72	33	7.7%	0.54 [0.05 , 1.03]
Danhauer 2009	114.8	19.1	13	98.4	31.8	14	5.3%	0.60 [-0.17 , 1.38]
Dimeo 2004	62	20	34	62	19	35	7.8%	0.00 [-0.47 , 0.47]
Donnelly 2011	80.19	16.93	16	78.71	19.1	17	6.0%	0.08 [-0.60 , 0.76]
Herrero 2006	81.3	10.7	8	62.5	16.1	8	3.5%	1.30 [0.19 , 2.41]
McNeely 2008a	83.9	15.6	27	78.1	19.3	25	7.1%	0.33 [-0.22 , 0.87]
Mehnert 2011	74.72	21.72	30	65.43	16.78	28	7.4%	0.47 [-0.05 , 0.99]
Milne 2008a	80.7	7.1	29	66.3	10.6	29	6.7%	1.57 [0.98 , 2.17]
Moadel 2007	78.07	17.17	45	70.38	22.75	26	7.7%	0.39 [-0.09 , 0.88]
Mustian 2004	121.7	20.8	11	124.3	22.5	10	4.8%	-0.12 [-0.97 , 0.74]
Rogers 2009	87.4	13.1	20	92	11.4	18	6.3%	-0.37 [-1.01 , 0.28]
Subtotal (95% CI)			**394**			**366**	**100.0%**	**0.49 [0.24 , 0.74]**
Heterogeneity: Tau² = 0.15; Chi² = 39.60, df = 15 (P = 0.0005); I² = 62%								
Test for overall effect: Z = 3.86 (P = 0.0001)								
1.2.2 More than 12 weeks' to less than 6 months' follow-up								
Bai 2004	61.15	26.57	24	52.46	21.71	21	13.0%	0.35 [-0.24 , 0.94]
Courneya 2003b	85.6	13	62	86.5	15	31	24.3%	-0.07 [-0.50 , 0.37]
Daley 2007a	90	14.9	31	85	14.69	31	18.0%	0.33 [-0.17 , 0.84]
Donnelly 2011	91.3	11	24	89.3	10.9	28	15.1%	0.18 [-0.37 , 0.73]
Thorsen 2005	74.2	17.3	52	74.7	16.7	49	29.7%	-0.03 [-0.42 , 0.36]
Subtotal (95% CI)			**193**			**160**	**100.0%**	**0.11 [-0.10 , 0.32]**
Heterogeneity: Tau² = 0.00; Chi² = 2.58, df = 4 (P = 0.63); I² = 0%								
Test for overall effect: Z = 1.00 (P = 0.32)								
1.2.3 6 months' follow-up								
Ohira 2006	-44.2	8.7	39	-47.4	9.4	40	68.4%	0.35 [-0.09 , 0.79]
Rogers 2009	86.1	13.6	19	85.4	18.4	17	31.6%	0.04 [-0.61 , 0.70]
Subtotal (95% CI)			**58**			**57**	**100.0%**	**0.25 [-0.12 , 0.62]**
Heterogeneity: Tau² = 0.00; Chi² = 0.58, df = 1 (P = 0.45); I² = 0%								
Test for overall effect: Z = 1.35 (P = 0.18)								
Test for subgroup differences: Chi² = 5.29, df = 2 (P = 0.07), I² = 62.2%								



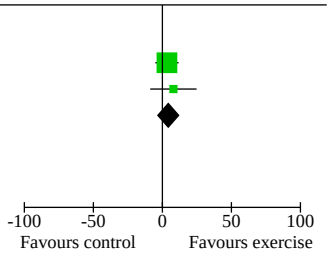
Analysis 1.3. Comparison 1: Health-related quality of life, Outcome 3: FACT An change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.3.1 Up to 12 weeks' follow-up								
Courneya 2009	10.6	22.1	60	1.1	22.1	62	51.1%	9.50 [1.66 , 17.34]
McNeely 2008a	8.5	19.3	27	3.9	10	34	48.9%	4.60 [-3.42 , 12.62]
Subtotal (95% CI)			**87**			**96**	**100.0%**	**7.10 [1.50 , 12.71]**
Heterogeneity: Tau² = 0.00; Chi² = 0.73, df = 1 (P = 0.39); I² = 0%								
Test for overall effect: Z = 2.48 (P = 0.01)								
Test for subgroup differences: Not applicable								



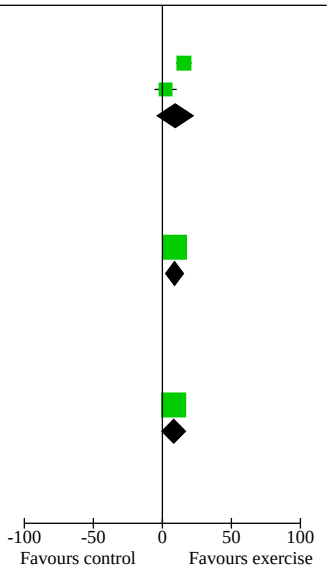
Analysis 1.4. Comparison 1: Health-related quality of life, Outcome 4: FACT-An follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.4.1 Up to 12 weeks' follow-up								
Courneya 2009	151.4	21.7	60	148.1	25.7	62	79.8%	3.30 [-5.13 , 11.73]
McNeely 2008a	142.4	27	27	134.4	34	25	20.2%	8.00 [-8.77 , 24.77]
Subtotal (95% CI)			**87**			**87**	**100.0%**	**4.25 [-3.28 , 11.78]**
Heterogeneity: Tau² = 0.00; Chi² = 0.24, df = 1 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.11 (P = 0.27)								
Test for subgroup differences: Not applicable								



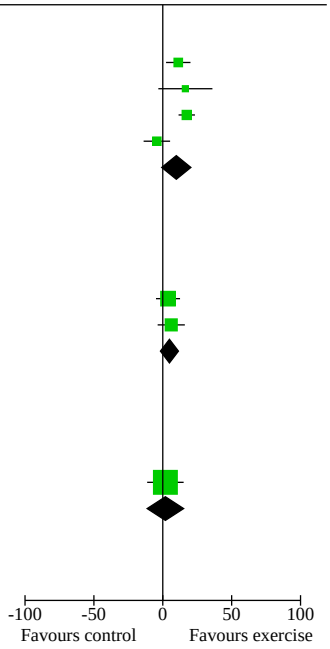
Analysis 1.5. Comparison 1: Health-related quality of life, Outcome 5: FACT-B change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.5.1 Up to 12 weeks' follow-up								
Milne 2008a	12.6	12.6	29	-3	8.4	29	52.5%	15.60 [10.09 , 21.11]
Rogers 2009	5.4	10.8	20	3.1	14.1	18	47.5%	2.30 [-5.75 , 10.35]
Subtotal (95% CI)			**49**			**47**	**100.0%**	**9.29 [-3.73 , 22.30]**
Heterogeneity: Tau² = 76.05; Chi² = 7.14, df = 1 (P = 0.008); I² = 86%								
Test for overall effect: Z = 1.40 (P = 0.16)								
1.5.2 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003c	9.1	14.1	24	0.3	8.5	28	100.0%	8.80 [2.34 , 15.26]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**8.80 [2.34 , 15.26]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.67 (P = 0.008)								
1.5.3 6 months' follow-up								
Rogers 2009	4.1	12	19	-4.1	13.8	17	100.0%	8.20 [-0.29 , 16.69]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**8.20 [-0.29 , 16.69]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.89 (P = 0.06)								
Test for subgroup differences: Chi² = 0.02, df = 2 (P = 0.99), I² = 0%								



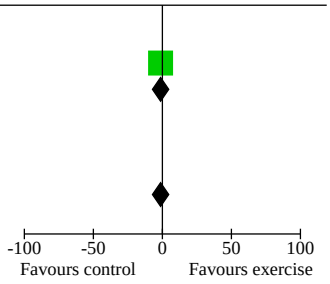
Analysis 1.6. Comparison 1: Health-related quality of life, Outcome 6: FACT-B follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.6.1 Up to 12 weeks' follow-up								
Daley 2007a	115.09	17.6	32	103.85	18.54	32	27.3%	11.24 [2.38 , 20.10]
Danhauer 2009	114.8	19.1	13	98.4	31.8	14	15.6%	16.40 [-3.23 , 36.03]
Milne 2008a	102.3	8.4	29	84.9	14	29	30.6%	17.40 [11.46 , 23.34]
Rogers 2009	114.4	16.2	20	118.7	14.4	19	26.5%	-4.30 [-13.91 , 5.31]
Subtotal (95% CI)			**94**			**94**	**100.0%**	**9.82 [-0.76 , 20.40]**
Heterogeneity: Tau² = 86.08; Chi² = 14.41, df = 3 (P = 0.002); I² = 79%								
Test for overall effect: Z = 1.82 (P = 0.07)								
1.6.2 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003c	119.6	16.9	24	115.8	14.9	28	56.1%	3.80 [-4.93 , 12.53]
Daley 2007a	115.39	20.11	30	109.25	18.85	30	43.9%	6.14 [-3.72 , 16.00]
Subtotal (95% CI)			**54**			**58**	**100.0%**	**4.83 [-1.71 , 11.36]**
Heterogeneity: Tau² = 0.00; Chi² = 0.12, df = 1 (P = 0.73); I² = 0%								
Test for overall effect: Z = 1.45 (P = 0.15)								
1.6.3 6 months' follow-up								
Rogers 2009	112.6	17.6	19	110.7	22.3	17	100.0%	1.90 [-11.33 , 15.13]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**1.90 [-11.33 , 15.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.28 (P = 0.78)								
Test for subgroup differences: Chi² = 0.97, df = 2 (P = 0.62), I² = 0%								



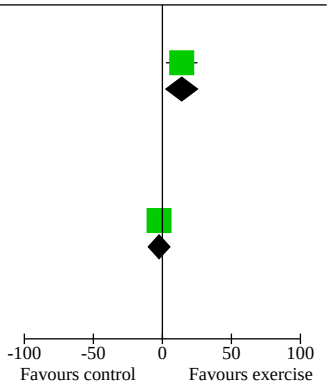
Analysis 1.7. Comparison 1: Health-related quality of life, Outcome 7: FACT-C change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.7.1 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003b	1.4	15.7	62	2.7	12.5	31	100.0%	-1.30 [-7.19 , 4.59]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-1.30 [-7.19 , 4.59]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.43 (P = 0.67)								
Total (95% CI)			**62**			**31**	**100.0%**	**-1.30 [-7.19 , 4.59]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.43 (P = 0.67)								
Test for subgroup differences: Not applicable								



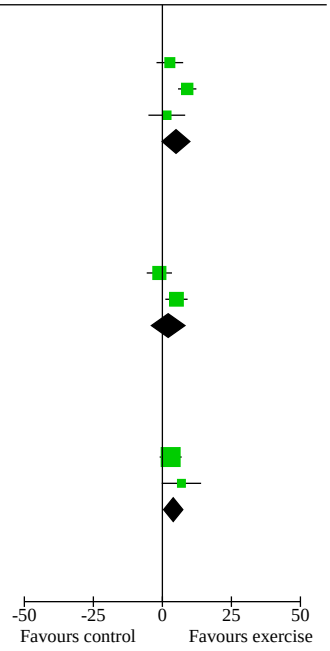
Analysis 1.8. Comparison 1: Health-related quality of life, Outcome 8: FACT-C follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.8.1 Up to 12 weeks' follow-up								
Bourke 2011	120	11	8	106	13	9	100.0%	14.00 [2.59 , 25.41]
Subtotal (95% CI)			**8**			**9**	**100.0%**	**14.00 [2.59 , 25.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.40 (P = 0.02)								
1.8.2 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003b	107.4	16.5	62	109.8	18.8	31	100.0%	-2.40 [-10.19 , 5.39]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-2.40 [-10.19 , 5.39]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.60 (P = 0.55)								
Test for subgroup differences: Chi² = 5.41, df = 1 (P = 0.02), I² = 81.5%								



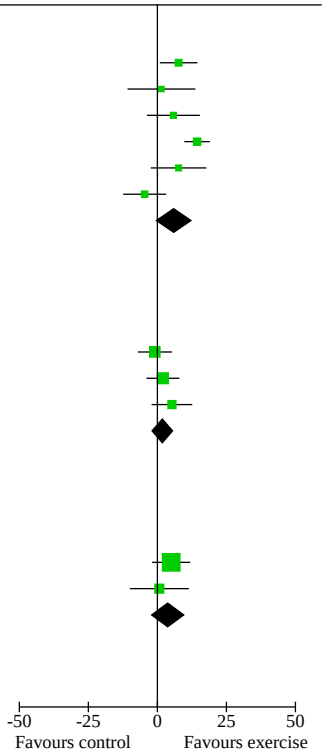
Analysis 1.9. Comparison 1: Health-related quality of life, Outcome 9: FACT-G change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.9.1 Up to 12 weeks' follow-up								
McNeely 2008a	4.4	10.6	27	1.7	6.9	25	33.6%	2.70 [-2.13 , 7.53]
Oh 2010	8.86	9	54	-0.13	8.6	54	40.2%	8.99 [5.67 , 12.31]
Rogers 2009	4.5	8.4	20	2.9	12	18	26.2%	1.60 [-5.05 , 8.25]
Subtotal (95% CI)			**101**			**97**	**100.0%**	**4.94 [-0.08 , 9.95]**
Heterogeneity: Tau² = 13.44; Chi² = 6.57, df = 2 (P = 0.04); I² = 70%								
Test for overall effect: Z = 1.93 (P = 0.05)								
1.9.2 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003b	0.9	12.2	62	2	9.7	31	48.4%	-1.10 [-5.67 , 3.47]
Courneya 2003c	5.7	7.4	24	0.6	7.4	28	51.6%	5.10 [1.07 , 9.13]
Subtotal (95% CI)			**86**			**59**	**100.0%**	**2.10 [-3.98 , 8.17]**
Heterogeneity: Tau² = 14.38; Chi² = 3.97, df = 1 (P = 0.05); I² = 75%								
Test for overall effect: Z = 0.68 (P = 0.50)								
1.9.3 6 months' follow-up								
Cadmus 2009	0.6	7.5	37	-2.4	9.8	37	76.3%	3.00 [-0.98 , 6.98]
Rogers 2009	3.4	11	19	-3.5	10.8	17	23.7%	6.90 [-0.23 , 14.03]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**3.93 [0.45 , 7.40]**
Heterogeneity: Tau² = 0.00; Chi² = 0.88, df = 1 (P = 0.35); I² = 0%								
Test for overall effect: Z = 2.22 (P = 0.03)								
Test for subgroup differences: Chi² = 0.50, df = 2 (P = 0.78), I² = 0%								



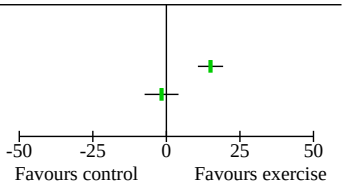
Analysis 1.10. Comparison 1: Health-related quality of life, Outcome 10: FACT-G follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.10.1 Up to 12 weeks' follow-up								
Daley 2007a	90.85	13.47	33	83.15	14.72	33	18.7%	7.70 [0.89 , 14.51]
Donnelly 2011	80.19	16.93	16	78.72	19.1	17	12.5%	1.47 [-10.83 , 13.77]
McNeely 2008a	83.9	15.6	27	78.1	19.3	25	15.4%	5.80 [-3.78 , 15.38]
Milne 2008a	80.7	7.1	29	66.3	10.6	29	21.2%	14.40 [9.76 , 19.04]
Moadel 2007	78.07	17.17	45	70.38	22.75	26	14.8%	7.69 [-2.39 , 17.77]
Rogers 2009	87.4	13.1	20	92	11.4	18	17.5%	-4.60 [-12.39 , 3.19]
Subtotal (95% CI)			**170**			**148**	**100.0%**	**5.90 [-0.36 , 12.16]**
Heterogeneity: Tau² = 42.46; Chi² = 18.64, df = 5 (P = 0.002); I² = 73%								
Test for overall effect: Z = 1.85 (P = 0.06)								
1.10.2 More than 12 weeks' to less than 6 months' follow-up								
Courneya 2003b	85.6	13	62	86.5	15	31	36.0%	-0.90 [-7.09 , 5.29]
Courneya 2003c	91.3	11	24	89.3	10.9	28	38.7%	2.00 [-3.97 , 7.97]
Daley 2007a	90.25	14.94	31	84.98	14.69	31	25.4%	5.27 [-2.11 , 12.65]
Subtotal (95% CI)			**117**			**90**	**100.0%**	**1.79 [-1.93 , 5.50]**
Heterogeneity: Tau² = 0.00; Chi² = 1.58, df = 2 (P = 0.45); I² = 0%								
Test for overall effect: Z = 0.94 (P = 0.35)								
1.10.3 6 months' follow-up								
Cadmus 2009	91.2	12.6	37	86.2	17.4	37	70.4%	5.00 [-1.92 , 11.92]
Rogers 2009	86.1	13.6	19	85.4	18.4	17	29.6%	0.70 [-9.97 , 11.37]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**3.73 [-2.08 , 9.53]**
Heterogeneity: Tau² = 0.00; Chi² = 0.44, df = 1 (P = 0.51); I² = 0%								
Test for overall effect: Z = 1.26 (P = 0.21)								
Test for subgroup differences: Chi² = 1.29, df = 2 (P = 0.52), I² = 0%								



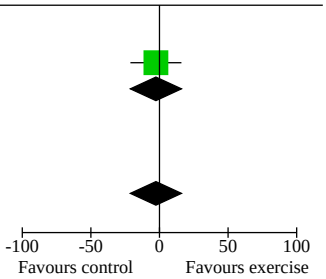
Analysis 1.11. Comparison 1: Health-related quality of life, Outcome 11: FACIT change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Mean Difference IV, Random, 95% CI
1.11.1 Up to 12 weeks' follow-up							
Mustian 2004	15	5	11	0	5	10	15.00 [10.72 , 19.28]
Targ 2002	4.98	16.1	79	6.62	21.7	88	-1.64 [-7.40 , 4.12]



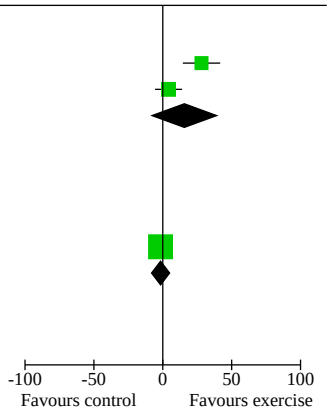
Analysis 1.12. Comparison 1: Health-related quality of life, Outcome 12: FACIT follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.12.1 Up to 12 weeks' follow-up								
Mustian 2004	121.7	20.8	11	124.3	22.5	10	100.0%	-2.60 [-21.19 , 15.99]
Subtotal (95% CI)			**11**			**10**	**100.0%**	**-2.60 [-21.19 , 15.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.27 (P = 0.78)								
Total (95% CI)			**11**			**10**	**100.0%**	**-2.60 [-21.19 , 15.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.27 (P = 0.78)								
Test for subgroup differences: Not applicable								



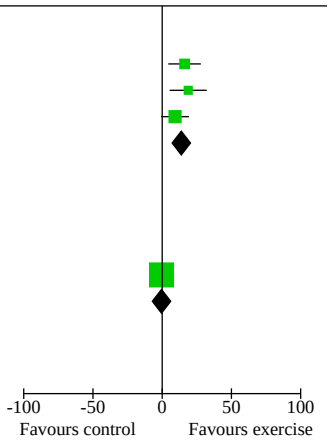
Analysis 1.13. Comparison 1: Health-related quality of life, Outcome 13: QLQ-C30 change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.13.1 Up to 12 weeks' follow-up								
Herrero 2006	17.7	8.3	8	-10.4	17.7	8	48.0%	28.10 [14.55 , 41.65]
Mehnert 2011	9.72	21.03	30	5.55	16.5	27	52.0%	4.17 [-5.60 , 13.94]
Subtotal (95% CI)			**38**			**35**	**100.0%**	**15.66 [-7.78 , 39.09]**
Heterogeneity: Tau² = 250.02; Chi² = 7.89, df = 1 (P = 0.005); I² = 87%								
Test for overall effect: Z = 1.31 (P = 0.19)								
1.13.2 More than 12 weeks' less than 6 months' follow-up								
Thorsen 2005	8.8	16.9	52	10.4	16.7	49	100.0%	-1.60 [-8.15 , 4.95]
Subtotal (95% CI)			**52**			**49**	**100.0%**	**-1.60 [-8.15 , 4.95]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.48 (P = 0.63)								
Test for subgroup differences: Chi² = 1.93, df = 1 (P = 0.16), I² = 48.2%								



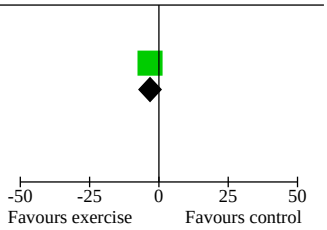
Analysis 1.14. Comparison 1: Health-related quality of life, Outcome 14: QLQ-C30 follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Fixed, 95% CI
1.14.1 Up to 12 weeks' follow-up								
Culos-Reed 2006	78.2	20.24	18	62.05	15.46	18	31.8%	16.15 [4.38 , 27.92]
Herrero 2006	81.3	10.7	8	62.5	16.1	8	24.5%	18.80 [5.40 , 32.20]
Mehnert 2011	74.72	21.72	30	65.43	16.78	27	43.8%	9.29 [-0.73 , 19.31]
Subtotal (95% CI)			**56**			**53**	**100.0%**	**13.80 [7.17 , 20.43]**
Heterogeneity: Chi² = 1.47, df = 2 (P = 0.48); I² = 0%								
Test for overall effect: Z = 4.08 (P < 0.0001)								
1.14.2 More than 12 weeks' less than 6 months' follow-up								
Thorsen 2005	74.2	17.3	52	74.7	16.7	49	100.0%	-0.50 [-7.13 , 6.13]
Subtotal (95% CI)			**52**			**49**	**100.0%**	**-0.50 [-7.13 , 6.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.15 (P = 0.88)								
Test for subgroup differences: Chi² = 8.93, df = 1 (P = 0.003), I² = 88.8%								



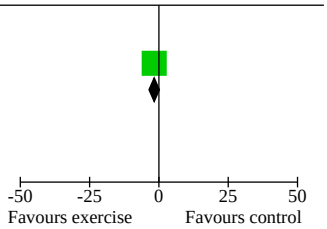
Analysis 1.15. Comparison 1: Health-related quality of life, Outcome 15: CARES follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.15.1 6 months' follow-up								
Ohira 2006	44.2	8.7	39	47.4	9.4	40	100.0%	-3.20 [-7.19 , 0.79]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-3.20 [-7.19 , 0.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.57 (P = 0.12)								
Test for subgroup differences: Not applicable								



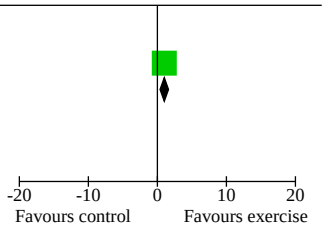
Analysis 1.16. Comparison 1: Health-related quality of life, Outcome 16: CARES change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.16.1 6 months' follow-up								
Ohira 2006	-2.3	4.5	39	-0.6	4	40	100.0%	-1.70 [-3.58 , 0.18]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-1.70 [-3.58 , 0.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.77 (P = 0.08)								
Test for subgroup differences: Not applicable								



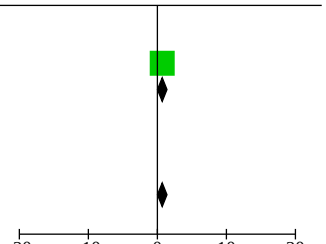
Analysis 1.17. Comparison 1: Health-related quality of life, Outcome 17: Chae and Cho change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.17.1 Up to 12 weeks' follow-up								
Cho 2006	0.9	1.3	28	-0.1	1	27	100.0%	1.00 [0.39 , 1.61]
Subtotal (95% CI)			**28**			**27**	**100.0%**	**1.00 [0.39 , 1.61]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.20 (P = 0.001)								
Test for subgroup differences: Not applicable								



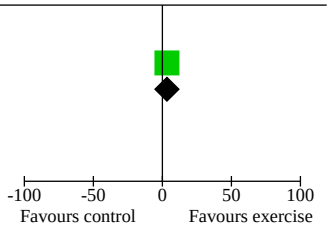
Analysis 1.18. Comparison 1: Health-related quality of life, Outcome 18: Chae and Cho follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.18.1 Up to 12 weeks' follow-up								
Cho 2006	7	1.3	28	6.3	1.3	27	100.0%	0.70 [0.01 , 1.39]
Subtotal (95% CI)			**28**			**27**	**100.0%**	**0.70 [0.01 , 1.39]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.00 (P = 0.05)								
Total (95% CI)			**28**			**27**	**100.0%**	**0.70 [0.01 , 1.39]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.00 (P = 0.05)								
Test for subgroup differences: Not applicable								



Analysis 1.19. Comparison 1: Health-related quality of life, Outcome 19: QoL Index follow-up values

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
1.19.1 Up to 12 weeks' follow-up								
Burnham 2002	89.8	12.9	14	86.5	7.3	7	100.0%	3.30 [-5.35 , 11.95]
Subtotal (95% CI)			**14**			**7**	**100.0%**	**3.30 [-5.35 , 11.95]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								
Test for subgroup differences: Not applicable								

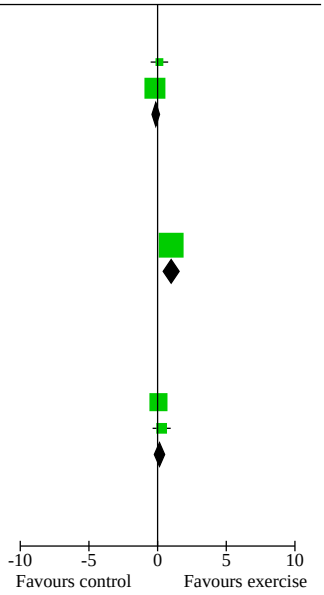


Comparison 2. Condition-specific quality of life

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
2.1 Breast cancer change	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1.1 Up to 12 weeks' follow-up	2	205	Std. Mean Difference (IV, Random, 95% CI)	-0.13 [-0.41, 0.14]
2.1.2 More than 12 weeks' up to 6 months' follow-up	1	52	Std. Mean Difference (IV, Random, 95% CI)	0.99 [0.41, 1.57]
2.1.3 6 months' follow-up	2	110	Std. Mean Difference (IV, Random, 95% CI)	0.14 [-0.24, 0.51]
2.2 FACT-B (breast) follow-up values	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.2.1 Up to 12 weeks' follow-up	4	175	Mean Difference (IV, Random, 95% CI)	2.20 [0.69, 3.72]
2.2.2 More than 12 weeks' up to 6 months' follow-up	2	113	Mean Difference (IV, Random, 95% CI)	1.81 [-0.35, 3.98]
2.2.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	2.05 [-0.20, 4.30]
2.3 FACT lymphoma follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.3.1 Up to 12 weeks' follow-up	1	122	Mean Difference (IV, Random, 95% CI)	1.00 [-1.74, 3.74]
2.4 FACT lymphoma subscale change	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.4.1 Up to 12 weeks' follow-up	1	122	Mean Difference (IV, Fixed, 95% CI)	1.20 [-1.02, 3.42]
2.5 FACT-C (colorectal) change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.5.1 More than 12 weeks' less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-0.20 [-2.10, 1.70]
2.6 FACT-C (colorectal) follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.6.1 More than 12 weeks' less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-1.40 [-3.33, 0.53]
2.7 Neck Dissection Impairment Index follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.7.1 Up to 12 weeks' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	8.40 [-3.54, 20.34]

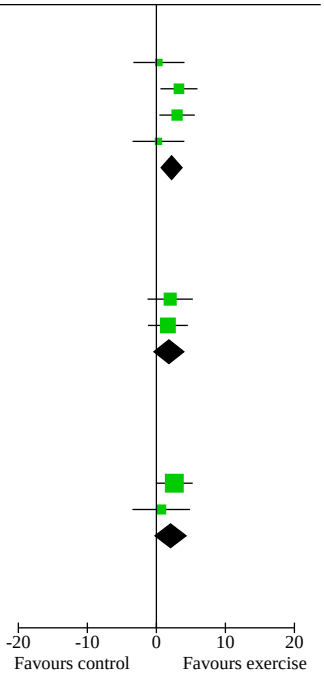
Analysis 2.1. Comparison 2: Condition-specific quality of life, Outcome 1: Breast cancer change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.1.1 Up to 12 weeks' follow-up								
Rogers 2009	1	5.3	20	0.3	5	18	18.6%	0.13 [-0.50 , 0.77]
Targ 2002	-0.29	5.8	79	0.62	3.3	88	81.4%	-0.19 [-0.50 , 0.11]
Subtotal (95% CI)			**99**			**106**	**100.0%**	**-0.13 [-0.41 , 0.14]**
Heterogeneity: Tau² = 0.00; Chi² = 0.83, df = 1 (P = 0.36); I² = 0%								
Test for overall effect: Z = 0.95 (P = 0.34)								
2.1.2 More than 12 weeks' up to 6 months' follow-up								
Courneya 2003c	3.3	4	24	-0.3	3.2	28	100.0%	0.99 [0.41 , 1.57]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**0.99 [0.41 , 1.57]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.34 (P = 0.0008)								
2.1.3 6 months' follow-up								
Cadmus 2009	1	3	37	0.8	3.4	37	67.6%	0.06 [-0.39 , 0.52]
Rogers 2009	0.7	3.7	19	-0.6	4.9	17	32.4%	0.30 [-0.36 , 0.95]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.14 [-0.24 , 0.51]**
Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.57); I² = 0%								
Test for overall effect: Z = 0.72 (P = 0.47)								
Test for subgroup differences: Chi² = 11.82, df = 2 (P = 0.003), I² = 83.1%								



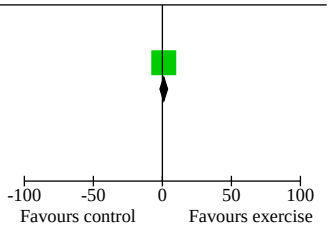
Analysis 2.2. Comparison 2: Condition-specific quality of life, Outcome 2: FACT-B (breast) follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.2.1 Up to 12 weeks' follow-up								
Banasik 2011	8.8	3.11	7	8.43	3.91	7	16.8%	0.37 [-3.33 , 4.07]
Daley 2007a	24.4	5.73	32	21.13	5.21	32	31.9%	3.27 [0.59 , 5.95]
Milne 2008a	21.6	4	29	18.6	5.8	29	35.0%	3.00 [0.44 , 5.56]
Rogers 2009	27	5.5	20	26.7	6.4	19	16.3%	0.30 [-3.45 , 4.05]
Subtotal (95% CI)			**88**			**87**	**100.0%**	**2.20 [0.69 , 3.72]**
Heterogeneity: Tau² = 0.00; Chi² = 2.91, df = 3 (P = 0.41); I² = 0%								
Test for overall effect: Z = 2.85 (P = 0.004)								
2.2.2 More than 12 weeks' up to 6 months' follow-up								
Courneya 2003c	28.4	6.7	24	26.4	5.1	28	43.5%	2.00 [-1.28 , 5.28]
Daley 2007a	25.3	5.88	30	23.63	5.58	31	56.5%	1.67 [-1.21 , 4.55]
Subtotal (95% CI)			**54**			**59**	**100.0%**	**1.81 [-0.35 , 3.98]**
Heterogeneity: Tau² = 0.00; Chi² = 0.02, df = 1 (P = 0.88); I² = 0%								
Test for overall effect: Z = 1.64 (P = 0.10)								
2.2.3 6 months' follow-up								
Cadmus 2009	26.7	5.8	37	24.1	5.9	37	71.0%	2.60 [-0.07 , 5.27]
Rogers 2009	26	6	19	25.3	6.7	17	29.0%	0.70 [-3.47 , 4.87]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**2.05 [-0.20 , 4.30]**
Heterogeneity: Tau² = 0.00; Chi² = 0.57, df = 1 (P = 0.45); I² = 0%								
Test for overall effect: Z = 1.79 (P = 0.07)								
Test for subgroup differences: Chi² = 0.08, df = 2 (P = 0.96), I² = 0%								



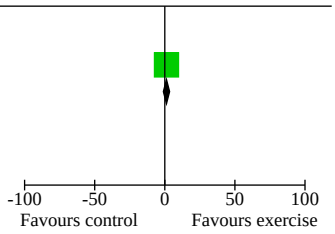
Analysis 2.3. Comparison 2: Condition-specific quality of life, Outcome 3: FACT lymphoma follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.3.1 Up to 12 weeks' follow-up								
Courneya 2009	50.3	7.2	60	49.3	8.2	62	100.0%	1.00 [-1.74 , 3.74]
Subtotal (95% CI)			**60**			**62**	**100.0%**	**1.00 [-1.74 , 3.74]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.72 (P = 0.47)								
Test for subgroup differences: Not applicable								



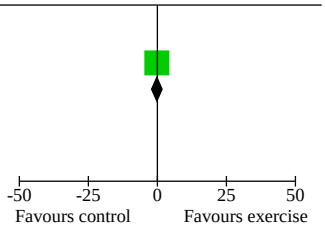
Analysis 2.4. Comparison 2: Condition-specific quality of life, Outcome 4: FACT lymphoma subscale change

Study or Subgroup	Exercise Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.4.1 Up to 12 weeks' follow-up								
Courneya 2009	2.7	6.2	60	1.5	6.3	62	100.0%	1.20 [-1.02 , 3.42]
Subtotal (95% CI)			**60**			**62**	**100.0%**	**1.20 [-1.02 , 3.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.06 (P = 0.29)								
Test for subgroup differences: Not applicable								



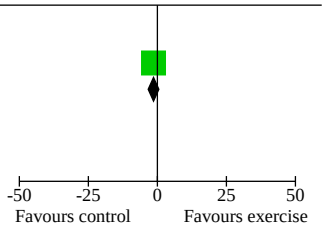
Analysis 2.5. Comparison 2: Condition-specific quality of life, Outcome 5: FACT-C (colorectal) change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.5.1 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	0.5	4.6	62	0.7	4.3	31	100.0%	-0.20 [-2.10 , 1.70]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-0.20 [-2.10 , 1.70]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.21 (P = 0.84)								
Test for subgroup differences: Not applicable								

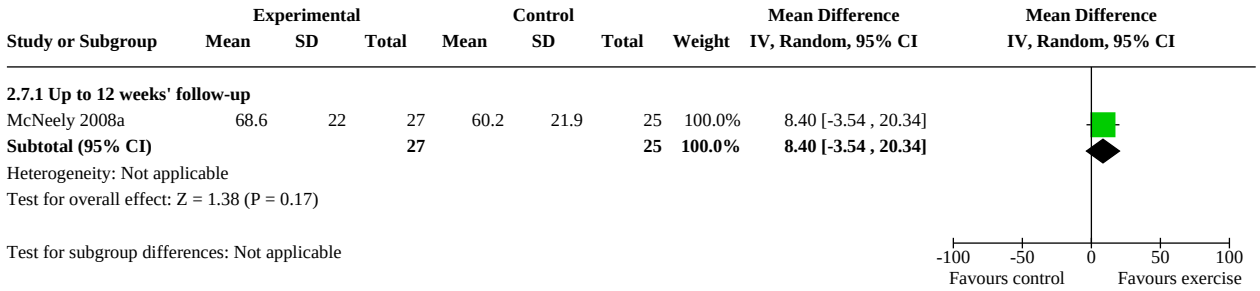


Analysis 2.6. Comparison 2: Condition-specific quality of life, Outcome 6: FACT-C (colorectal) follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.6.1 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	21.8	4.6	62	23.2	4.4	31	100.0%	-1.40 [-3.33 , 0.53]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-1.40 [-3.33 , 0.53]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.42 (P = 0.15)								
Test for subgroup differences: Not applicable								



Analysis 2.7. Comparison 2: Condition-specific quality of life, Outcome 7: Neck Dissection Impairment Index follow-up values



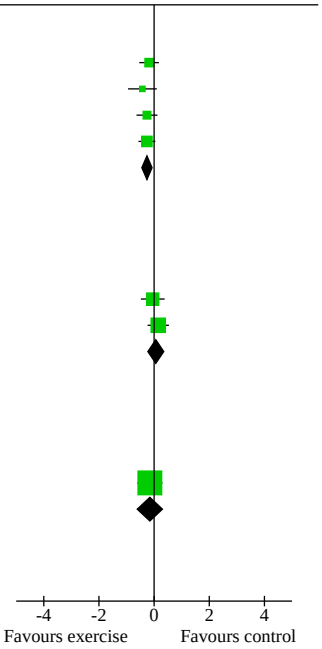
Comparison 3. Anxiety

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Overall anxiety change	7		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1.1 Up to 12 weeks' follow-up	4	455	Std. Mean Difference (IV, Random, 95% CI)	-0.26 [-0.44, -0.07]
3.1.2 More than 12 weeks' less than 6 months' follow-up	2	196	Std. Mean Difference (IV, Random, 95% CI)	0.06 [-0.23, 0.35]
3.1.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.61, 0.30]
3.2 Overall anxiety follow-up scores	11		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
3.2.1 Up to 12 weeks' follow-up	8	396	Std. Mean Difference (IV, Random, 95% CI)	-0.40 [-0.77, -0.03]
3.2.2 More than 12 weeks' less than 6 months' follow-up	2	196	Std. Mean Difference (IV, Random, 95% CI)	0.13 [-0.15, 0.42]
3.2.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.60, 0.31]
3.3 Hospital Anxiety and Depression; anxiety subscale change score	2	159	Mean Difference (IV, Random, 95% CI)	-0.50 [-2.66, 1.66]
3.3.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-1.85 [-4.09, 0.39]
3.3.2 More than 12 weeks' up to 6 months' follow-up	1	101	Mean Difference (IV, Random, 95% CI)	0.40 [-0.63, 1.43]
3.4 Hospital Anxiety and Depression; anxiety subscale follow-up scores	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.4.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-2.31 [-4.53, -0.09]
3.4.2 More than 12 weeks' follow-up	1	103	Mean Difference (IV, Random, 95% CI)	0.90 [-0.53, 2.33]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.5 State Trait Anxiety Inventory change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.5.1 Up to 12 weeks' follow-up	1	122	Mean Difference (IV, Random, 95% CI)	-1.00 [-2.93, 0.93]
3.5.2 More than 12 weeks' up to 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-0.60 [-5.57, 4.37]
3.5.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-1.20 [-4.70, 2.30]
3.6 State Trait Anxiety Inventory follow-up scores	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.6.1 Up to 12 weeks' follow-up	4	263	Mean Difference (IV, Random, 95% CI)	-2.40 [-6.90, 2.10]
3.6.2 More than 12 weeks' less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	0.00 [-5.28, 5.28]
3.6.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-2.00 [-8.35, 4.35]
3.7 POMS tension anxiety subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.7.1 Up to 12 weeks' follow-up	2	275	Mean Difference (IV, Random, 95% CI)	1.55 [0.08, 3.02]
3.8 POMS tension anxiety subscale follow-up scores	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.8.1 Up to 12 weeks' follow-up	3	125	Mean Difference (IV, Fixed, 95% CI)	-3.20 [-5.40, -1.00]
3.9 Linear Analog Self-Assessment Scale follow-up scores	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.9.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	-13.80 [-33.36, 5.76]

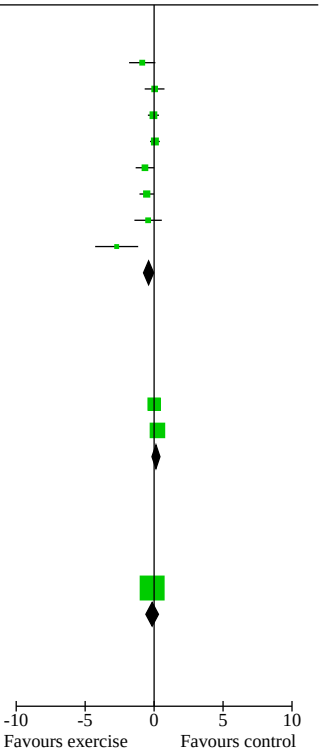
Analysis 3.1. Comparison 3: Anxiety, Outcome 1: Overall anxiety change

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
3.1.1 Up to 12 weeks' follow-up								
Courneya 2009	-1.6	5.4	60	-0.6	5.5	62	27.0%	-0.18 [-0.54 , 0.17]
Mehnert 2011	-1.67	3.68	30	0.18	4.89	28	12.6%	-0.42 [-0.95 , 0.10]
Oh 2010	-1.71	4.5	54	-0.47	5	54	23.8%	-0.26 [-0.64 , 0.12]
Targ 2002	-3.98	9.2	79	-1.82	7.3	88	36.7%	-0.26 [-0.57 , 0.04]
Subtotal (95% CI)			**223**			**232**	**100.0%**	**-0.26 [-0.44 , -0.07]**
Heterogeneity: Tau² = 0.00; Chi² = 0.56, df = 3 (P = 0.90); I² = 0%								
Test for overall effect: Z = 2.75 (P = 0.006)								
3.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-4.3	13	62	-3.7	10.7	31	44.6%	-0.05 [-0.48 , 0.38]
Thorsen 2005	-0.38	2.6	53	-0.78	2.7	50	55.4%	0.15 [-0.24 , 0.54]
Subtotal (95% CI)			**115**			**81**	**100.0%**	**0.06 [-0.23 , 0.35]**
Heterogeneity: Tau² = 0.00; Chi² = 0.45, df = 1 (P = 0.50); I² = 0%								
Test for overall effect: Z = 0.42 (P = 0.68)								
3.1.3 6 months' follow-up								
Cadmus 2009	-0.7	9.2	37	0.5	5.8	37	100.0%	-0.15 [-0.61 , 0.30]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.15 [-0.61 , 0.30]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.66 (P = 0.51)								
Test for subgroup differences: Chi² = 3.38, df = 2 (P = 0.18), I² = 40.9%								



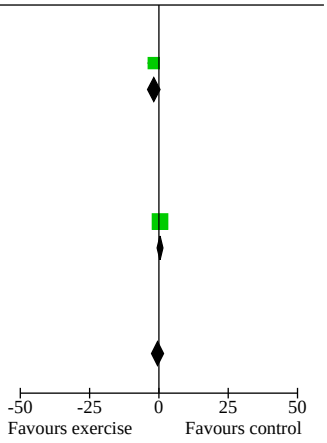
Analysis 3.2. Comparison 3: Anxiety, Outcome 2: Overall anxiety follow-up scores

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
3.2.1 Up to 12 weeks' follow-up								
Burnham 2002	5.9	5.6	14	19.7	26.1	7	9.0%	-0.86 [-1.81 , 0.09]
Cohen 2004	34.1	8.4	16	33.8	8.5	14	12.3%	0.03 [-0.68 , 0.75]
Courneya 2003a	43.81	12	51	44.34	8.98	45	18.1%	-0.05 [-0.45 , 0.35]
Courneya 2009	16.5	5.2	60	16.2	5.2	62	18.9%	0.06 [-0.30 , 0.41]
Culos-Reed 2006	1.39	3.29	18	4.56	5.72	18	13.0%	-0.66 [-1.34 , 0.01]
Mehnert 2011	4.83	3.46	30	7.14	4.97	28	15.7%	-0.54 [-1.06 , -0.01]
Pinto 2003	7.58	7.3	12	10.5	3.7	6	8.6%	-0.43 [-1.43 , 0.56]
Segar 1998	28	3	10	40	6	5	4.5%	-2.71 [-4.27 , -1.16]
Subtotal (95% CI)			**211**			**185**	**100.0%**	**-0.40 [-0.77 , -0.03]**
Heterogeneity: Tau² = 0.15; Chi² = 18.32, df = 7 (P = 0.01); I² = 62%								
Test for overall effect: Z = 2.14 (P = 0.03)								
3.2.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	35.5	11.5	62	35.5	12.6	31	44.7%	0.00 [-0.43 , 0.43]
Thorsen 2005	5.3	3.5	53	4.4	3.9	50	55.3%	0.24 [-0.15 , 0.63]
Subtotal (95% CI)			**115**			**81**	**100.0%**	**0.13 [-0.15 , 0.42]**
Heterogeneity: Tau² = 0.00; Chi² = 0.67, df = 1 (P = 0.41); I² = 0%								
Test for overall effect: Z = 0.91 (P = 0.36)								
3.2.3 6 months' follow-up								
Cadmus 2009	32.1	12.3	37	34.1	15.4	37	100.0%	-0.14 [-0.60 , 0.31]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.14 [-0.60 , 0.31]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.61 (P = 0.54)								
Test for subgroup differences: Chi² = 5.10, df = 2 (P = 0.08), I² = 60.8%								



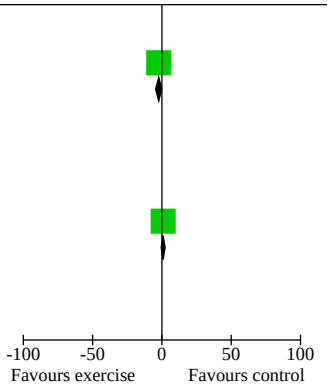
Analysis 3.3. Comparison 3: Anxiety, Outcome 3: Hospital Anxiety and Depression; anxiety subscale change score

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
3.3.1 Up to 12 weeks' follow-up								
Mehnert 2011	-1.67	3.68	30	0.18	4.89	28	39.9%	-1.85 [-4.09 , 0.39]
Subtotal (95% CI)			**30**			**28**	**39.9%**	**-1.85 [-4.09 , 0.39]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.62 (P = 0.11)								
3.3.2 More than 12 weeks' up to 6 months' follow-up								
Thorsen 2005	-0.38	2.6	52	-0.78	2.7	49	60.1%	0.40 [-0.63 , 1.43]
Subtotal (95% CI)			**52**			**49**	**60.1%**	**0.40 [-0.63 , 1.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.76 (P = 0.45)								
Total (95% CI)			**82**			**77**	**100.0%**	**-0.50 [-2.66 , 1.66]**
Heterogeneity: Tau² = 1.74; Chi² = 3.20, df = 1 (P = 0.07); I² = 69%								
Test for overall effect: Z = 0.45 (P = 0.65)								
Test for subgroup differences: Chi² = 3.20, df = 1 (P = 0.07), I² = 68.7%								



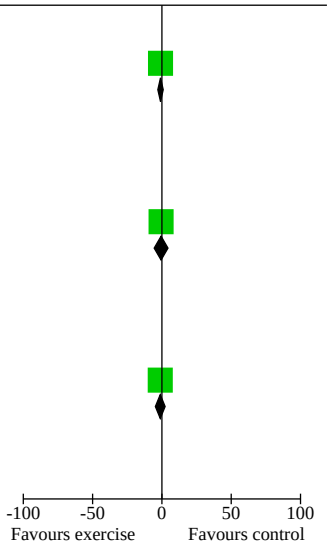
Analysis 3.4. Comparison 3: Anxiety, Outcome 4: Hospital Anxiety and Depression; anxiety subscale follow-up scores

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
3.4.1 Up to 12 weeks' follow-up								
Mehnert 2011	4.83	3.46	30	7.14	4.97	28	100.0%	-2.31 [-4.53 , -0.09]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**-2.31 [-4.53 , -0.09]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.04 (P = 0.04)								
3.4.2 More than 12 weeks' follow-up								
Thorsen 2005	5.3	3.5	53	4.4	3.9	50	100.0%	0.90 [-0.53 , 2.33]
Subtotal (95% CI)			**53**			**50**	**100.0%**	**0.90 [-0.53 , 2.33]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.23 (P = 0.22)								
Test for subgroup differences: Chi² = 5.67, df = 1 (P = 0.02), I² = 82.4%								



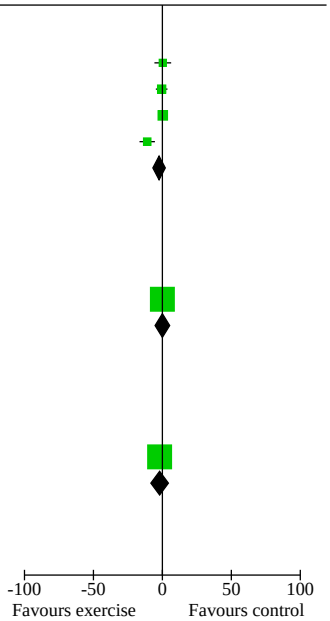
Analysis 3.5. Comparison 3: Anxiety, Outcome 5: State Trait Anxiety Inventory change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
3.5.1 Up to 12 weeks' follow-up								
Courneya 2009	-1.6	5.4	60	-0.6	5.5	62	100.0%	-1.00 [-2.93 , 0.93]
Subtotal (95% CI)			**60**			**62**	**100.0%**	**-1.00 [-2.93 , 0.93]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.01 (P = 0.31)								
3.5.2 More than 12 weeks' up to 6 months' follow-up								
Courneya 2003b	-4.3	13	62	-3.7	10.7	31	100.0%	-0.60 [-5.57 , 4.37]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-0.60 [-5.57 , 4.37]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
3.5.3 6 months' follow-up								
Cadmus 2009	-0.7	9.2	37	0.5	5.8	37	100.0%	-1.20 [-4.70 , 2.30]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-1.20 [-4.70 , 2.30]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.67 (P = 0.50)								
Test for subgroup differences: Chi^2 = 0.04, df = 2 (P = 0.98), I^2 = 0%								



Analysis 3.6. Comparison 3: Anxiety, Outcome 6: State Trait Anxiety Inventory follow-up scores

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
3.6.1 Up to 12 weeks' follow-up								
Cohen 2004	34.1	8.4	16	33.8	8.5	14	20.7%	0.30 [-5.76 , 6.36]
Courneya 2003a	43.81	12	51	44.34	8.98	45	25.8%	-0.53 [-4.74 , 3.68]
Courneya 2009	16.5	5.2	60	16.2	5.2	62	31.5%	0.30 [-1.55 , 2.15]
Segar 1998	28.5	3	10	39.5	6	5	22.0%	-11.00 [-16.58 , -5.42]
Subtotal (95% CI)			**137**			**126**	**100.0%**	**-2.40 [-6.90 , 2.10]**
Heterogeneity: Tau^2 = 15.81; Chi^2 = 14.34, df = 3 (P = 0.002); I^2 = 79%								
Test for overall effect: Z = 1.05 (P = 0.30)								
3.6.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	35.5	11.5	62	35.5	12.6	31	100.0%	0.00 [-5.28 , 5.28]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**0.00 [-5.28 , 5.28]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
3.6.3 6 months' follow-up								
Cadmus 2009	32.1	12.3	37	34.1	15.4	37	100.0%	-2.00 [-8.35 , 4.35]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-2.00 [-8.35 , 4.35]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.62 (P = 0.54)								
Test for subgroup differences: Chi^2 = 0.49, df = 2 (P = 0.78), I^2 = 0%								

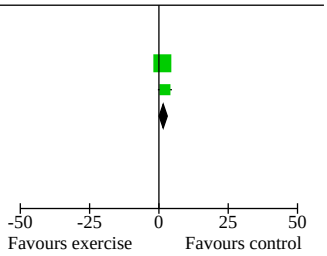


Analysis 3.7. Comparison 3: Anxiety, Outcome 7: POMS tension anxiety subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
3.7.1 Up to 12 weeks' follow-up								
Oh 2010	1.71	4.5064	54	0.47	5.0193	54	66.3%	1.24 [-0.56 , 3.04]
Targ 2002	3.98	9.1523	79	1.82	7.2683	88	33.7%	2.16 [-0.37 , 4.69]
Subtotal (95% CI)			**133**			**142**	**100.0%**	**1.55 [0.08 , 3.02]**

Heterogeneity: Tau² = 0.00; Chi² = 0.34, df = 1 (P = 0.56); I² = 0%
Test for overall effect: Z = 2.07 (P = 0.04)

Test for subgroup differences: Not applicable

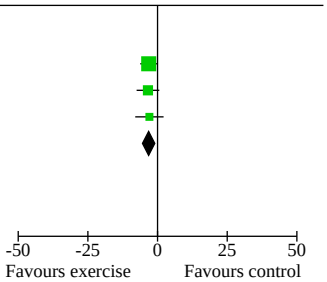


Analysis 3.8. Comparison 3: Anxiety, Outcome 8: POMS tension anxiety subscale follow-up scores

Study or Subgroup	Exercise Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.8.1 Up to 12 weeks' follow-up								
Culos-Reed 2006	1.39	3.29	18	4.56	5.72	18	52.1%	-3.17 [-6.22 , -0.12]
Moadel 2007	8.79	7.54	45	12.23	8.91	26	29.2%	-3.44 [-7.51 , 0.63]
Pinto 2003	7.58	7.3	12	10.5	3.7	6	18.7%	-2.92 [-8.00 , 2.16]
Subtotal (95% CI)			**75**			**50**	**100.0%**	**-3.20 [-5.40 , -1.00]**

Heterogeneity: Chi² = 0.03, df = 2 (P = 0.99); I² = 0%
Test for overall effect: Z = 2.85 (P = 0.004)

Test for subgroup differences: Not applicable

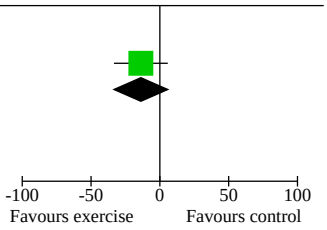


Analysis 3.9. Comparison 3: Anxiety, Outcome 9: Linear Analog Self-Assessment Scale follow-up scores

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
3.9.1 Up to 12 weeks' follow-up								
Burnham 2002	5.9	5.6	14	19.7	26.1	7	100.0%	-13.80 [-33.36 , 5.76]
Subtotal (95% CI)			**14**			**7**	**100.0%**	**-13.80 [-33.36 , 5.76]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.38 (P = 0.17)

Test for subgroup differences: Not applicable



Comparison 4. Body Image/self-esteem

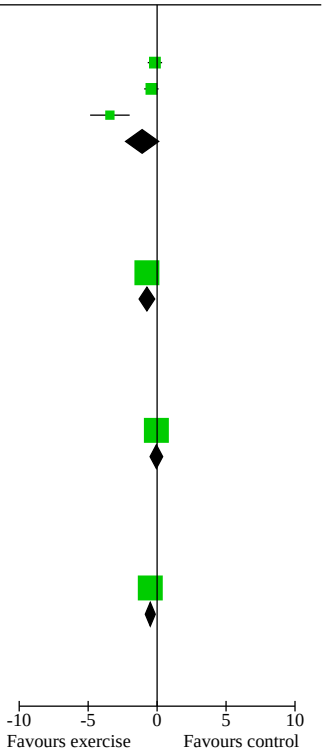
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Overall body image/self-esteem change	6		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1.1 Up to 12 weeks' follow-up	3	136	Std. Mean Difference (IV, Random, 95% CI)	-1.09 [-2.29, 0.11]
4.1.2 More than 12 weeks' less than 6 months' follow-up	1	52	Std. Mean Difference (IV, Random, 95% CI)	-0.74 [-1.30, -0.18]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	-0.05 [-0.51, 0.40]
4.1.4 More than 6 months' follow-up	1	122	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.86, -0.13]
4.2 Overall body image/self-esteem follow-up scores	9		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
4.2.1 Up to 12 weeks' follow-up	5	233	Std. Mean Difference (IV, Random, 95% CI)	-0.50 [-0.79, -0.20]
4.2.2 More than 12 weeks' less than 6 months' follow-up	2	113	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.83, 0.20]
4.2.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	-0.20 [-0.65, 0.26]
4.2.4 More than 6 months' follow-up	1	122	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.42, 0.29]
4.3 Body Esteem Scale - weight follow-up scores	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.3.1 Up to 12 weeks' follow-up	2	106	Mean Difference (IV, Random, 95% CI)	4.36 [-0.91, 9.63]
4.4 Body Image Questionnaire individual body image subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.4.1 Up to 12 weeks' follow-up	1	57	Mean Difference (IV, Random, 95% CI)	-0.34 [-1.44, 0.76]
4.5 Body Image Questionnaire individual body image subscale follow-up values	1	57	Mean Difference (IV, Random, 95% CI)	-1.35 [-2.50, -0.20]
4.5.1 Up to 12 weeks' follow-up	1	57	Mean Difference (IV, Random, 95% CI)	-1.35 [-2.50, -0.20]
4.6 Body Image and Relationship Scale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.6.1 More than 6 months' follow-up	1	122	Mean Difference (IV, Random, 95% CI)	7.90 [2.27, 13.53]
4.7 Body Image and Relationship Scale follow-up values	1	122	Mean Difference (IV, Random, 95% CI)	-1.10 [-7.19, 4.99]
4.7.1 More than 6 months' follow-up	1	122	Mean Difference (IV, Random, 95% CI)	-1.10 [-7.19, 4.99]
4.8 Social Physique Anxiety Scale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.8.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-1.60 [-3.56, 0.36]
4.9 Social Physique Anxiety Scale follow-up values	1	58	Mean Difference (IV, Random, 95% CI)	-2.20 [-5.57, 1.17]
4.9.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-2.20 [-5.57, 1.17]
4.10 Rosenberg Self-Esteem Scale change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.10.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	4.50 [3.40, 5.60]
4.10.2 More than 12 weeks' less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	2.70 [0.73, 4.67]
4.10.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	0.20 [-1.50, 1.90]
4.11 Rosenberg Self-Esteem Scale follow-up values	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.11.1 Up to 12 weeks' follow-up	1	15	Mean Difference (IV, Random, 95% CI)	2.50 [-0.12, 5.12]
4.11.2 More than 12 weeks' less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	0.20 [-2.29, 2.69]
4.11.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	1.10 [-1.43, 3.63]
4.12 Physical Self-Perception Profile attractiveness of body subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.12.1 Up to 12 weeks' follow-up	1	66	Mean Difference (IV, Random, 95% CI)	0.23 [-0.10, 0.56]
4.12.2 More than 12 weeks' less than 6 months' follow-up	1	62	Mean Difference (IV, Random, 95% CI)	0.26 [-19.10, 19.62]
4.13 Physical Self-Perception Profile physical self-worth subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.13.1 Up to 12 weeks' follow-up	1	65	Mean Difference (IV, Random, 95% CI)	0.46 [0.13, 0.79]
4.13.2 More than 12 weeks' less than 6 months' follow-up	1	61	Mean Difference (IV, Random, 95% CI)	0.29 [0.04, 0.54]

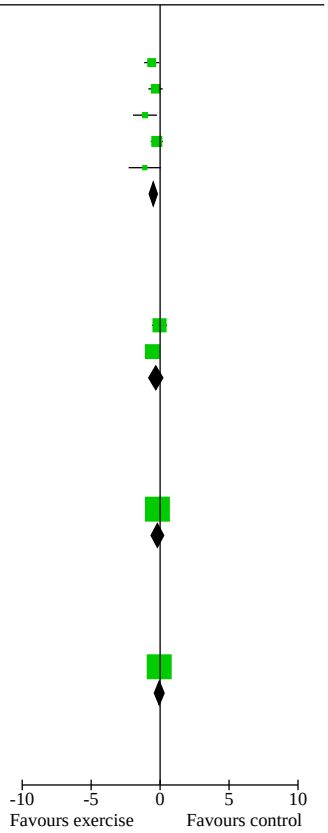
Analysis 4.1. Comparison 4: Body Image/self-esteem, Outcome 1: Overall body image/self-esteem change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.1.1 Up to 12 weeks' follow-up								
Mehnert 2011	-0.73	2.12	30	-0.39	2.12	27	37.2%	-0.16 [-0.68 , 0.36]
Milne 2008a	-1.5	3.6805	29	0.1	3.9434	29	37.2%	-0.41 [-0.93 , 0.11]
Mustian 2004	-3	1	11	1.5	1.5	10	25.5%	-3.42 [-4.86 , -1.99]
Subtotal (95% CI)			**70**			**66**	**100.0%**	**-1.09 [-2.29 , 0.11]**
Heterogeneity: Tau² = 0.94; Chi² = 17.74, df = 2 (P = 0.0001); I² = 89%								
Test for overall effect: Z = 1.77 (P = 0.08)								
4.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	-2.6	3.7	24	0.1	3.5	28	100.0%	-0.74 [-1.30 , -0.18]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**-0.74 [-1.30 , -0.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.57 (P = 0.01)								
4.1.3 6 months' follow-up								
Cadmus 2009	-0.4	4.4	37	-0.2	2.9	37	100.0%	-0.05 [-0.51 , 0.40]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.05 [-0.51 , 0.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.23 (P = 0.82)								
4.1.4 More than 6 months' follow-up								
Speck 2010	-12	15.5	59	-4.1	16.2	63	100.0%	-0.49 [-0.86 , -0.13]
Subtotal (95% CI)			**59**			**63**	**100.0%**	**-0.49 [-0.86 , -0.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.69 (P = 0.007)								
Test for subgroup differences: Chi² = 5.05, df = 3 (P = 0.17), I² = 40.6%								



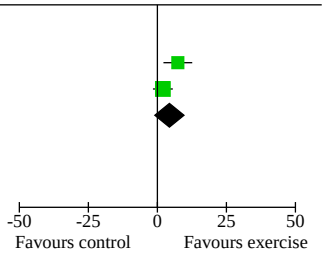
Analysis 4.2. Comparison 4: Body Image/self-esteem, Outcome 2: Overall body image/self-esteem follow-up scores

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.2.1 Up to 12 weeks' follow-up								
Mehnert 2011	5.47	1.66	27	6.82	2.6	27	23.7%	-0.61 [-1.16 , -0.06]
Milne 2008a	17.9	6.8	29	20.1	6.3	29	25.8%	-0.33 [-0.85 , 0.19]
Pinto 2003	-30.91	8.2	12	-23.5	4.2	12	10.6%	-1.10 [-1.97 , -0.23]
Pinto 2005	-27.02	9.02	39	-25.03	7.42	43	33.7%	-0.24 [-0.67 , 0.20]
Segar 1998	-33	1.8	10	-30.5	2.7	5	6.1%	-1.11 [-2.28 , 0.06]
Subtotal (95% CI)			**117**			**116**	**100.0%**	**-0.50 [-0.79 , -0.20]**
Heterogeneity: Tau² = 0.02; Chi² = 4.76, df = 4 (P = 0.31); I² = 16%								
Test for overall effect: Z = 3.28 (P = 0.001)								
4.2.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	-34.8	4.7	24	-34.6	4.4	28	48.4%	-0.04 [-0.59 , 0.50]
Daley 2007a	-2.17	0.53	30	-1.88	0.48	31	51.6%	-0.57 [-1.08 , -0.05]
Subtotal (95% CI)			**54**			**59**	**100.0%**	**-0.31 [-0.83 , 0.20]**
Heterogeneity: Tau² = 0.06; Chi² = 1.88, df = 1 (P = 0.17); I² = 47%								
Test for overall effect: Z = 1.20 (P = 0.23)								
4.2.3 6 months' follow-up								
Cadmus 2009	-34.5	5.2	37	-33.4	5.9	37	100.0%	-0.20 [-0.65 , 0.26]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.20 [-0.65 , 0.26]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.84 (P = 0.40)								
4.2.4 More than 6 months' follow-up								
Speck 2010	70.4	16.3	59	71.5	18	63	100.0%	-0.06 [-0.42 , 0.29]
Subtotal (95% CI)			**59**			**63**	**100.0%**	**-0.06 [-0.42 , 0.29]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.73)								
Test for subgroup differences: Chi² = 3.58, df = 3 (P = 0.31), I² = 16.1%								



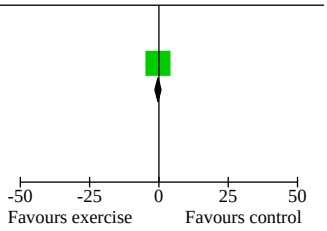
Analysis 4.3. Comparison 4: Body Image/self-esteem, Outcome 3: Body Esteem Scale - weight follow-up scores

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.3.1 Up to 12 weeks' follow-up								
Pinto 2003	30.91	8.2	12	23.5	4.2	12	43.7%	7.41 [2.20 , 12.62]
Pinto 2005	27.02	9.02	39	25.03	7.42	43	56.3%	1.99 [-1.61 , 5.59]
Subtotal (95% CI)			**51**			**55**	**100.0%**	**4.36 [-0.91 , 9.63]**
Heterogeneity: Tau² = 9.47; Chi² = 2.81, df = 1 (P = 0.09); I² = 64%								
Test for overall effect: Z = 1.62 (P = 0.10)								
Test for subgroup differences: Not applicable								



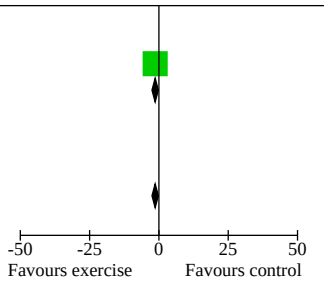
Analysis 4.4. Comparison 4: Body Image/self-esteem, Outcome 4: Body Image Questionnaire individual body image subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.4.1 Up to 12 weeks' follow-up								
Mehnert 2011	-0.73	2.12	30	-0.39	2.12	27	100.0%	-0.34 [-1.44 , 0.76]
Subtotal (95% CI)			**30**			**27**	**100.0%**	**-0.34 [-1.44 , 0.76]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.60 (P = 0.55)								
Test for subgroup differences: Not applicable								



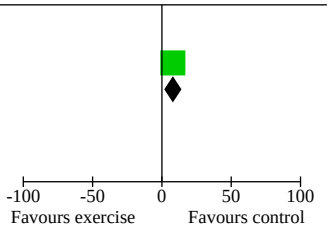
Analysis 4.5. Comparison 4: Body Image/self-esteem, Outcome 5: Body Image Questionnaire individual body image subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.5.1 Up to 12 weeks' follow-up								
Mehnert 2011	5.47	1.66	30	6.82	2.6	27	100.0%	-1.35 [-2.50 , -0.20]
Subtotal (95% CI)			**30**			**27**	**100.0%**	**-1.35 [-2.50 , -0.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.31 (P = 0.02)								
Total (95% CI)			**30**			**27**	**100.0%**	**-1.35 [-2.50 , -0.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.31 (P = 0.02)								
Test for subgroup differences: Not applicable								



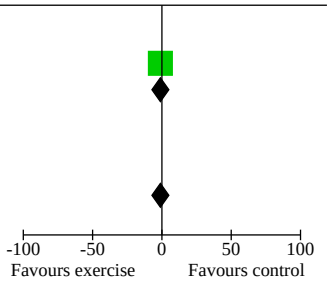
Analysis 4.6. Comparison 4: Body Image/self-esteem, Outcome 6: Body Image and Relationship Scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.6.1 More than 6 months' follow-up								
Speck 2010	12	15.5	59	4.1	16.2	63	100.0%	7.90 [2.27 , 13.53]
Subtotal (95% CI)			**59**			**63**	**100.0%**	**7.90 [2.27 , 13.53]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.75 (P = 0.006)								
Test for subgroup differences: Not applicable								



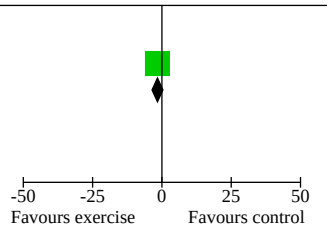
Analysis 4.7. Comparison 4: Body Image/self-esteem, Outcome 7: Body Image and Relationship Scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.7.1 More than 6 months' follow-up								
Speck 2010	70.4	16.3	59	71.5	18	63	100.0%	-1.10 [-7.19 , 4.99]
Subtotal (95% CI)			**59**			**63**	**100.0%**	**-1.10 [-7.19 , 4.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.72)								
Total (95% CI)			**59**			**63**	**100.0%**	**-1.10 [-7.19 , 4.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.72)								
Test for subgroup differences: Not applicable								



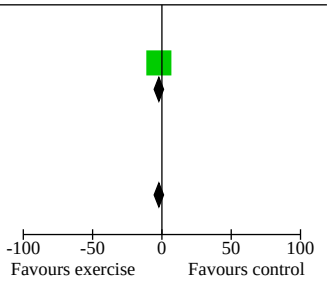
Analysis 4.8. Comparison 4: Body Image/self-esteem, Outcome 8: Social Physique Anxiety Scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.8.1 Up to 12 weeks' follow-up								
Milne 2008a	-1.5	3.7	29	0.1	3.9	29	100.0%	-1.60 [-3.56 , 0.36]
Subtotal (95% CI)			**29**			**29**	**100.0%**	**-1.60 [-3.56 , 0.36]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.60 (P = 0.11)								
Test for subgroup differences: Not applicable								



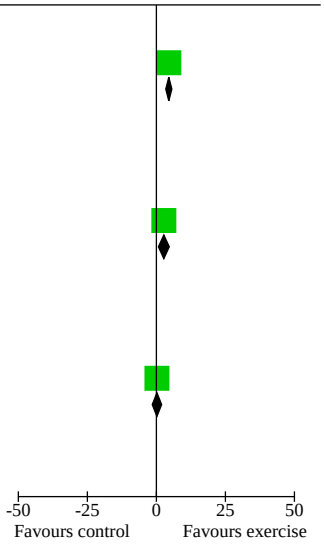
Analysis 4.9. Comparison 4: Body Image/self-esteem, Outcome 9: Social Physique Anxiety Scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.9.1 Up to 12 weeks' follow-up								
Milne 2008a	17.9	6.8	29	20.1	6.3	29	100.0%	-2.20 [-5.57 , 1.17]
Subtotal (95% CI)			**29**			**29**	**100.0%**	**-2.20 [-5.57 , 1.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.28 (P = 0.20)								
Total (95% CI)			**29**			**29**	**100.0%**	**-2.20 [-5.57 , 1.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.28 (P = 0.20)								
Test for subgroup differences: Not applicable								



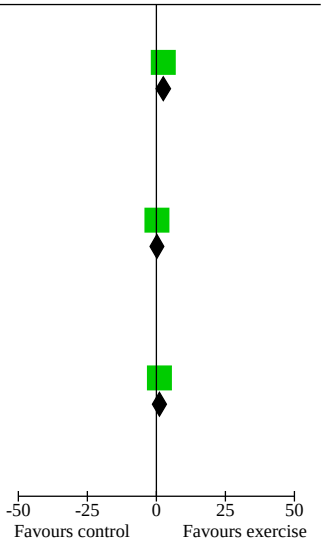
Analysis 4.10. Comparison 4: Body Image/self-esteem, Outcome 10: Rosenberg Self-Esteem Scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.10.1 Up to 12 weeks' follow-up								
Mustian 2004	3	1	11	-1.5	1.5	10	100.0%	4.50 [3.40 , 5.60]
Subtotal (95% CI)			**11**			**10**	**100.0%**	**4.50 [3.40 , 5.60]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 8.01 (P < 0.00001)								
4.10.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	2.6	3.7	24	-0.1	3.5	28	100.0%	2.70 [0.73 , 4.67]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**2.70 [0.73 , 4.67]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.69 (P = 0.007)								
4.10.3 6 months' follow-up								
Cadmus 2009	0.4	4.4	37	0.2	2.9	37	100.0%	0.20 [-1.50 , 1.90]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.20 [-1.50 , 1.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.23 (P = 0.82)								
Test for subgroup differences: Chi² = 17.56, df = 2 (P = 0.0002), I² = 88.6%								



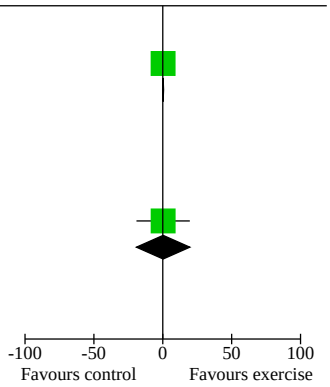
Analysis 4.11. Comparison 4: Body Image/self-esteem, Outcome 11: Rosenberg Self-Esteem Scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.11.1 Up to 12 weeks' follow-up								
Segar 1998	33	1.8	10	30.5	2.7	5	100.0%	2.50 [-0.12 , 5.12]
Subtotal (95% CI)			**10**			**5**	**100.0%**	**2.50 [-0.12 , 5.12]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.87 (P = 0.06)								
4.11.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	34.8	4.7	24	34.6	4.4	28	100.0%	0.20 [-2.29 , 2.69]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**0.20 [-2.29 , 2.69]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.16 (P = 0.87)								
4.11.3 6 months' follow-up								
Cadmus 2009	34.5	5.2	37	33.4	5.9	37	100.0%	1.10 [-1.43 , 3.63]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**1.10 [-1.43 , 3.63]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.85 (P = 0.39)								
Test for subgroup differences: Chi² = 1.57, df = 2 (P = 0.46), I² = 0%								



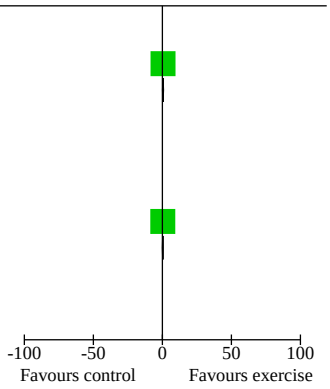
Analysis 4.12. Comparison 4: Body Image/self-esteem, Outcome 12: Physical Self-Perception Profile attractiveness of body subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.12.1 Up to 12 weeks' follow-up								
Daley 2007a	1.79	0.71	33	1.56	0.67	33	100.0%	0.23 [-0.10 , 0.56]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.23 [-0.10 , 0.56]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.35 (P = 0.18)								
4.12.2 More than 12 weeks' less than 6 months' follow-up								
Daley 2007a	2.02	55	31	1.76	0.57	31	100.0%	0.26 [-19.10 , 19.62]
Subtotal (95% CI)			**31**			**31**	**100.0%**	**0.26 [-19.10 , 19.62]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.03 (P = 0.98)								
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%								



Analysis 4.13. Comparison 4: Body Image/self-esteem, Outcome 13: Physical Self-Perception Profile physical self-worth subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
4.13.1 Up to 12 weeks' follow-up								
Daley 2007a	2.03	0.83	33	1.57	0.5	32	100.0%	0.46 [0.13 , 0.79]
Subtotal (95% CI)			**33**			**32**	**100.0%**	**0.46 [0.13 , 0.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.72 (P = 0.007)								
4.13.2 More than 12 weeks' less than 6 months' follow-up								
Daley 2007a	2.17	0.53	30	1.88	0.48	31	100.0%	0.29 [0.04 , 0.54]
Subtotal (95% CI)			**30**			**31**	**100.0%**	**0.29 [0.04 , 0.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.24 (P = 0.03)								
Test for subgroup differences: Chi² = 0.64, df = 1 (P = 0.43), I² = 0%								



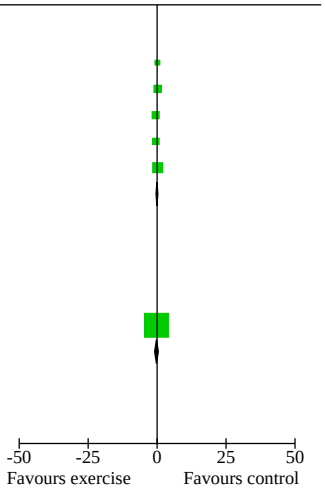
Comparison 5. Cognitive functioning

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
5.1 Overall cognitive functioning change	5		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
5.1.1 Up to 12 weeks' follow-up	5	332	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.38, 0.26]
5.1.2 6 months' follow-up	1	36	Std. Mean Difference (IV, Random, 95% CI)	-0.23 [-0.89, 0.43]
5.2 Overall cognitive functioning follow-up values	10		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
5.2.1 Up to 12 weeks' follow-up	9	439	Std. Mean Difference (IV, Random, 95% CI)	0.29 [-0.57, 1.16]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.2.2 More than 12 weeks' less than 6 months' follow-up	3	186	Std. Mean Difference (IV, Random, 95% CI)	-0.26 [-0.55, 0.03]
5.3 QLQ-C30 cognitive function change	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
5.3.1 Up to 12 weeks' follow-up	3	127	Mean Difference (IV, Fixed, 95% CI)	3.31 [-2.92, 9.53]
5.4 QLQ-C30 cognitive function follow-up values	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.4.1 Up to 12 weeks' follow-up	4	256	Mean Difference (IV, Random, 95% CI)	4.47 [-0.35, 9.28]
5.4.2 More than 12 weeks' less than 6 months' follow-up	2	150	Mean Difference (IV, Random, 95% CI)	6.15 [-0.30, 12.59]
5.5 FACT Cog change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.5.1 Up to 12 weeks' follow-up	1	38	Mean Difference (IV, Random, 95% CI)	-8.90 [-22.95, 5.15]
5.5.2 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	-4.72 [-18.12, 8.68]
5.6 FACT-Cog follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.6.1 12 weeks' follow-up	1	37	Mean Difference (IV, Random, 95% CI)	-11.00 [-27.52, 5.52]
5.6.2 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	-2.80 [-23.50, 17.90]
5.7 Linear Analog Self-Assessment Scale - confusion follow-up values	1	21	Mean Difference (IV, Random, 95% CI)	-4.90 [-17.70, 7.90]
5.7.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	-4.90 [-17.70, 7.90]
5.8 POMS confusion subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.8.1 Up to 12 weeks' follow-up	1	167	Mean Difference (IV, Random, 95% CI)	1.15 [-0.71, 3.01]
5.9 Profile of Mood State confusion subscale follow-up values	4	233	Mean Difference (IV, Random, 95% CI)	-1.62 [-2.53, -0.71]
5.9.1 Up to 12 weeks' follow-up	4	233	Mean Difference (IV, Random, 95% CI)	-1.62 [-2.53, -0.71]
5.10 Symptoms of Stress Inventory - cognitive disorganization subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.10.1 Up to 12 weeks' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	-1.67 [-3.66, 0.32]

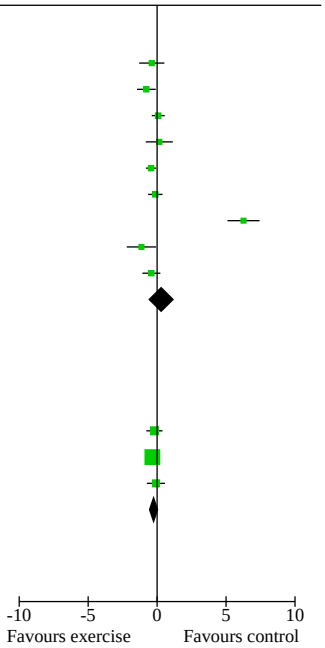
Analysis 5.1. Comparison 5: Cognitive functioning, Outcome 1: Overall cognitive functioning change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
5.1.1 Up to 12 weeks' follow-up								
Herrero 2006	-6.2	19.8	8	-8.3	14.8	8	8.7%	0.11 [-0.87 , 1.09]
Mehnert 2011	-1.67	22.36	30	-6.18	24.6	27	21.2%	0.19 [-0.33 , 0.71]
Oh 2010	-7.25	14.9	23	0.54	14.6	31	20.0%	-0.52 [-1.07 , 0.03]
Rogers 2009	-4.2	16.8	20	4.7	25.9	18	16.4%	-0.40 [-1.05 , 0.24]
Targ 2002	-2.03	7.1	79	-3.18	4.8	88	33.7%	0.19 [-0.11 , 0.50]
Subtotal (95% CI)			**160**			**172**	**100.0%**	**-0.06 [-0.38 , 0.26]**
Heterogeneity: Tau² = 0.06; Chi² = 7.04, df = 4 (P = 0.13); I² = 43%								
Test for overall effect: Z = 0.34 (P = 0.73)								
5.1.2 6 months' follow-up								
Rogers 2009	-5.4	16.3	19	-0.68	23.6	17	100.0%	-0.23 [-0.89 , 0.43]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**-0.23 [-0.89 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.69 (P = 0.49)								
Test for subgroup differences: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%								



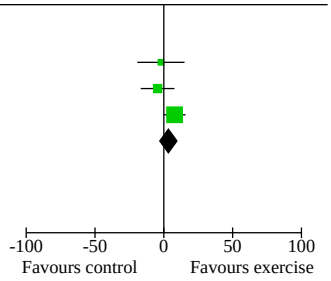
Analysis 5.2. Comparison 5: Cognitive functioning, Outcome 2: Overall cognitive functioning follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
5.2.1 Up to 12 weeks' follow-up								
Burnham 2002	6.6	10.2	14	11.5	15.7	7	10.8%	-0.39 [-1.30 , 0.53]
Culos-Reed 2006	-0.39	1.82	18	2.11	4.06	18	11.4%	-0.78 [-1.46 , -0.10]
Dimeo 2004	-78	25	34	-80	23	35	11.8%	0.08 [-0.39 , 0.55]
Herrero 2006	91.7	12.6	8	89.6	12.4	8	10.6%	0.16 [-0.82 , 1.14]
Knols 2011	-87.6	16.7	57	-79.2	21.1	57	12.0%	-0.44 [-0.81 , -0.07]
Mehnert 2011	-81.1	16.8	30	-78.4	24.38	27	11.7%	-0.13 [-0.65 , 0.39]
Moadel 2007	29.2	3.86	45	4.35	4.03	26	10.0%	6.27 [5.10 , 7.43]
Pinto 2003	4.33	2.5	12	7.17	2.1	6	10.3%	-1.14 [-2.20 , -0.07]
Rogers 2009	124.5	30.8	19	135.5	19.5	18	11.4%	-0.42 [-1.07 , 0.24]
Subtotal (95% CI)			**237**			**202**	**100.0%**	**0.29 [-0.57 , 1.16]**
Heterogeneity: Tau² = 1.59; Chi² = 127.39, df = 8 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 0.66 (P = 0.51)								
5.2.2 More than 12 weeks' less than 6 months' follow-up								
Bai 2004	-79.11	19.37	24	-74.75	25.19	21	24.3%	-0.19 [-0.78 , 0.39]
Knols 2011	-85.6	17.8	51	-78.9	20.7	54	56.2%	-0.34 [-0.73 , 0.04]
Rogers 2009	126.4	31.8	19	129.2	31.5	17	19.5%	-0.09 [-0.74 , 0.57]
Subtotal (95% CI)			**94**			**92**	**100.0%**	**-0.26 [-0.55 , 0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 0.50, df = 2 (P = 0.78); I² = 0%								
Test for overall effect: Z = 1.74 (P = 0.08)								
Test for subgroup differences: Chi² = 1.39, df = 1 (P = 0.24), I² = 28.2%								



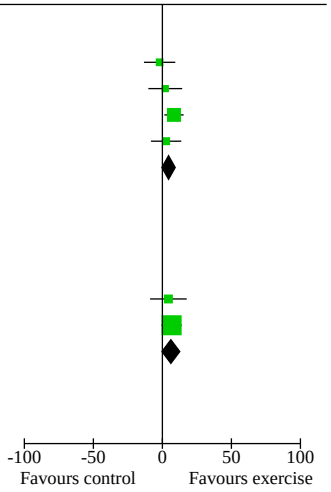
Analysis 5.3. Comparison 5: Cognitive functioning, Outcome 3: QLQ-C30 cognitive function change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
5.3.1 Up to 12 weeks' follow-up								
Herrero 2006	6.2	19.8	8	8.3	14.8	8	13.2%	-2.10 [-19.23 , 15.03]
Mehnert 2011	1.67	22.36	30	6.18	24.6	27	25.8%	-4.51 [-16.76 , 7.74]
Oh 2010	7.25	14.9388	23	-0.54	14.5855	31	60.9%	7.79 [-0.19 , 15.77]
Subtotal (95% CI)			**61**			**66**	**100.0%**	**3.31 [-2.92 , 9.53]**
Heterogeneity: Chi² = 3.16, df = 2 (P = 0.21); I² = 37%								
Test for overall effect: Z = 1.04 (P = 0.30)								
Test for subgroup differences: Not applicable								



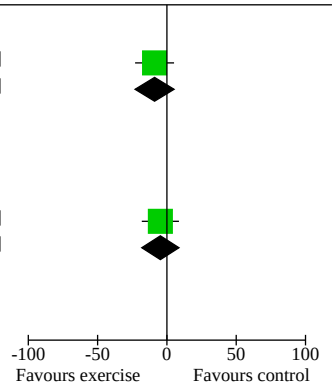
Analysis 5.4. Comparison 5: Cognitive functioning, Outcome 4: QLQ-C30 cognitive function follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.4.1 Up to 12 weeks' follow-up								
Dimeo 2004	78	25	34	80	23	35	18.0%	-2.00 [-13.34 , 9.34]
Herrero 2006	91.7	12.6	8	89.6	12.4	8	15.4%	2.10 [-10.15 , 14.35]
Knols 2011	87.6	16.7	57	79.2	21.1	57	47.4%	8.40 [1.41 , 15.39]
Mehnert 2011	81.1	16.8	30	78.4	24.38	27	19.2%	2.70 [-8.29 , 13.69]
Subtotal (95% CI)			**129**			**127**	**100.0%**	**4.47 [-0.35 , 9.28]**
Heterogeneity: Tau² = 0.00; Chi² = 2.71, df = 3 (P = 0.44); I² = 0%								
Test for overall effect: Z = 1.82 (P = 0.07)								
5.4.2 More than 12 weeks' less than 6 months' follow-up								
Bai 2004	79.11	19.37	24	74.75	25.19	21	23.6%	4.36 [-8.91 , 17.63]
Knols 2011	85.6	17.8	51	78.9	20.7	54	76.4%	6.70 [-0.67 , 14.07]
Subtotal (95% CI)			**75**			**75**	**100.0%**	**6.15 [-0.30 , 12.59]**
Heterogeneity: Tau² = 0.00; Chi² = 0.09, df = 1 (P = 0.76); I² = 0%								
Test for overall effect: Z = 1.87 (P = 0.06)								
Test for subgroup differences: Chi² = 0.17, df = 1 (P = 0.68), I² = 0%								



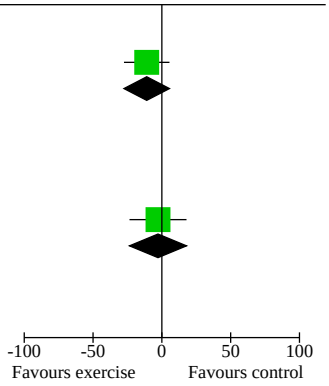
Analysis 5.5. Comparison 5: Cognitive functioning, Outcome 5: FACT Cog change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.5.1 Up to 12 weeks' follow-up								
Rogers 2009	-4.2	16.8	20	4.7	25.9	18	100.0%	-8.90 [-22.95 , 5.15]
Subtotal (95% CI)			**20**			**18**	**100.0%**	**-8.90 [-22.95 , 5.15]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.24 (P = 0.21)								
5.5.2 6 months' follow-up								
Rogers 2009	-5.4	16.3	19	-0.68	23.6	17	100.0%	-4.72 [-18.12 , 8.68]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**-4.72 [-18.12 , 8.68]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.69 (P = 0.49)								
Test for subgroup differences: Chi² = 0.18, df = 1 (P = 0.67), I² = 0%								



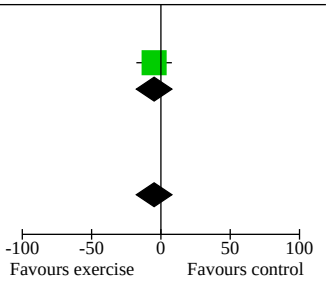
Analysis 5.6. Comparison 5: Cognitive functioning, Outcome 6: FACT-Cog follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.6.1 12 weeks' follow-up								
Rogers 2009	124.5	30.8	19	135.5	19.5	18	100.0%	-11.00 [-27.52 , 5.52]
Subtotal (95% CI)			**19**			**18**	**100.0%**	**-11.00 [-27.52 , 5.52]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.30 (P = 0.19)								
5.6.2 6 months' follow-up								
Rogers 2009	126.4	31.8	19	129.2	31.5	17	100.0%	-2.80 [-23.50 , 17.90]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**-2.80 [-23.50 , 17.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.27 (P = 0.79)								
Test for subgroup differences: Chi² = 0.37, df = 1 (P = 0.54), I² = 0%								



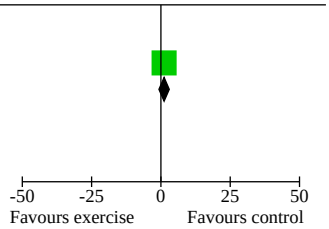
Analysis 5.7. Comparison 5: Cognitive functioning, Outcome 7: Linear Analog Self-Assessment Scale - confusion follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.7.1 Up to 12 weeks' follow-up								
Burnham 2002	6.6	10.2	14	11.5	15.7	7	100.0%	-4.90 [-17.70 , 7.90]
Subtotal (95% CI)			**14**			**7**	**100.0%**	**-4.90 [-17.70 , 7.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								
Total (95% CI)			**14**			**7**	**100.0%**	**-4.90 [-17.70 , 7.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								
Test for subgroup differences: Not applicable								

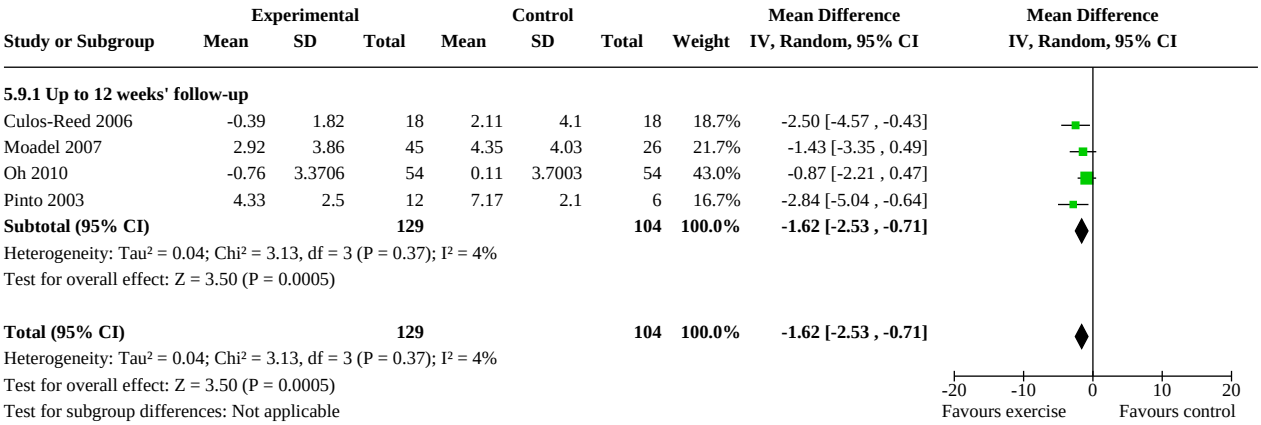


Analysis 5.8. Comparison 5: Cognitive functioning, Outcome 8: POMS confusion subscale change

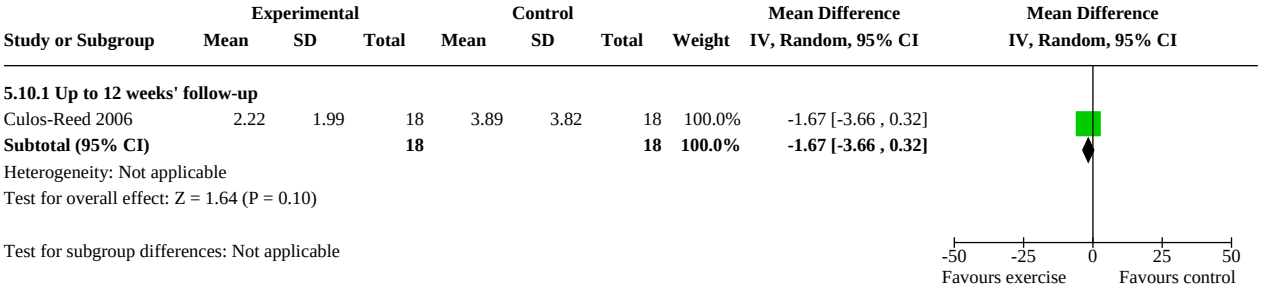
Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.8.1 Up to 12 weeks' follow-up								
Targ 2002	-2.03	7.1	79	-3.18	4.8	88	100.0%	1.15 [-0.71 , 3.01]
Subtotal (95% CI)			**79**			**88**	**100.0%**	**1.15 [-0.71 , 3.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.23)								
Test for subgroup differences: Not applicable								



Analysis 5.9. Comparison 5: Cognitive functioning, Outcome 9: Profile of Mood State confusion subscale follow-up values



Analysis 5.10. Comparison 5: Cognitive functioning, Outcome 10: Symptoms of Stress Inventory - cognitive disorganization subscale follow-up values



Comparison 6. Depression

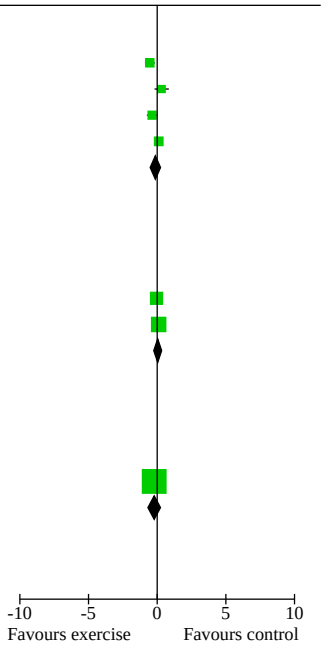
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 Overall depression change	7		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
6.1.1 Up to 12 weeks' follow-up	4	455	Std. Mean Difference (IV, Random, 95% CI)	-0.13 [-0.51, 0.24]
6.1.2 More than 12 weeks' less than 6 months' follow-up	2	196	Std. Mean Difference (IV, Random, 95% CI)	0.04 [-0.25, 0.33]
6.1.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	-0.22 [-0.68, 0.24]
6.2 Overall depression follow-up values	15		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
6.2.1 Up to 12 weeks' follow-up	12	707	Std. Mean Difference (IV, Random, 95% CI)	-0.41 [-0.65, -0.17]
6.2.2 More than 12 weeks' less than 6 months' follow-up	3	258	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.41, 0.20]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.2.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	0.12 [-0.33, 0.58]
6.3 Centers for Epidemiologic Studies - Depression Scale change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.3.1 Up to 12 weeks' follow-up	1	122	Mean Difference (IV, Random, 95% CI)	-2.40 [-4.05, -0.75]
6.3.2 More than 12 weeks' up to 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-0.50 [-4.59, 3.59]
6.3.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-1.40 [-4.29, 1.49]
6.4 Centers for Epidemiologic Studies - Depression Scale follow-up values	5	346	Mean Difference (IV, Random, 95% CI)	-0.98 [-2.44, 0.49]
6.4.1 Up to 12 weeks' follow-up	3	179	Mean Difference (IV, Random, 95% CI)	-2.68 [-8.28, 2.93]
6.4.2 More than 12 weeks' less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-1.00 [-5.41, 3.41]
6.4.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-1.20 [-5.62, 3.22]
6.5 Hospital Anxiety and Depression Scale - depression subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.5.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-1.32 [-3.38, 0.74]
6.5.2 More than 12 weeks' up to 6 months' follow-up	1	103	Mean Difference (IV, Random, 95% CI)	0.50 [-1.33, 2.33]
6.6 Hospital Anxiety Depression Scale- depression subscale follow-up values	2	161	Mean Difference (IV, Random, 95% CI)	-0.65 [-2.82, 1.52]
6.6.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-1.94 [-3.89, 0.01]
6.6.2 More than 12 weeks' less than 6 months' follow-up	1	103	Mean Difference (IV, Random, 95% CI)	0.30 [-0.63, 1.23]
6.7 Profile of Moods Scale - depression subscale follow-up values	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.7.1 Up to 12 weeks' follow-up	4	335	Mean Difference (IV, Random, 95% CI)	-2.51 [-4.28, -0.74]
6.8 Beck Depression Inventory II follow-up values	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.8.1 Up to 12 weeks' follow-up	3	114	Mean Difference (IV, Random, 95% CI)	-4.28 [-6.01, -2.55]
6.8.2 More than 12 weeks' less than 6 months' follow-up	1	62	Mean Difference (IV, Random, 95% CI)	-3.03 [-6.50, 0.44]
6.9 Linear Analog Self-Assessment scale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
6.9.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	-12.70 [-24.86, -0.54]

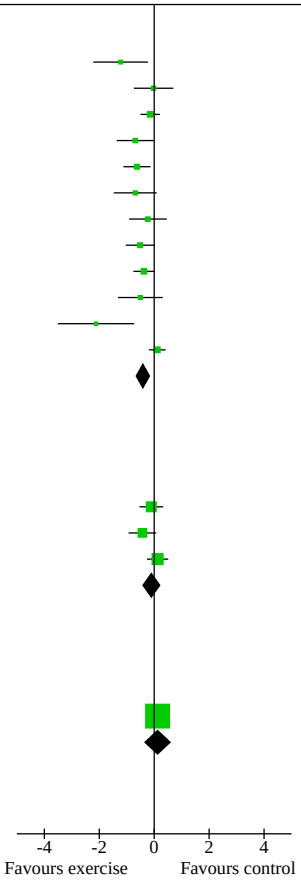
Analysis 6.1. Comparison 6: Depression, Outcome 1: Overall depression change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
6.1.1 Up to 12 weeks' follow-up								
Courneya 2009	-2.2	4.6453	60	0.2	4.3315	62	26.0%	-0.53 [-0.89 , -0.17]
Mehnert 2011	1.47	3.39	30	0.15	4.5	28	20.7%	0.33 [-0.19 , 0.85]
Oh 2010	-1.01	5.8986	54	1.54	7.5472	54	25.3%	-0.37 [-0.75 , 0.01]
Targ 2002	-0.377	25.3273	79	-2.55	11.6576	88	27.9%	0.11 [-0.19 , 0.42]
Subtotal (95% CI)			**223**			**232**	**100.0%**	**-0.13 [-0.51 , 0.24]**
Heterogeneity: Tau² = 0.11; Chi² = 11.72, df = 3 (P = 0.008); I² = 74%								
Test for overall effect: Z = 0.69 (P = 0.49)								
6.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-1	10.7	62	-0.5	8.8	31	44.6%	-0.05 [-0.48 , 0.38]
Thorsen 2005	-0.68	2.8	53	-1.18	6	50	55.4%	0.11 [-0.28 , 0.49]
Subtotal (95% CI)			**115**			**81**	**100.0%**	**0.04 [-0.25 , 0.33]**
Heterogeneity: Tau² = 0.00; Chi² = 0.28, df = 1 (P = 0.60); I² = 0%								
Test for overall effect: Z = 0.26 (P = 0.80)								
6.1.3 6 months' follow-up								
Cadmus 2009	0.3	6.4	37	1.7	6.3	37	100.0%	-0.22 [-0.68 , 0.24]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.22 [-0.68 , 0.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.94 (P = 0.35)								
Test for subgroup differences: Chi² = 1.04, df = 2 (P = 0.59), I² = 0%								



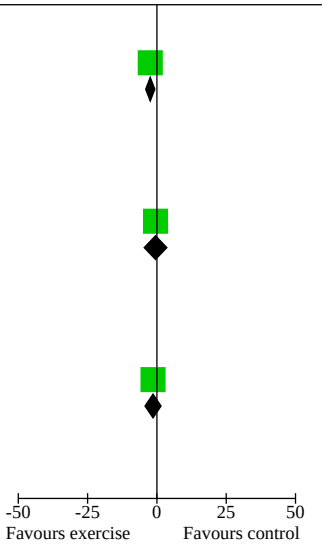
Analysis 6.2. Comparison 6: Depression, Outcome 2: Overall depression follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
6.2.1 Up to 12 weeks' follow-up								
Burnham 2002	4.8	5.2	14	17.5	16	7	4.4%	-1.22 [-2.22 , -0.23]
Cohen 2004	9	42	16	9.7	3.8	14	6.9%	-0.02 [-0.74 , 0.70]
Courneya 2009	5.4	4.5	60	6.1	5	62	12.9%	-0.15 [-0.50 , 0.21]
Culos-Reed 2006	2.22	2.65	18	5.5	6.03	18	7.4%	-0.69 [-1.36 , -0.01]
Daley 2007a	6	6.47	33	10.33	7.19	33	10.2%	-0.63 [-1.12 , -0.13]
Danhauer 2009	8.1	8.9	13	17.8	16.9	14	6.2%	-0.69 [-1.47 , 0.09]
Donnelly 2011	11.25	8.29	16	13.41	10.1	17	7.3%	-0.23 [-0.91 , 0.46]
Mehnert 2011	2.7	2.95	30	4.64	4.43	28	9.7%	-0.51 [-1.04 , 0.01]
Oh 2010	-1.01	5.8986	54	1.54	7.5472	54	12.4%	-0.37 [-0.75 , 0.01]
Pinto 2003	6.17	7.2	12	9.83	6.8	12	5.9%	-0.50 [-1.32 , 0.31]
Segar 1998	5.5	2	10	10	2	5	2.6%	-2.12 [-3.51 , -0.73]
Targ 2002	-0.377	25.3273	79	-2.55	11.6576	88	14.0%	0.11 [-0.19 , 0.42]
Subtotal (95% CI)			**355**			**352**	**100.0%**	**-0.41 [-0.65 , -0.17]**
Heterogeneity: Tau² = 0.08; Chi² = 23.16, df = 11 (P = 0.02); I² = 53%								
Test for overall effect: Z = 3.37 (P = 0.0007)								
6.2.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	8.6	8.7	62	9.6	10.9	31	33.8%	-0.10 [-0.54 , 0.33]
Daley 2007a	6.52	6.95	31	9.55	7	31	27.1%	-0.43 [-0.93 , 0.07]
Thorsen 2005	2.2	2.3	53	1.9	2.5	50	39.0%	0.12 [-0.26 , 0.51]
Subtotal (95% CI)			**146**			**112**	**100.0%**	**-0.10 [-0.41 , 0.20]**
Heterogeneity: Tau² = 0.02; Chi² = 2.92, df = 2 (P = 0.23); I² = 31%								
Test for overall effect: Z = 0.66 (P = 0.51)								
6.2.3 6 months' follow-up								
Cadmus 2009	-9.6	9.3	37	-10.8	10.1	37	100.0%	0.12 [-0.33 , 0.58]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.12 [-0.33 , 0.58]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.60)								
Test for subgroup differences: Chi² = 5.19, df = 2 (P = 0.07), I² = 61.5%								



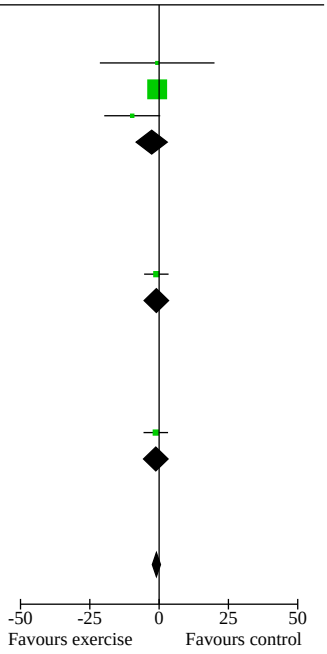
Analysis 6.3. Comparison 6: Depression, Outcome 3: Centers for Epidemiologic Studies - Depression Scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.3.1 Up to 12 weeks' follow-up								
Courneya 2009	-2.2	4.6	60	0.2	4.7	62	100.0%	-2.40 [-4.05 , -0.75]
Subtotal (95% CI)			**60**			**62**	**100.0%**	**-2.40 [-4.05 , -0.75]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.85 (P = 0.004)								
6.3.2 More than 12 weeks' up to 6 months' follow-up								
Courneya 2003b	-1	10.7	62	-0.5	8.8	31	100.0%	-0.50 [-4.59 , 3.59]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-0.50 [-4.59 , 3.59]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
6.3.3 6 months' follow-up								
Cadmus 2009	0.3	6.4	37	1.7	6.3	37	100.0%	-1.40 [-4.29 , 1.49]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-1.40 [-4.29 , 1.49]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.95 (P = 0.34)								
Test for subgroup differences: Chi² = 0.91, df = 2 (P = 0.64), I² = 0%								



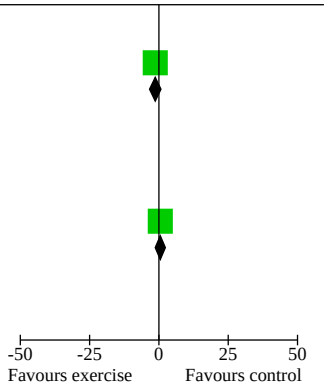
Analysis 6.4. Comparison 6: Depression, Outcome 4: Centers for Epidemiologic Studies - Depression Scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.4.1 Up to 12 weeks' follow-up								
Cohen 2004	9	42	16	9.7	3.8	14	0.5%	-0.70 [-21.38 , 19.98]
Courneya 2009	5.4	4.5	60	6.1	5	62	75.4%	-0.70 [-2.39 , 0.99]
Danhauer 2009	8.1	8.9	13	17.8	16.9	14	2.1%	-9.70 [-19.79 , 0.39]
Subtotal (95% CI)			**89**			**90**	**78.0%**	**-2.68 [-8.28 , 2.93]**
Heterogeneity: Tau² = 10.73; Chi² = 2.97, df = 2 (P = 0.23); I² = 33%								
Test for overall effect: Z = 0.94 (P = 0.35)								
6.4.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	8.6	8.7	62	9.6	10.9	31	11.0%	-1.00 [-5.41 , 3.41]
Subtotal (95% CI)			**62**			**31**	**11.0%**	**-1.00 [-5.41 , 3.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.44 (P = 0.66)								
6.4.3 6 months' follow-up								
Cadmus 2009	9.6	9.3	37	10.8	10.1	37	11.0%	-1.20 [-5.62 , 3.22]
Subtotal (95% CI)			**37**			**37**	**11.0%**	**-1.20 [-5.62 , 3.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.59)								
Total (95% CI)			**188**			**158**	**100.0%**	**-0.98 [-2.44 , 0.49]**
Heterogeneity: Tau² = 0.00; Chi² = 2.99, df = 4 (P = 0.56); I² = 0%								
Test for overall effect: Z = 1.31 (P = 0.19)								
Test for subgroup differences: Chi² = 0.24, df = 2 (P = 0.89), I² = 0%								



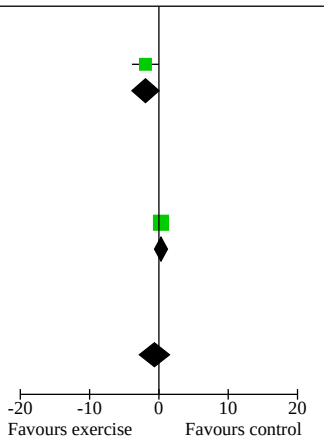
Analysis 6.5. Comparison 6: Depression, Outcome 5: Hospital Anxiety and Depression Scale - depression subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.5.1 Up to 12 weeks' follow-up								
Mehnert 2011	-1.47	3.39	30	-0.15	4.5	28	100.0%	-1.32 [-3.38 , 0.74]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**-1.32 [-3.38 , 0.74]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.25 (P = 0.21)								
6.5.2 More than 12 weeks' up to 6 months' follow-up								
Thorsen 2005	-0.68	2.8	53	-1.18	6	50	100.0%	0.50 [-1.33 , 2.33]
Subtotal (95% CI)			**53**			**50**	**100.0%**	**0.50 [-1.33 , 2.33]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.54 (P = 0.59)								
Test for subgroup differences: Chi² = 1.68, df = 1 (P = 0.20), I² = 40.4%								



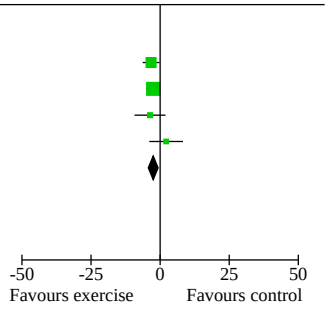
Analysis 6.6. Comparison 6: Depression, Outcome 6: Hospital Anxiety Depression Scale- depression subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.6.1 Up to 12 weeks' follow-up								
Mehnert 2011	2.7	2.95	30	4.64	4.43	28	42.4%	-1.94 [-3.89 , 0.01]
Subtotal (95% CI)			**30**			**28**	**42.4%**	**-1.94 [-3.89 , 0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.95 (P = 0.05)								
6.6.2 More than 12 weeks' less than 6 months' follow-up								
Thorsen 2005	2.2	2.3	53	1.9	2.5	50	57.6%	0.30 [-0.63 , 1.23]
Subtotal (95% CI)			**53**			**50**	**57.6%**	**0.30 [-0.63 , 1.23]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.63 (P = 0.53)								
Total (95% CI)			**83**			**78**	**100.0%**	**-0.65 [-2.82 , 1.52]**
Heterogeneity: Tau² = 1.90; Chi² = 4.13, df = 1 (P = 0.04); I² = 76%								
Test for overall effect: Z = 0.59 (P = 0.56)								
Test for subgroup differences: Chi² = 4.13, df = 1 (P = 0.04), I² = 75.8%								



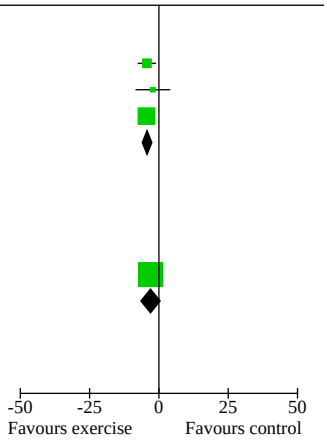
Analysis 6.7. Comparison 6: Depression, Outcome 7: Profile of Moods Scale - depression subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.7.1 Up to 12 weeks' follow-up								
Culos-Reed 2006	2.22	2.65	18	5.5	6.03	18	33.8%	-3.28 [-6.32 , -0.24]
Oh 2010	-1.01	5.8986	54	1.54	7.5472	54	47.9%	-2.55 [-5.10 , 0.00]
Pinto 2003	6.17	7.2	12	9.83	6.8	12	10.0%	-3.66 [-9.26 , 1.94]
Targ 2002	-0.377	25.3273	79	-2.55	11.6576	88	8.4%	2.17 [-3.92 , 8.27]
Subtotal (95% CI)			**163**			**172**	**100.0%**	**-2.51 [-4.28 , -0.74]**
Heterogeneity: Tau² = 0.00; Chi² = 2.68, df = 3 (P = 0.44); I² = 0%								
Test for overall effect: Z = 2.78 (P = 0.005)								
Test for subgroup differences: Not applicable								

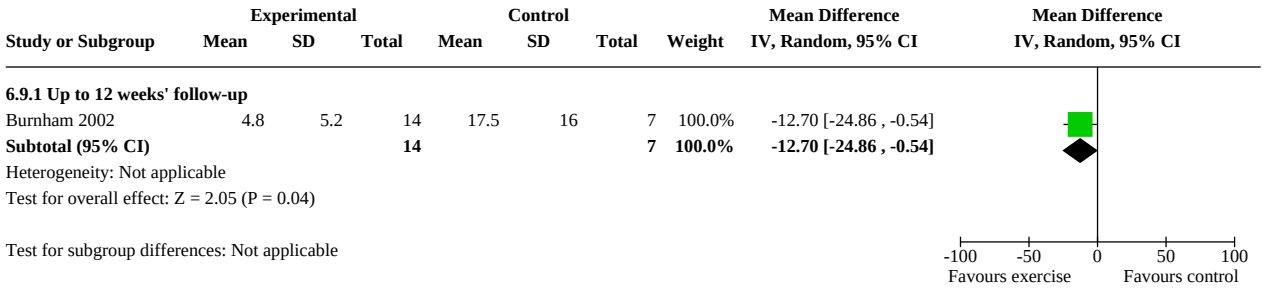


Analysis 6.8. Comparison 6: Depression, Outcome 8: Beck Depression Inventory II follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.8.1 Up to 12 weeks' follow-up								
Daley 2007a	6	6.47	33	10.33	7.19	33	27.5%	-4.33 [-7.63 , -1.03]
Donnelly 2011	11.25	8.29	16	13.41	10.1	17	7.6%	-2.16 [-8.45 , 4.13]
Segar 1998	5.5	2	10	10	2	5	64.9%	-4.50 [-6.65 , -2.35]
Subtotal (95% CI)			**59**			**55**	**100.0%**	**-4.28 [-6.01 , -2.55]**
Heterogeneity: Tau² = 0.00; Chi² = 0.48, df = 2 (P = 0.79); I² = 0%								
Test for overall effect: Z = 4.84 (P < 0.00001)								
6.8.2 More than 12 weeks' less than 6 months' follow-up								
Daley 2007a	6.52	6.95	31	9.55	7	31	100.0%	-3.03 [-6.50 , 0.44]
Subtotal (95% CI)			**31**			**31**	**100.0%**	**-3.03 [-6.50 , 0.44]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.71 (P = 0.09)								
Test for subgroup differences: Chi² = 0.40, df = 1 (P = 0.53), I² = 0%								



Analysis 6.9. Comparison 6: Depression, Outcome 9: Linear Analog Self-Assessment scale follow-up values



Comparison 7. Emotional well-being/mental health

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 Overall emotional well-being/mental health change	15		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
7.1.1 Up to 12 weeks' follow-up	8	632	Std. Mean Difference (IV, Random, 95% CI)	0.33 [0.05, 0.61]
7.1.2 More than 12 weeks' less than 6 months' follow-up	3	246	Std. Mean Difference (IV, Random, 95% CI)	0.13 [-0.34, 0.60]
7.1.3 6 months' follow-up	4	271	Std. Mean Difference (IV, Random, 95% CI)	0.60 [0.17, 1.03]
7.1.4 More than 6 months' follow-up	2	202	Std. Mean Difference (IV, Random, 95% CI)	0.08 [-0.19, 0.36]
7.2 Overall emotional well-being/mental health follow-up values	27		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
7.2.1 Up to 12 weeks' follow-up	19	1086	Std. Mean Difference (IV, Fixed, 95% CI)	0.24 [0.12, 0.37]
7.2.2 More than 12 weeks' less than 6 months' follow-up	9	666	Std. Mean Difference (IV, Fixed, 95% CI)	0.17 [0.01, 0.32]
7.2.3 6 months' follow-up	3	189	Std. Mean Difference (IV, Fixed, 95% CI)	0.13 [-0.16, 0.41]
7.2.4 More than 6 months' follow-up	1	120	Std. Mean Difference (IV, Fixed, 95% CI)	-0.22 [-0.58, 0.14]
7.3 FACT emotional subscale change	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.3.1 Up to 12 weeks' follow-up	3	313	Mean Difference (IV, Random, 95% CI)	0.67 [-0.71, 2.05]
7.3.2 More than 12 weeks' less than 6 months' follow-up	2	145	Mean Difference (IV, Random, 95% CI)	0.18 [-2.26, 2.63]
7.3.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	0.80 [-0.17, 1.76]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.4 FACT emotional subscale follow-up values	10		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.4.1 Up to 12 weeks' follow-up	7	372	Mean Difference (IV, Random, 95% CI)	1.14 [-0.08, 2.35]
7.4.2 More than 12 weeks' less than 6 months' follow-up	4	263	Mean Difference (IV, Random, 95% CI)	0.36 [-0.47, 1.19]
7.4.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	-0.22 [-2.18, 1.73]
7.5 QLQ-C30 subscale change	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.5.1 Up to 12 weeks' follow-up	3	142	Mean Difference (IV, Random, 95% CI)	4.26 [-4.19, 12.72]
7.5.2 More than 12 weeks' less than 6 months' follow-up	1	101	Mean Difference (IV, Random, 95% CI)	-4.80 [-11.64, 2.04]
7.6 QLQ-C30 subscale follow-up values	7		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.6.1 Up to 12 weeks' follow-up	5	292	Mean Difference (IV, Random, 95% CI)	5.56 [0.55, 10.56]
7.6.2 More than 12 weeks' less than 6 months' follow-up	3	251	Mean Difference (IV, Random, 95% CI)	-0.55 [-6.09, 4.99]
7.7 MOS SF-36 emotional role subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.7.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-1.27 [-15.55, 13.01]
7.7.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	1.80 [-2.32, 5.92]
7.8 MOS SF-36 emotional role subscale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.8.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	7.96 [-11.76, 27.68]
7.8.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	2.80 [-2.32, 7.92]
7.9 MOS SF-36 mental health component change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.9.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	2.26 [-7.22, 11.74]
7.9.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	1.70 [-1.82, 5.22]
7.9.3 More than 6 months' follow-up	1	120	Mean Difference (IV, Random, 95% CI)	0.20 [-7.15, 7.55]
7.10 MOS SF-36 mental health component follow-up values	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.10.1 Up to 12 weeks' follow-up	2	129	Mean Difference (IV, Random, 95% CI)	9.45 [4.75, 14.16]

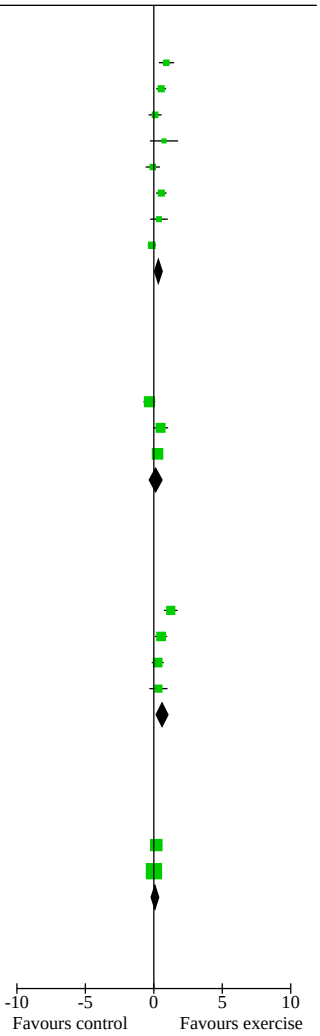
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.10.2 More than 12 weeks' less than 6 months' follow-up	1	60	Mean Difference (IV, Random, 95% CI)	13.85 [8.11, 19.59]
7.10.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	3.20 [-2.02, 8.42]
7.10.4 More than 6 months' follow-up	1	120	Mean Difference (IV, Random, 95% CI)	-2.00 [-5.23, 1.23]
7.11 MOS SF-12 mental health component follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.11.1 Up to 12 weeks' follow-up	2	114	Mean Difference (IV, Random, 95% CI)	1.65 [-1.62, 4.92]
7.11.2 More than 12 weeks' less than 6 months' follow-up	1	87	Mean Difference (IV, Random, 95% CI)	3.42 [-0.15, 6.99]
7.12 POMS total mood disturbance change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.12.1 Up to 12 weeks' follow-up	2	275	Mean Difference (IV, Random, 95% CI)	-8.08 [-15.03, -1.12]
7.12.2 6 months' follow-up	1	82	Mean Difference (IV, Random, 95% CI)	-7.67 [-19.29, 3.95]
7.12.3 More than 6 months' follow-up	1	82	Mean Difference (IV, Random, 95% CI)	-1.09 [-12.84, 10.66]
7.13 POMS total mood disturbance follow-up values	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.13.1 Up to 12 weeks' follow-up	4	213	Mean Difference (IV, Random, 95% CI)	-10.43 [-16.36, -4.51]
7.14 POMS - anxiety and depression subscales follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.14.1 Up to 12 weeks' follow-up	1	87	Mean Difference (IV, Random, 95% CI)	-0.13 [-1.95, 1.69]
7.14.2 More than 12 weeks' less than 6 months' follow-up	1	87	Mean Difference (IV, Random, 95% CI)	-1.98 [-3.97, 0.01]
7.15 POMS anger subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.15.1 Up to 12 weeks' follow-up	2	275	Mean Difference (IV, Random, 95% CI)	0.05 [-1.48, 1.57]
7.16 POMS anger subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.16.1 Up to 12 weeks' follow-up	1	18	Mean Difference (IV, Random, 95% CI)	-1.50 [-5.20, 2.20]
7.17 CARES subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.17.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-2.20 [-4.04, -0.36]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.18 CARES subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.18.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-2.60 [-6.22, 1.02]
7.19 Cohen's perceived stress scale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.19.1 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-0.60 [-3.36, 2.16]
7.20 Cohen's perceived stress scale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.20.1 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-0.90 [-4.29, 2.49]
7.21 Fordyce change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.21.1 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	3.10 [-5.36, 11.56]
7.22 Fordyce follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.22.1 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	0.50 [-9.78, 10.78]
7.23 Happiness change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.23.1 More than 12 weeks' less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	16.50 [3.02, 29.98]
7.24 Happiness follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.24.1 More than 12 weeks' less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	7.10 [-5.68, 19.88]
7.25 Linear Analog Self-Assessment Scale - anger follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.25.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	2.50 [-17.49, 22.49]
7.26 Symptoms of Stress Index - emotional irritability subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.26.1 Up to 12 weeks' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	-2.39 [-4.79, 0.01]
7.27 PANAS - positivity follow-up values	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.27.1 Up to 12 weeks' follow-up	3	78	Mean Difference (IV, Random, 95% CI)	3.59 [-0.18, 7.37]
7.28 PANAS - negativity follow-up values	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.28.1 Up to 12 weeks' follow-up	3	78	Mean Difference (IV, Random, 95% CI)	-4.01 [-7.26, -0.77]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
7.29 Satisfaction with Life Scale change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.29.1 Up to 12 weeks' follow-up	1	122	Mean Difference (IV, Random, 95% CI)	10.30 [3.51, 17.09]
7.29.2 More than 12 weeks' less than 6 months' follow-up	2	149	Mean Difference (IV, Random, 95% CI)	-0.17 [-1.37, 1.03]
7.30 Satisfaction with Life Scale follow-up values	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.30.1 Up to 12 weeks' follow-up	3	284	Mean Difference (IV, Random, 95% CI)	-0.16 [-1.89, 1.58]
7.30.2 More than 12 weeks' less than 6 months' follow-up	1	93	Mean Difference (IV, Random, 95% CI)	-0.90 [-3.46, 1.66]
7.31 Lee Psychosocial Adjustment instrument change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.31.1 Up to 12 weeks' follow-up	1	55	Mean Difference (IV, Random, 95% CI)	5.90 [2.57, 9.23]
7.32 Lee Psychosocial Adjustmen-t Instrument follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.32.1 Up to 12 weeks' follow-up	1	55	Mean Difference (IV, Random, 95% CI)	4.80 [1.21, 8.39]

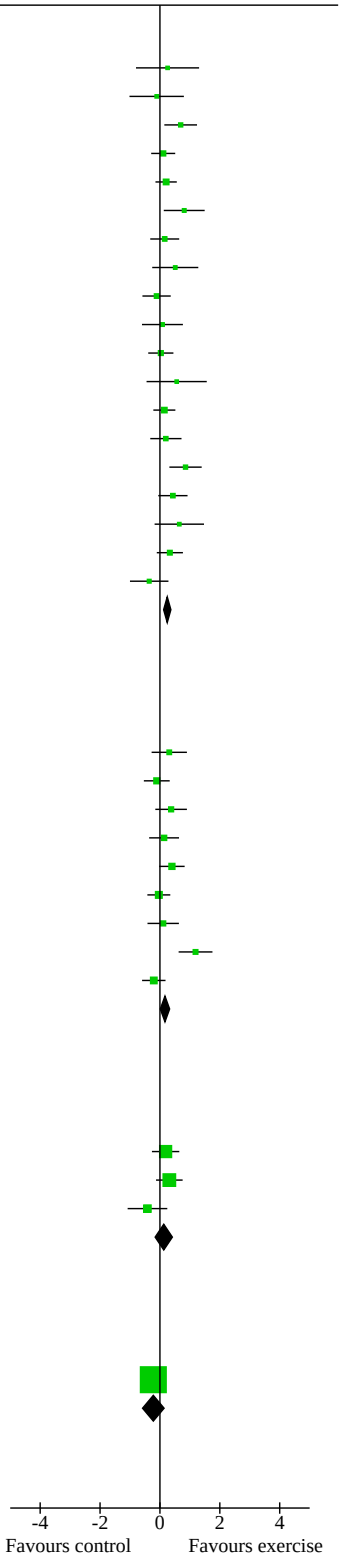
Analysis 7.1. Comparison 7: Emotional well-being/mental health, Outcome 1: Overall emotional well-being/mental health change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
7.1.1 Up to 12 weeks' follow-up								
Cho 2006	2.9	6.3	28	-3	6.3	27	11.5%	0.92 [0.37 , 1.48]
Courneya 2009	11	19.7424	60	0.7	18.5074	62	15.6%	0.54 [0.17 , 0.90]
Dimeo 2004	11	30	34	8	34	35	13.2%	0.09 [-0.38 , 0.56]
Herrero 2006	9.3	19.6	8	-2.1	5.9	8	5.6%	0.74 [-0.28 , 1.77]
Mehnert 2011	3.06	27.13	30	5.25	29.2	27	12.2%	-0.08 [-0.60 , 0.44]
Oh 2010	1.6	2.7	54	0.05	2.9	54	15.1%	0.55 [0.16 , 0.93]
Rogers 2009	0.9	2.5	20	-0.1	2.7	18	10.0%	0.38 [-0.27 , 1.02]
Targ 2002	1.68	4.4199	79	2.37	5.1916	88	16.8%	-0.14 [-0.45 , 0.16]
Subtotal (95% CI)			**313**			**319**	**100.0%**	**0.33 [0.05 , 0.61]**
Heterogeneity: Tau² = 0.10; Chi² = 19.58, df = 7 (P = 0.007); I² = 64%								
Test for overall effect: Z = 2.30 (P = 0.02)								
7.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	0.3	3.4	62	1.3	2.3	31	34.5%	-0.32 [-0.76 , 0.11]
Courneya 2003c	2.3	3.1	24	0.8	2.9	28	29.0%	0.49 [-0.06 , 1.05]
Thorsen 2005	-1	18.8027	52	-5.8	15.8045	49	36.5%	0.27 [-0.12 , 0.67]
Subtotal (95% CI)			**138**			**108**	**100.0%**	**0.13 [-0.34 , 0.60]**
Heterogeneity: Tau² = 0.12; Chi² = 6.32, df = 2 (P = 0.04); I² = 68%								
Test for overall effect: Z = 0.55 (P = 0.58)								
7.1.3 6 months' follow-up								
Cadmus 2009	3	2.2	37	-0.5	3.3	37	25.3%	1.23 [0.74 , 1.73]
Ohira 2006	2.5	4.4	39	0.3	3.9	40	27.0%	0.52 [0.08 , 0.97]
Pinto 2005	-3.8	27.7	39	-11.47	25.8	43	27.5%	0.28 [-0.15 , 0.72]
Rogers 2009	0.5	2.7	19	-0.29	1.7	17	20.2%	0.34 [-0.32 , 1.00]
Subtotal (95% CI)			**134**			**137**	**100.0%**	**0.60 [0.17 , 1.03]**
Heterogeneity: Tau² = 0.13; Chi² = 8.94, df = 3 (P = 0.03); I² = 66%								
Test for overall effect: Z = 2.72 (P = 0.007)								
7.1.4 More than 6 months' follow-up								
Pinto 2005	3.09	27.9	39	-2	26.2	43	40.5%	0.19 [-0.25 , 0.62]
Speck 2010	3.3	23.2	58	3.1	17.2	62	59.5%	0.01 [-0.35 , 0.37]
Subtotal (95% CI)			**97**			**105**	**100.0%**	**0.08 [-0.19 , 0.36]**
Heterogeneity: Tau² = 0.00; Chi² = 0.38, df = 1 (P = 0.54); I² = 0%								
Test for overall effect: Z = 0.58 (P = 0.56)								
Test for subgroup differences: Chi² = 4.51, df = 3 (P = 0.21), I² = 33.4%								



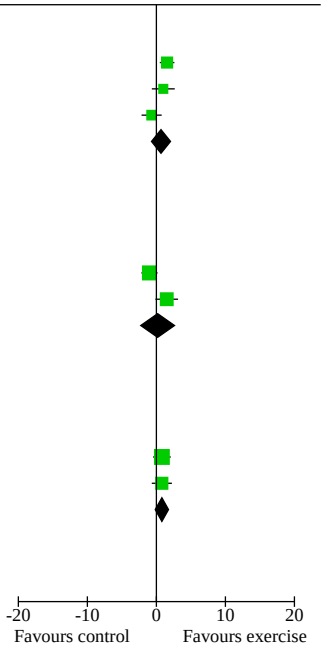
Analysis 7.2. Comparison 7: Emotional well-being/mental health, Outcome 2: Overall emotional well-being/mental health follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
7.2.1 Up to 12 weeks' follow-up								
Banasik 2011	0.47	0.36	7	0.38	0.3	7	1.3%	0.25 [-0.80 , 1.31]
Burnham 2002	-15.5	22.1	14	-13	22	7	1.8%	-0.11 [-1.02 , 0.80]
Cho 2006	52.1	7.9	28	47.3	5.5	27	4.9%	0.69 [0.15 , 1.24]
Courneya 2003a	17.09	4.66	51	16.6	4.2	45	9.0%	0.11 [-0.29 , 0.51]
Courneya 2009	70.9	16.4	60	67.2	19	62	11.5%	0.21 [-0.15 , 0.56]
Culos-Reed 2006	83.8	14.14	18	70.83	17	18	3.1%	0.81 [0.13 , 1.49]
Daley 2007a	19.06	4.18	33	18.45	3.31	33	6.2%	0.16 [-0.32 , 0.64]
Danhauer 2009	20.8	3.2	13	18.2	6.1	14	2.5%	0.51 [-0.26 , 1.28]
Dimeo 2004	69	27	34	72	26	35	6.5%	-0.11 [-0.58 , 0.36]
Donnelly 2011	30.43	8.81	16	29.65	9.3	17	3.1%	0.08 [-0.60 , 0.77]
Fillion 2008	-12.51	4.63	44	-12.64	4	43	8.2%	0.03 [-0.39 , 0.45]
Herrero 2006	90.6	9.3	8	83.3	14.8	8	1.4%	0.56 [-0.45 , 1.56]
Knols 2011	76.1	24.9	57	72.6	22	57	10.7%	0.15 [-0.22 , 0.52]
Mehnert 2011	67.5	26.92	30	62.04	27.28	27	5.3%	0.20 [-0.32 , 0.72]
Milne 2008a	19.8	2.4	29	17.2	3.5	29	5.0%	0.85 [0.32 , 1.39]
Moadel 2007	18.19	4.18	45	16.09	5.71	26	6.1%	0.43 [-0.05 , 0.92]
Pinto 2003	-10.83	28.1	12	-27.17	20	12	2.1%	0.65 [-0.18 , 1.47]
Pinto 2005	-8.02	20.69	39	-16.51	29	43	7.6%	0.33 [-0.11 , 0.77]
Rogers 2009	20	3.1	20	21.1	2.9	18	3.5%	-0.36 [-1.00 , 0.28]
Subtotal (95% CI)			**558**			**528**	**100.0%**	**0.24 [0.12 , 0.37]**
Heterogeneity: Chi² = 20.90, df = 18 (P = 0.28); I² = 14%								
Test for overall effect: Z = 3.98 (P < 0.0001)								
7.2.2 More than 12 weeks' less than 6 months' follow-up								
Bai 2004	75.19	21.3	24	69.16	16	21	6.9%	0.31 [-0.28 , 0.90]
Courneya 2003b	20.3	3.8	62	20.7	3.7	31	12.9%	-0.11 [-0.54 , 0.33]
Courneya 2003c	21.5	3.4	24	20.3	3	34	8.7%	0.37 [-0.15 , 0.90]
Daley 2007a	19.52	4.47	31	18.97	3.26	31	9.7%	0.14 [-0.36 , 0.64]
Fillion 2008	51.38	7.57	44	47.96	9.3	43	13.3%	0.40 [-0.02 , 0.82]
Knols 2011	75.6	24.2	51	76.4	19	54	16.4%	-0.04 [-0.42 , 0.35]
Milne 2008a	19.7	3	28	19.4	2.4	28	8.8%	0.11 [-0.42 , 0.63]
Tang 2010	57.87	10.7	24	44.02	12	35	7.5%	1.19 [0.62 , 1.75]
Thorsen 2005	73.9	20.1	52	78.4	23.5	49	15.7%	-0.20 [-0.60 , 0.19]
Subtotal (95% CI)			**340**			**326**	**100.0%**	**0.17 [0.01 , 0.32]**
Heterogeneity: Chi² = 20.70, df = 8 (P = 0.008); I² = 61%								
Test for overall effect: Z = 2.12 (P = 0.03)								
7.2.3 6 months' follow-up								
Cadmus 2009	20.5	3	37	19.8	4.2	37	39.5%	0.19 [-0.27 , 0.65]
Ohira 2006	-45.6	8.2	39	-48.2	8.2	40	41.8%	0.31 [-0.13 , 0.76]
Rogers 2009	19.5	3.1	19	20.8	3	17	18.8%	-0.42 [-1.08 , 0.25]
Subtotal (95% CI)			**95**			**94**	**100.0%**	**0.13 [-0.16 , 0.41]**
Heterogeneity: Chi² = 3.34, df = 2 (P = 0.19); I² = 40%								
Test for overall effect: Z = 0.87 (P = 0.38)								
7.2.4 More than 6 months' follow-up								
Speck 2010	52.2	9.5	58	54.2	8.5	62	100.0%	-0.22 [-0.58 , 0.14]
Subtotal (95% CI)			**58**			**62**	**100.0%**	**-0.22 [-0.58 , 0.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.23)								



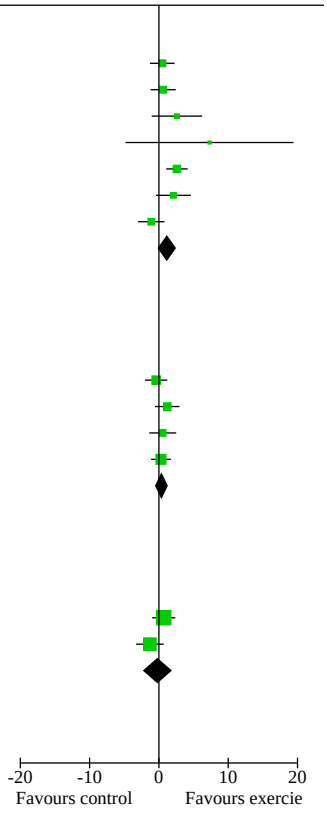
Analysis 7.3. Comparison 7: Emotional well-being/mental health, Outcome 3: FACT emotional subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.3.1 Up to 12 weeks' follow-up								
Oh 2010	1.6	2.7	54	0.05	2.9	54	38.8%	1.55 [0.49 , 2.61]
Rogers 2009	0.9	2.5	20	-0.1	2.7	18	29.0%	1.00 [-0.66 , 2.66]
Targ 2002	1.68	4.4	79	2.37	5.2	88	32.2%	-0.69 [-2.15 , 0.77]
Subtotal (95% CI)			**153**			**160**	**100.0%**	**0.67 [-0.71 , 2.05]**
Heterogeneity: Tau² = 0.98; Chi² = 6.01, df = 2 (P = 0.05); I² = 67%								
Test for overall effect: Z = 0.95 (P = 0.34)								
7.3.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	0.3	3.4	62	1.3	2.3	31	52.7%	-1.00 [-2.17 , 0.17]
Courneya 2003c	2.3	3.1	24	0.8	2.9	28	47.3%	1.50 [-0.14 , 3.14]
Subtotal (95% CI)			**86**			**59**	**100.0%**	**0.18 [-2.26 , 2.63]**
Heterogeneity: Tau² = 2.60; Chi² = 5.91, df = 1 (P = 0.02); I² = 83%								
Test for overall effect: Z = 0.15 (P = 0.88)								
7.3.3 6 months' follow-up								
Cadmus 2009	0.3	2.2	37	-0.5	3.3	37	56.6%	0.80 [-0.48 , 2.08]
Rogers 2009	0.5	2.7	19	-0.29	1.7	17	43.4%	0.79 [-0.67 , 2.25]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.80 [-0.17 , 1.76]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.99); I² = 0%								
Test for overall effect: Z = 1.62 (P = 0.10)								
Test for subgroup differences: Chi² = 0.21, df = 2 (P = 0.90), I² = 0%								



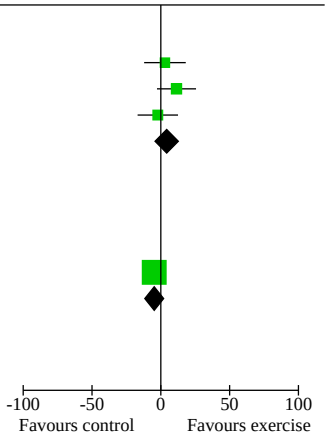
Analysis 7.4. Comparison 7: Emotional well-being/mental health, Outcome 4: FACT emotional subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.4.1 Up to 12 weeks' follow-up								
Courneya 2003a	17.09	4.66	51	16.6	4.2	45	19.2%	0.49 [-1.28 , 2.26]
Daley 2007a	19.06	4.18	33	18.45	3.31	33	18.7%	0.61 [-1.21 , 2.43]
Danhauer 2009	20.8	3.2	13	18.2	6.1	14	8.3%	2.60 [-1.04 , 6.24]
Herrero 2006	90.6	9.3	8	83.3	14.8	8	1.0%	7.30 [-4.81 , 19.41]
Milne 2008a	19.8	2.4	29	17.2	3.5	29	21.2%	2.60 [1.06 , 4.14]
Moadel 2007	18.19	4.18	45	16.09	5.71	26	13.6%	2.10 [-0.41 , 4.61]
Rogers 2009	20	3.1	20	21.1	2.9	18	18.0%	-1.10 [-3.01 , 0.81]
Subtotal (95% CI)			**199**			**173**	**100.0%**	**1.14 [-0.08 , 2.35]**
Heterogeneity: Tau² = 1.19; Chi² = 11.73, df = 6 (P = 0.07); I² = 49%								
Test for overall effect: Z = 1.83 (P = 0.07)								
7.4.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	20.3	3.8	62	20.7	3.7	31	26.3%	-0.40 [-2.01 , 1.21]
Courneya 2003c	21.5	3.4	24	20.3	3	28	22.1%	1.20 [-0.56 , 2.96]
Daley 2007a	19.52	4.47	31	18.97	3.26	31	18.0%	0.55 [-1.40 , 2.50]
Milne 2008a	19.7	3	28	19.4	2.4	28	33.6%	0.30 [-1.12 , 1.72]
Subtotal (95% CI)			**145**			**118**	**100.0%**	**0.36 [-0.47 , 1.19]**
Heterogeneity: Tau² = 0.00; Chi² = 1.78, df = 3 (P = 0.62); I² = 0%								
Test for overall effect: Z = 0.85 (P = 0.39)								
7.4.3 6 months' follow-up								
Cadmus 2009	20.5	3	37	19.8	4.2	37	53.9%	0.70 [-0.96 , 2.36]
Rogers 2009	19.5	3.1	19	20.8	3	17	46.1%	-1.30 [-3.29 , 0.69]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**-0.22 [-2.18 , 1.73]**
Heterogeneity: Tau² = 1.12; Chi² = 2.28, df = 1 (P = 0.13); I² = 56%								
Test for overall effect: Z = 0.22 (P = 0.82)								
Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43), I² = 0%								



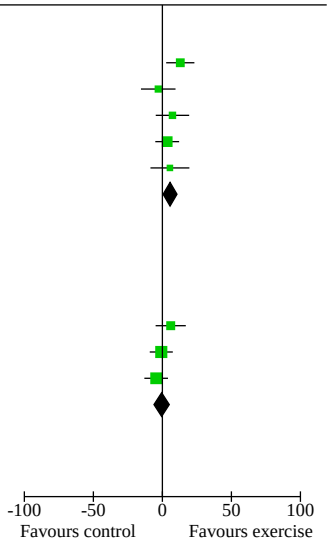
Analysis 7.5. Comparison 7: Emotional well-being/mental health, Outcome 5: QLQ-C30 subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.5.1 Up to 12 weeks' follow-up								
Dimeo 2004	11	30	34	8	34	35	31.3%	3.00 [-12.12 , 18.12]
Herrero 2006	9.3	19.6	8	-2.1	5.9	8	35.5%	11.40 [-2.78 , 25.58]
Mehnert 2011	3.06	27.13	30	5.25	29.2	27	33.2%	-2.19 [-16.87 , 12.49]
Subtotal (95% CI)			**72**			**70**	**100.0%**	**4.26 [-4.19 , 12.72]**
Heterogeneity: Tau² = 0.00; Chi² = 1.74, df = 2 (P = 0.42); I² = 0%								
Test for overall effect: Z = 0.99 (P = 0.32)								
7.5.2 More than 12 weeks' less than 6 months' follow-up								
Thorsen 2005	1	18.4672	52	5.8	16.5828	49	100.0%	-4.80 [-11.64 , 2.04]
Subtotal (95% CI)			**52**			**49**	**100.0%**	**-4.80 [-11.64 , 2.04]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.38 (P = 0.17)								
Test for subgroup differences: Chi² = 2.67, df = 1 (P = 0.10), I² = 62.5%								



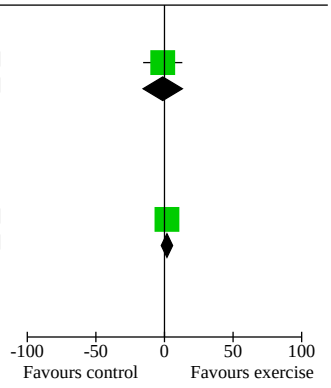
Analysis 7.6. Comparison 7: Emotional well-being/mental health, Outcome 6: QLQ-C30 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.6.1 Up to 12 weeks' follow-up								
Culos-Reed 2006	83.8	14.14	18	70.83	17	18	23.2%	12.97 [2.75 , 23.19]
Dimeo 2004	69	27	34	72	26	35	15.6%	-3.00 [-15.51 , 9.51]
Herrero 2006	90.6	9.3	8	83.3	14.8	8	16.7%	7.30 [-4.81 , 19.41]
Knols 2011	76.1	24.9	57	72.6	22	57	32.1%	3.50 [-5.13 , 12.13]
Mehnert 2011	67.5	26.92	30	62.04	27.28	27	12.4%	5.46 [-8.64 , 19.56]
Subtotal (95% CI)			**147**			**145**	**100.0%**	**5.56 [0.55 , 10.56]**
Heterogeneity: Tau² = 0.96; Chi² = 4.12, df = 4 (P = 0.39); I² = 3%								
Test for overall effect: Z = 2.18 (P = 0.03)								
7.6.2 More than 12 weeks' less than 6 months' follow-up								
Bai 2004	75.19	21.3	24	69.16	16	21	23.9%	6.03 [-4.90 , 16.96]
Knols 2011	75.6	24.2	51	76.4	19	54	38.9%	-0.80 [-9.15 , 7.55]
Thorsen 2005	73.9	20.1	52	78.4	23.5	49	37.3%	-4.50 [-13.05 , 4.05]
Subtotal (95% CI)			**127**			**124**	**100.0%**	**-0.55 [-6.09 , 4.99]**
Heterogeneity: Tau² = 2.36; Chi² = 2.21, df = 2 (P = 0.33); I² = 10%								
Test for overall effect: Z = 0.19 (P = 0.85)								
Test for subgroup differences: Chi² = 2.57, df = 1 (P = 0.11), I² = 61.1%								



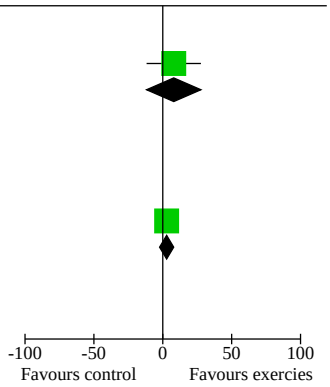
Analysis 7.7. Comparison 7: Emotional well-being/mental health, Outcome 7: MOS SF-36 emotional role subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.7.1 Up to 12 weeks' follow-up								
Mehnert 2011	2.3	35.5	30	3.57	17.6	28	100.0%	-1.27 [-15.55 , 13.01]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**-1.27 [-15.55 , 13.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.17 (P = 0.86)								
7.7.2 6 months' follow-up								
Cadmus 2009	0.9	12	37	-0.9	4.4	37	100.0%	1.80 [-2.32 , 5.92]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**1.80 [-2.32 , 5.92]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.86 (P = 0.39)								
Test for subgroup differences: Chi² = 0.16, df = 1 (P = 0.69), I² = 0%								



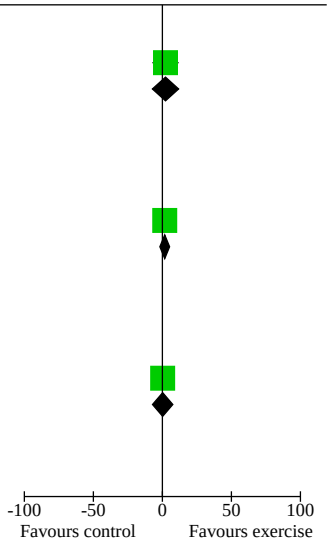
Analysis 7.8. Comparison 7: Emotional well-being/mental health, Outcome 8: MOS SF-36 emotional role subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.8.1 Up to 12 weeks' follow-up								
Mehnert 2011	77.01	35.75	30	69.05	40.5	28	100.0%	7.96 [-11.76 , 27.68]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**7.96 [-11.76 , 27.68]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.79 (P = 0.43)								
7.8.2 6 months' follow-up								
Cadmus 2009	50.2	10.4	37	47.4	12	37	100.0%	2.80 [-2.32 , 7.92]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**2.80 [-2.32 , 7.92]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 0.25, df = 1 (P = 0.62), I² = 0%								



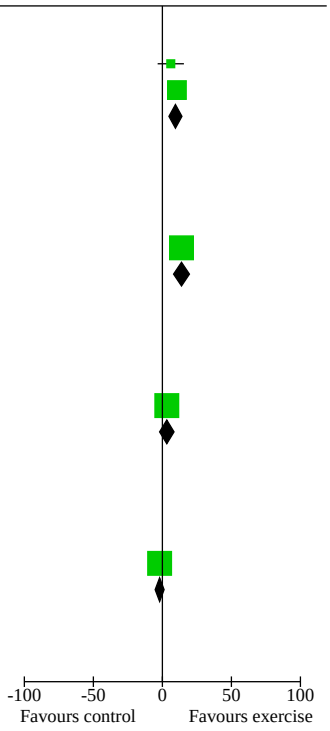
Analysis 7.9. Comparison 7: Emotional well-being/mental health, Outcome 9: MOS SF-36 mental health component change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.9.1 Up to 12 weeks' follow-up								
Mehnert 2011	4.8	17.64	30	2.54	19.1	28	100.0%	2.26 [-7.22 , 11.74]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**2.26 [-7.22 , 11.74]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.47 (P = 0.64)								
7.9.2 6 months' follow-up								
Cadmus 2009	0.8	6.6	37	-0.9	8.7	37	100.0%	1.70 [-1.82 , 5.22]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**1.70 [-1.82 , 5.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.95 (P = 0.34)								
7.9.3 More than 6 months' follow-up								
Speck 2010	3.3	23.2	58	3.1	17.2	62	100.0%	0.20 [-7.15 , 7.55]
Subtotal (95% CI)			**58**			**62**	**100.0%**	**0.20 [-7.15 , 7.55]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.05 (P = 0.96)								
Test for subgroup differences: Chi² = 0.16, df = 2 (P = 0.92), I² = 0%								



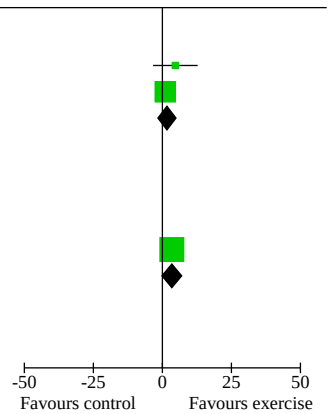
Analysis 7.10. Comparison 7: Emotional well-being/mental health, Outcome 10: MOS SF-36 mental health component follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.10.1 Up to 12 weeks' follow-up								
Mehnert 2011	69.33	18.3	30	63.25	18.61	28	24.5%	6.08 [-3.43 , 15.59]
Tang 2010	55.47	10.07	36	44.92	13	35	75.5%	10.55 [5.13 , 15.97]
Subtotal (95% CI)			**66**			**63**	**100.0%**	**9.45 [4.75 , 14.16]**
Heterogeneity: Tau² = 0.00; Chi² = 0.64, df = 1 (P = 0.42); I² = 0%								
Test for overall effect: Z = 3.94 (P < 0.0001)								
7.10.2 More than 12 weeks' less than 6 months' follow-up								
Tang 2010	57.87	10.68	24	44.02	11.75	36	100.0%	13.85 [8.11 , 19.59]
Subtotal (95% CI)			**24**			**36**	**100.0%**	**13.85 [8.11 , 19.59]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.73 (P < 0.00001)								
7.10.3 6 months' follow-up								
Cadmus 2009	50.6	10.9	37	47.4	12	37	100.0%	3.20 [-2.02 , 8.42]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**3.20 [-2.02 , 8.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.20 (P = 0.23)								
7.10.4 More than 6 months' follow-up								
Speck 2010	52.2	9.5	58	54.2	8.5	62	100.0%	-2.00 [-5.23 , 1.23]
Subtotal (95% CI)			**58**			**62**	**100.0%**	**-2.00 [-5.23 , 1.23]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.23)								
Test for subgroup differences: Chi² = 29.71, df = 3 (P < 0.00001), I² = 89.9%								



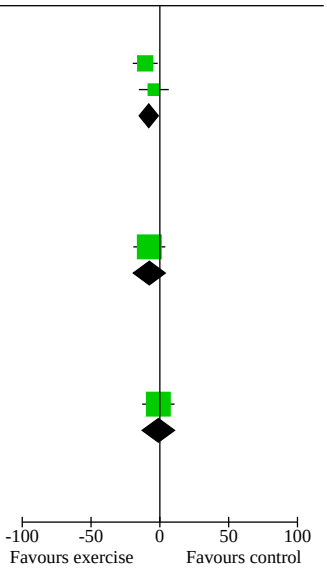
Analysis 7.11. Comparison 7: Emotional well-being/mental health, Outcome 11: MOS SF-12 mental health component follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.11.1 Up to 12 weeks' follow-up								
Danhauer 2009	52.2	6.6	13	47.5	13.8	14	16.5%	4.70 [-3.37 , 12.77]
Fillion 2008	48.54	7.91	44	47.49	9.08	43	83.5%	1.05 [-2.53 , 4.63]
Subtotal (95% CI)			**57**			**57**	**100.0%**	**1.65 [-1.62 , 4.92]**
Heterogeneity: Tau² = 0.00; Chi² = 0.66, df = 1 (P = 0.42); I² = 0%								
Test for overall effect: Z = 0.99 (P = 0.32)								
7.11.2 More than 12 weeks' less than 6 months' follow-up								
Fillion 2008	51.38	7.57	44	47.96	9.3	43	100.0%	3.42 [-0.15 , 6.99]
Subtotal (95% CI)			**44**			**43**	**100.0%**	**3.42 [-0.15 , 6.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.88 (P = 0.06)								
Test for subgroup differences: Chi² = 0.51, df = 1 (P = 0.47), I² = 0%								



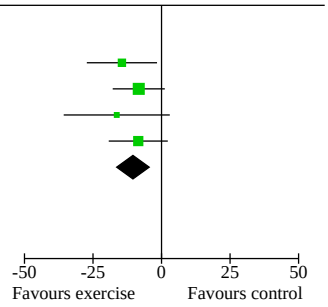
Analysis 7.12. Comparison 7: Emotional well-being/mental health, Outcome 12: POMS total mood disturbance change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.12.1 Up to 12 weeks' follow-up								
Oh 2010	-8.73	21.6	54	1.91	26.2	54	59.0%	-10.64 [-19.70 , -1.58]
Targ 2002	-12.34	36.7	79	-7.95	34.7	88	41.0%	-4.39 [-15.26 , 6.48]
Subtotal (95% CI)			**133**			**142**	**100.0%**	**-8.08 [-15.03 , -1.12]**
Heterogeneity: Tau² = 0.00; Chi² = 0.75, df = 1 (P = 0.39); I² = 0%								
Test for overall effect: Z = 2.28 (P = 0.02)								
7.12.2 6 months' follow-up								
Pinto 2005	3.8	27.7	39	11.47	25.8	43	100.0%	-7.67 [-19.29 , 3.95]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**-7.67 [-19.29 , 3.95]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.29 (P = 0.20)								
7.12.3 More than 6 months' follow-up								
Pinto 2005	-3.09	27.9	39	-2	26.2	43	100.0%	-1.09 [-12.84 , 10.66]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**-1.09 [-12.84 , 10.66]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.18 (P = 0.86)								
Test for subgroup differences: Chi² = 1.05, df = 2 (P = 0.59), I² = 0%								



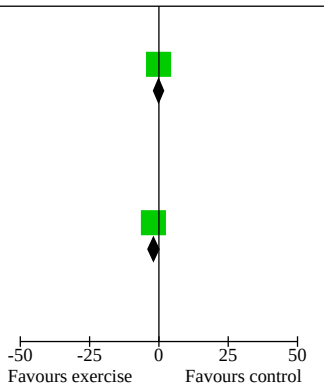
Analysis 7.13. Comparison 7: Emotional well-being/mental health, Outcome 13: POMS total mood disturbance follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.13.1 Up to 12 weeks' follow-up								
Culos-Reed 2006	-6.78	16.75	18	7.67	22.09	18	21.4%	-14.45 [-27.26 , -1.64]
Moadel 2007	18.26	16.26	45	26.57	21.41	26	38.9%	-8.31 [-17.81 , 1.19]
Pinto 2003	10.83	28.1	12	27.17	19.5	12	9.4%	-16.34 [-35.69 , 3.01]
Pinto 2005	8.02	20.69	39	16.51	28.75	43	30.3%	-8.49 [-19.26 , 2.28]
Subtotal (95% CI)			**114**			**99**	**100.0%**	**-10.43 [-16.36 , -4.51]**
Heterogeneity: Tau² = 0.00; Chi² = 1.05, df = 3 (P = 0.79); I² = 0%								
Test for overall effect: Z = 3.45 (P = 0.0006)								
Test for subgroup differences: Not applicable								



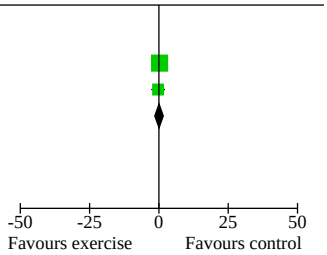
Analysis 7.14. Comparison 7: Emotional well-being/mental health, Outcome 14: POMS - anxiety and depression subscales follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.14.1 Up to 12 weeks' follow-up								
Fillion 2008	12.51	4.63	44	12.64	4.02	43	100.0%	-0.13 [-1.95 , 1.69]
Subtotal (95% CI)			**44**			**43**	**100.0%**	**-0.13 [-1.95 , 1.69]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.14 (P = 0.89)								
7.14.2 More than 12 weeks' less than 6 months' follow-up								
Fillion 2008	11.15	3.9	44	13.13	5.44	43	100.0%	-1.98 [-3.97 , 0.01]
Subtotal (95% CI)			**44**			**43**	**100.0%**	**-1.98 [-3.97 , 0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.95 (P = 0.05)								
Test for subgroup differences: Chi² = 1.80, df = 1 (P = 0.18), I² = 44.6%								



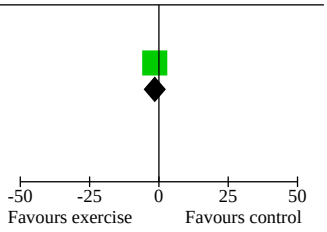
Analysis 7.15. Comparison 7: Emotional well-being/mental health, Outcome 15: POMS anger subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.15.1 Up to 12 weeks' follow-up								
Oh 2010	-0.05	4.6	54	-0.3	5.6	54	62.3%	0.25 [-1.68 , 2.18]
Targ 2002	-1.3	8.9	79	-1.01	7.3	88	37.7%	-0.29 [-2.78 , 2.20]
Subtotal (95% CI)			**133**			**142**	**100.0%**	**0.05 [-1.48 , 1.57]**
Heterogeneity: Tau² = 0.00; Chi² = 0.11, df = 1 (P = 0.74); I² = 0%								
Test for overall effect: Z = 0.06 (P = 0.95)								
Test for subgroup differences: Not applicable								



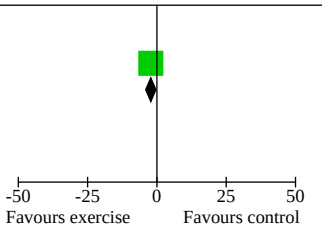
Analysis 7.16. Comparison 7: Emotional well-being/mental health, Outcome 16: POMS anger subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.16.1 Up to 12 weeks' follow-up								
Pinto 2003	3.33	4.1	12	4.83	3.6	6	100.0%	-1.50 [-5.20 , 2.20]
Subtotal (95% CI)			**12**			**6**	**100.0%**	**-1.50 [-5.20 , 2.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.79 (P = 0.43)								
Test for subgroup differences: Not applicable								



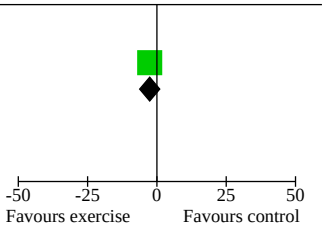
Analysis 7.17. Comparison 7: Emotional well-being/mental health, Outcome 17: CARES subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.17.1 6 months' follow-up								
Ohira 2006	-2.5	4.4	39	-0.3	3.9	40	100.0%	-2.20 [-4.04 , -0.36]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-2.20 [-4.04 , -0.36]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.35 (P = 0.02)								
Test for subgroup differences: Not applicable								



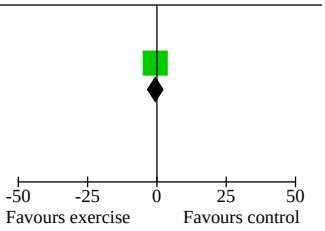
Analysis 7.18. Comparison 7: Emotional well-being/mental health, Outcome 18: CARES subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.18.1 6 months' follow-up								
Ohira 2006	45.6	8.2	39	48.2	8.2	40	100.0%	-2.60 [-6.22 , 1.02]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-2.60 [-6.22 , 1.02]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.41 (P = 0.16)								
Test for subgroup differences: Not applicable								



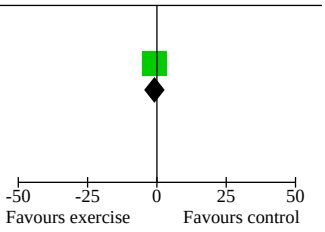
Analysis 7.19. Comparison 7: Emotional well-being/mental health, Outcome 19: Cohen's perceived stress scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.19.1 6 months' follow-up								
Cadmus 2009	0.2	5.7	37	0.8	6.4	37	100.0%	-0.60 [-3.36 , 2.16]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.60 [-3.36 , 2.16]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.43 (P = 0.67)								
Test for subgroup differences: Not applicable								



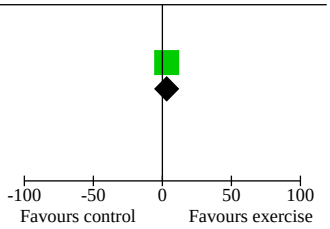
Analysis 7.20. Comparison 7: Emotional well-being/mental health, Outcome 20: Cohen's perceived stress scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.20.1 6 months' follow-up								
Cadmus 2009	12.9	6.2	37	13.8	8.5	37	100.0%	-0.90 [-4.29 , 2.49]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.90 [-4.29 , 2.49]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.52 (P = 0.60)								
Test for subgroup differences: Not applicable								



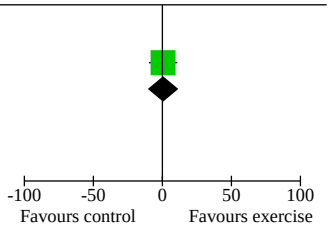
Analysis 7.21. Comparison 7: Emotional well-being/mental health, Outcome 21: Fordyce change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.21.1 6 months' follow-up								
Cadmus 2009	3	18	37	-0.1	19.1	37	100.0%	3.10 [-5.36 , 11.56]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**3.10 [-5.36 , 11.56]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.72 (P = 0.47)								
Test for subgroup differences: Not applicable								



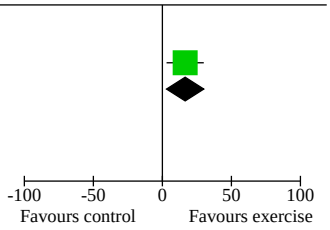
Analysis 7.22. Comparison 7: Emotional well-being/mental health, Outcome 22: Fordyce follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.22.1 6 months' follow-up								
Cadmus 2009	68.6	22.4	37	68.1	22.7	37	100.0%	0.50 [-9.78 , 10.78]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.50 [-9.78 , 10.78]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.10 (P = 0.92)								
Test for subgroup differences: Not applicable								



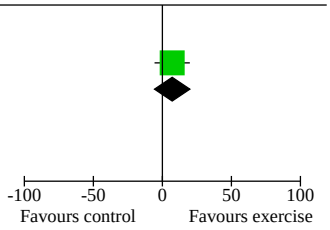
Analysis 7.23. Comparison 7: Emotional well-being/mental health, Outcome 23: Happiness change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.23.1 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	17.3	26.1	24	0.8	23	28	100.0%	16.50 [3.02 , 29.98]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**16.50 [3.02 , 29.98]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.40 (P = 0.02)								
Test for subgroup differences: Not applicable								



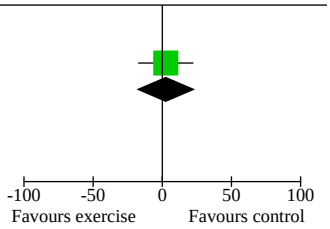
Analysis 7.24. Comparison 7: Emotional well-being/mental health, Outcome 24: Happiness follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.24.1 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	65.4	24.6	24	58.3	22	28	100.0%	7.10 [-5.68 , 19.88]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**7.10 [-5.68 , 19.88]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.09 (P = 0.28)								
Test for subgroup differences: Not applicable								



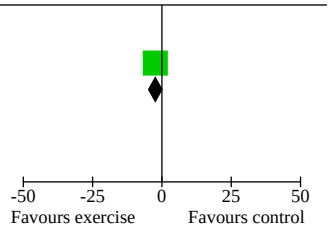
Analysis 7.25. Comparison 7: Emotional well-being/mental health, Outcome 25: Linear Analog Self-Assessment Scale - anger follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.25.1 Up to 12 weeks' follow-up								
Burnham 2002	15.5	22.1	14	13	22	7	100.0%	2.50 [-17.49 , 22.49]
Subtotal (95% CI)			**14**			**7**	**100.0%**	**2.50 [-17.49 , 22.49]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.25 (P = 0.81)								
Test for subgroup differences: Not applicable								



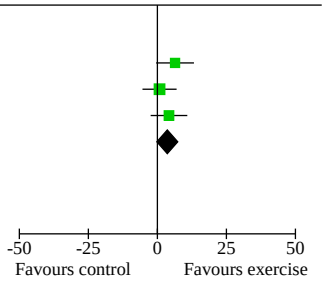
Analysis 7.26. Comparison 7: Emotional well-being/mental health, Outcome 26: Symptoms of Stress Index - emotional irritability subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.26.1 Up to 12 weeks' follow-up								
Culos-Reed 2006	1.83	1.72	18	4.22	4.9	18	100.0%	-2.39 [-4.79 , 0.01]
Subtotal (95% CI)			**18**			**18**	**100.0%**	**-2.39 [-4.79 , 0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.95 (P = 0.05)								
Test for subgroup differences: Not applicable								



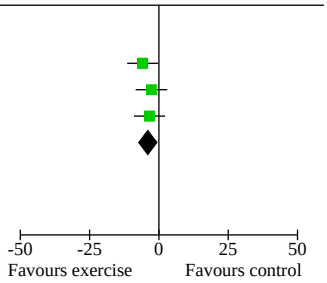
Analysis 7.27. Comparison 7: Emotional well-being/mental health, Outcome 27: PANAS - positivity follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.27.1 Up to 12 weeks' follow-up								
Danhauer 2009	38.2	6.8	13	31.8	11	14	30.4%	6.40 [-0.45 , 13.25]
Donnelly 2011	30.43	8.81	16	29.65	9.3	17	37.3%	0.78 [-5.40 , 6.96]
Pinto 2003	35	8.1	12	30.8	6	6	32.3%	4.20 [-2.44 , 10.84]
Subtotal (95% CI)			**41**			**37**	**100.0%**	**3.59 [-0.18 , 7.37]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.47$, df = 2 (P = 0.48); $I^2 = 0\%$								
Test for overall effect: Z = 1.87 (P = 0.06)								
Test for subgroup differences: Not applicable								



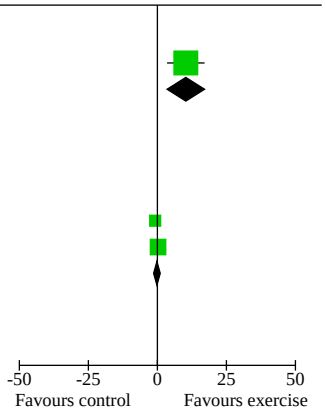
Analysis 7.28. Comparison 7: Emotional well-being/mental health, Outcome 28: PANAS - negativity follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.28.1 Up to 12 weeks' follow-up								
Danhauer 2009	14	3.9	13	19.9	9.8	14	34.1%	-5.90 [-11.45 , -0.35]
Donnelly 2011	15.44	6.46	16	18.12	10	17	32.2%	-2.68 [-8.39 , 3.03]
Pinto 2003	14.42	5.9	12	17.8	5.6	6	33.7%	-3.38 [-8.97 , 2.21]
Subtotal (95% CI)			**41**			**37**	**100.0%**	**-4.01 [-7.26 , -0.77]**
Heterogeneity: Tau² = 0.00; Chi² = 0.70, df = 2 (P = 0.70); I² = 0%								
Test for overall effect: Z = 2.43 (P = 0.02)								
Test for subgroup differences: Not applicable								



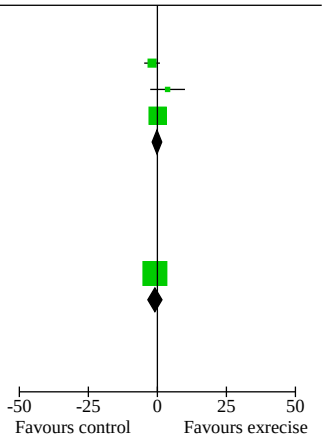
Analysis 7.29. Comparison 7: Emotional well-being/mental health, Outcome 29: Satisfaction with Life Scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.29.1 Up to 12 weeks' follow-up								
Courneya 2009	11	19.7	60	0.7	18.5	62	100.0%	10.30 [3.51 , 17.09]
Subtotal (95% CI)			**60**			**62**	**100.0%**	**10.30 [3.51 , 17.09]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.97 (P = 0.003)								
7.29.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	0.7	5	62	1.5	4.1	31	39.4%	-0.80 [-2.71 , 1.11]
Daley 2007a	12.56	2.7	31	12.32	3.08	25	60.6%	0.24 [-1.30 , 1.78]
Subtotal (95% CI)			**93**			**56**	**100.0%**	**-0.17 [-1.37 , 1.03]**
Heterogeneity: Tau² = 0.00; Chi² = 0.69, df = 1 (P = 0.41); I² = 0%								
Test for overall effect: Z = 0.28 (P = 0.78)								
Test for subgroup differences: Chi² = 8.87, df = 1 (P = 0.003), I² = 88.7%								

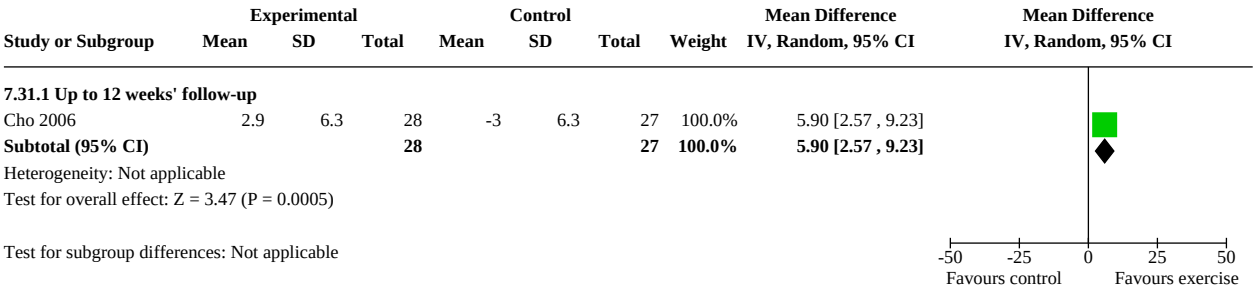


Analysis 7.30. Comparison 7: Emotional well-being/mental health, Outcome 30: Satisfaction with Life Scale follow-up values

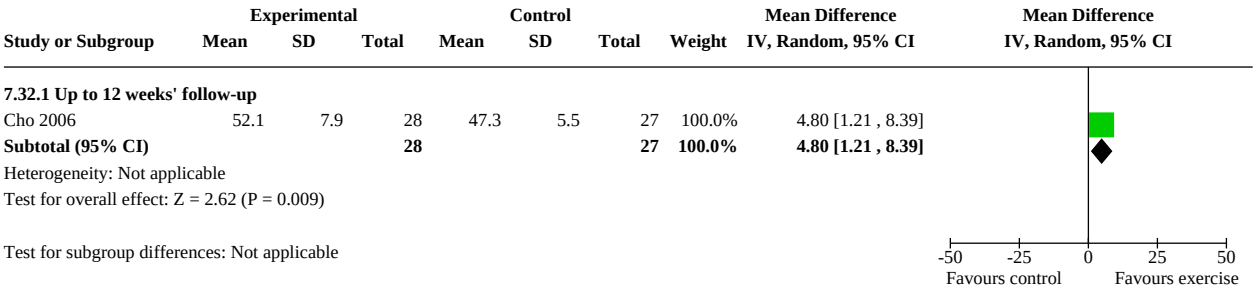
Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.30.1 Up to 12 weeks' follow-up								
Courneya 2003a	20.76	7.59	51	22.68	6.68	45	24.7%	-1.92 [-4.77 , 0.93]
Courneya 2009	70.9	16.4	60	67.2	19	62	6.9%	3.70 [-2.59 , 9.99]
Daley 2007a	5.21	1.27	33	5.12	1.27	33	68.4%	0.09 [-0.52 , 0.70]
Subtotal (95% CI)			**144**			**140**	**100.0%**	**-0.16 [-1.89 , 1.58]**
Heterogeneity: Tau² = 1.05; Chi² = 3.14, df = 2 (P = 0.21); I² = 36%								
Test for overall effect: Z = 0.18 (P = 0.86)								
7.30.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	26.9	6	62	27.8	5.9	31	100.0%	-0.90 [-3.46 , 1.66]
Subtotal (95% CI)			**62**			**31**	**100.0%**	**-0.90 [-3.46 , 1.66]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.69 (P = 0.49)								
Test for subgroup differences: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%								



Analysis 7.31. Comparison 7: Emotional well-being/mental health, Outcome 31: Lee Psychosocial Adjustment instrument change



Analysis 7.32. Comparison 7: Emotional well-being/mental health, Outcome 32: Lee Psychosocial Adjustment Instrument follow-up values



Comparison 8. Fatigue

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.1 Overall fatigue change	14		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
8.1.1 Up to 12 weeks' follow-up	10	745	Std. Mean Difference (IV, Random, 95% CI)	-0.82 [-1.50, -0.14]
8.1.2 More than 12 weeks' less than 6 months' follow-up	3	246	Std. Mean Difference (IV, Random, 95% CI)	-0.42 [-0.83, -0.02]
8.1.3 6 months' follow-up	3	514	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.31, 0.19]
8.1.4 More than 6 months' follow-up	1	82	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.82, 0.06]
8.2 Overall fatigue follow-up values	22		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
8.2.1 Up to 12 weeks' follow-up	18	994	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.46, -0.14]
8.2.2 More than 12 weeks' less than 6 months' follow-up	5	436	Std. Mean Difference (IV, Random, 95% CI)	-0.03 [-0.31, 0.25]
8.2.3 6 months' follow-up	2	110	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.48, 0.27]

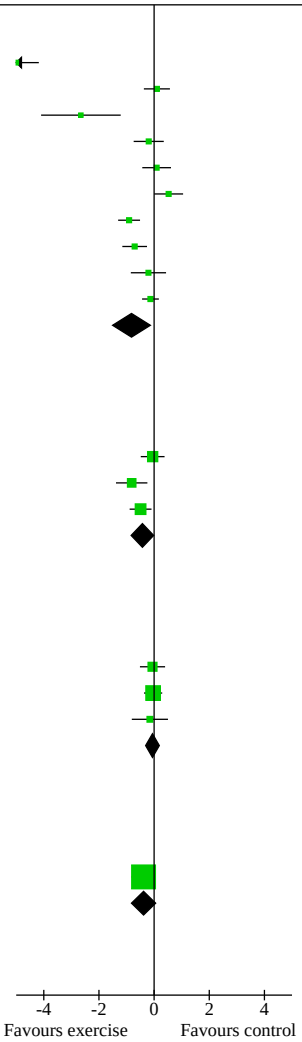
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.3 FACT fatigue subscale change	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.3.1 Up to 12 weeks' follow-up	4	296	Mean Difference (IV, Random, 95% CI)	4.33 [2.43, 6.22]
8.3.2 More than 12 weeks' less than 6 months' follow-up	2	145	Mean Difference (IV, Random, 95% CI)	-3.97 [-10.53, 2.60]
8.3.3 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	1.28 [-4.11, 6.67]
8.4 FACT fatigue subscale follow-up values	11		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.4.1 Up to 12 weeks' follow-up	9	550	Mean Difference (IV, Random, 95% CI)	2.00 [0.00, 3.99]
8.4.2 More than 12 weeks' up to 6 months' follow-up	3	250	Mean Difference (IV, Random, 95% CI)	0.21 [-2.04, 2.45]
8.4.3 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	0.30 [-7.22, 7.82]
8.5 QLQ-C30 fatigue subscale change	4	259	Mean Difference (IV, Random, 95% CI)	-15.38 [-39.15, 8.38]
8.5.1 Up to 12 weeks' follow-up	3	158	Mean Difference (IV, Random, 95% CI)	-22.45 [-50.66, 5.77]
8.5.2 More than 12 weeks' less than 6 months' follow-up	1	101	Mean Difference (IV, Random, 95% CI)	11.20 [2.36, 20.04]
8.6 QLQ-C30 fatigue subscale follow-up values	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.6.1 Up to 12 weeks' follow-up	4	256	Mean Difference (IV, Random, 95% CI)	-7.50 [-12.92, -2.07]
8.6.2 More than 12 weeks' less than 6 months' follow-up	2	206	Mean Difference (IV, Random, 95% CI)	3.03 [-6.67, 12.73]
8.7 Multidimensional Fatigue Inventory follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.7.1 Up to 12 weeks' follow-up	1	87	Mean Difference (IV, Random, 95% CI)	-0.21 [-0.58, 0.16]
8.8 MOS SF-36 vitality subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.8.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	2.90 [-6.33, 12.13]
8.8.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	0.40 [-2.72, 3.52]
8.9 MOS SF-36 vitality subscale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.9.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	8.09 [-1.39, 17.57]
8.9.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	1.30 [-3.03, 5.63]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.10 Piper Revised Fatigue Scale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.10.1 Up to 12 weeks' follow-up	1	66	Mean Difference (IV, Random, 95% CI)	-1.30 [-2.17, -0.43]
8.10.2 More than 12 weeks' less than 6 months' follow-up	1	59	Mean Difference (IV, Random, 95% CI)	-0.78 [-1.83, 0.27]
8.11 Schwartz Cancer Fatigue scale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.11.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-2.20 [-4.32, -0.08]
8.12 Linear Analog Self-Assessment energy scale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.12.1 Up to 12 weeks' follow-up	1	21	Mean Difference (IV, Random, 95% CI)	-19.00 [-31.89, -6.11]
8.13 Linear Analog Self-Assessment scale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.13.1 Up to 12 weeks' follow-up	1	82	Mean Difference (IV, Random, 95% CI)	-13.14 [-23.32, -2.96]
8.13.2 6 months' follow-up	1	404	Mean Difference (IV, Random, 95% CI)	-2.44 [-13.06, 8.18]
8.13.3 More than 6 months' follow-up	1	82	Mean Difference (IV, Random, 95% CI)	-9.47 [-20.13, 1.19]
8.14 Linear Analog Self-Assessment scale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.14.1 Up to 12 weeks' follow-up	2	103	Mean Difference (IV, Random, 95% CI)	-15.40 [-25.08, -5.73]
8.15 Schwartz Cancer Fatigue scale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.15.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-4.00 [-6.03, -1.97]
8.16 POMS fatigue subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.16.1 Up to 12 weeks' follow-up	2	275	Mean Difference (IV, Random, 95% CI)	-0.17 [-2.25, 1.92]
8.17 POMS - fatigue subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.17.1 Up to 12 weeks' follow-up	1	24	Mean Difference (IV, Random, 95% CI)	-1.84 [-6.96, 3.28]
8.18 POMS vigor subscale change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.18.1 Up to 12 weeks' follow-up	2	275	Mean Difference (IV, Random, 95% CI)	2.37 [-1.70, 6.44]
8.18.2 6 months' follow-up	1	82	Mean Difference (IV, Random, 95% CI)	2.23 [-0.48, 4.94]
8.18.3 More than 6 months' follow-up	1	82	Mean Difference (IV, Random, 95% CI)	0.84 [-1.89, 3.57]
8.19 POMS vigor subscale follow-up values	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
8.19.1 Up to 12 weeks' follow-up	1	82	Mean Difference (IV, Fixed, 95% CI)	4.77 [2.36, 7.18]
8.20 POMS short form vigor follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.20.1 Up to 12 weeks' follow-up	1	86	Mean Difference (IV, Random, 95% CI)	0.27 [-0.07, 0.61]
8.20.2 More than 12 weeks' less than 6 months' follow-up	1	85	Mean Difference (IV, Random, 95% CI)	0.39 [0.05, 0.73]
8.21 Brief Fatigue Inventory follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.21.1 Up to 12 weeks' follow-up	1	30	Mean Difference (IV, Random, 95% CI)	0.00 [-1.08, 1.08]

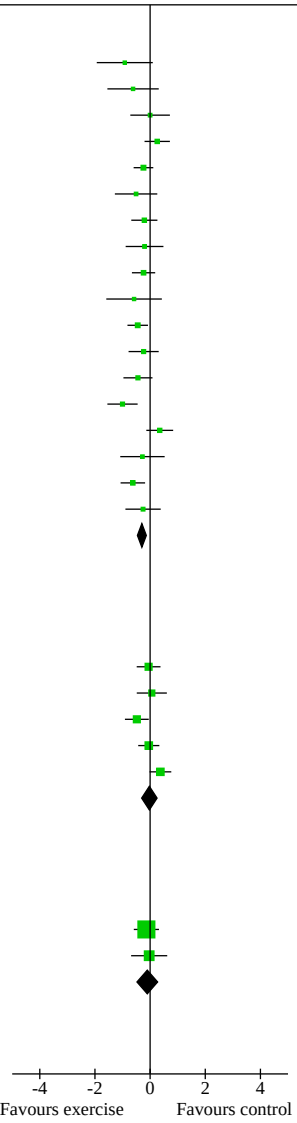
Analysis 8.1. Comparison 8: Fatigue, Outcome 1: Overall fatigue change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
8.1.1 Up to 12 weeks' follow-up								
Courneya 2009	-45	10.1	60	0.1	6.7	38	9.5%	-5.00 [-5.83 , -4.18]
Dimeo 2004	9	25	34	6	33	35	10.4%	0.10 [-0.37 , 0.57]
Herrero 2006	-42	19.6	8	6.4	14.5	8	7.4%	-2.65 [-4.10 , -1.21]
McNeely 2008a	-3.1	9	27	-1.6	5.6	25	10.3%	-0.20 [-0.74 , 0.35]
Mehnert 2011	-0.74	20.73	30	-2.88	25.3	27	10.3%	0.09 [-0.43 , 0.61]
Milne 2008a	2.7	3.7	29	0.5	4.5	29	10.3%	0.53 [0.00 , 1.05]
Oh 2010	-6.34	7.2	54	-0.64	5.1	54	10.6%	-0.91 [-1.30 , -0.51]
Pinto 2005	-14.93	23.5	39	1.79	23.5	43	10.5%	-0.70 [-1.15 , -0.26]
Rogers 2009	1.9	9.6	20	4.2	12.3	18	10.0%	-0.21 [-0.84 , 0.43]
Targ 2002	-1.27	7.054	79	-0.25	8.2	88	10.7%	-0.13 [-0.44 , 0.17]
Subtotal (95% CI)			**380**			**365**	**100.0%**	**-0.82 [-1.50 , -0.14]**
Heterogeneity: Tau² = 1.11; Chi² = 161.32, df = 9 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 2.35 (P = 0.02)								
8.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-0.4	11	62	0.2	12	31	35.2%	-0.05 [-0.48 , 0.38]
Courneya 2003c	-9.3	10.2	24	-2	7.5	28	27.3%	-0.81 [-1.38 , -0.24]
Thorsen 2005	5.8	22.3	52	17	23	49	37.5%	-0.49 [-0.89 , -0.09]
Subtotal (95% CI)			**138**			**108**	**100.0%**	**-0.42 [-0.83 , -0.02]**
Heterogeneity: Tau² = 0.07; Chi² = 4.72, df = 2 (P = 0.09); I² = 58%								
Test for overall effect: Z = 2.04 (P = 0.04)								
8.1.3 6 months' follow-up								
Cadmus 2009	-1.6	6.6	37	-1.2	7.1	37	29.5%	-0.06 [-0.51 , 0.40]
Pinto 2005	-10.1	25.1	39	-7.66	69.7	365	56.2%	-0.04 [-0.37 , 0.29]
Rogers 2009	-0.08	7.8	19	1.2	8.6	17	14.3%	-0.15 [-0.81 , 0.50]
Subtotal (95% CI)			**95**			**419**	**100.0%**	**-0.06 [-0.31 , 0.19]**
Heterogeneity: Tau² = 0.00; Chi² = 0.10, df = 2 (P = 0.95); I² = 0%								
Test for overall effect: Z = 0.47 (P = 0.64)								
8.1.4 More than 6 months' follow-up								
Pinto 2005	-13.31	25	39	-3.84	24.1	43	100.0%	-0.38 [-0.82 , 0.06]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**-0.38 [-0.82 , 0.06]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.71 (P = 0.09)								
Test for subgroup differences: Chi² = 6.00, df = 3 (P = 0.11), I² = 50.0%								



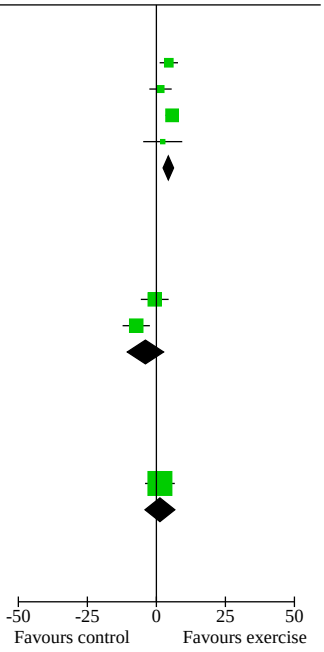
Analysis 8.2. Comparison 8: Fatigue, Outcome 2: Overall fatigue follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
8.2.1 Up to 12 weeks' follow-up								
Bourke 2011	-48	4	8	-43	6	9	2.2%	-0.92 [-1.93 , 0.10]
Burnham 2002	15.3	21.4	14	32.2	34.5	7	2.6%	-0.62 [-1.55 , 0.31]
Cohen 2004	3.1	1.5	16	3.1	1.5	14	3.9%	0.00 [-0.72 , 0.72]
Courneya 2003a	-19.67	11.1	31	-22.37	9.84	45	7.3%	0.26 [-0.20 , 0.72]
Courneya 2009	-40.5	9.4	60	-38	11.1	62	9.6%	-0.24 [-0.60 , 0.12]
Danhauer 2009	-39.8	11.5	13	-32.6	15.5	14	3.5%	-0.51 [-1.28 , 0.26]
Dimeo 2004	34	21	34	39	26	35	7.0%	-0.21 [-0.68 , 0.26]
Donnelly 2011	-36.5	10.61	16	-34	13.3	17	4.2%	-0.20 [-0.89 , 0.48]
Fillion 2008	2.65	0.92	44	2.86	0.83	43	8.0%	-0.24 [-0.66 , 0.18]
Herrero 2006	-4.2	19.6	8	6.4	14.5	8	2.2%	-0.58 [-1.59 , 0.43]
Knols 2011	-43.6	8.4	57	-40	7.5	57	9.2%	-0.45 [-0.82 , -0.08]
McNeely 2008a	-36.7	9	27	-34.3	11.1	25	5.9%	-0.23 [-0.78 , 0.31]
Mehnert 2011	28.52	25.88	30	39.92	25.09	27	6.1%	-0.44 [-0.97 , 0.09]
Milne 2008a	12.9	3	29	16.9	4.7	29	5.8%	-1.00 [-1.55 , -0.45]
Moadel 2007	-30.56	10.85	45	-34.77	13.58	26	6.8%	0.35 [-0.14 , 0.84]
Pinto 2003	7.16	6.4	12	9	6.4	12	3.3%	-0.28 [-1.08 , 0.53]
Pinto 2005	27.08	21.41	39	42.28	26.2	43	7.6%	-0.63 [-1.07 , -0.18]
Rogers 2009	-12.4	10.4	20	-10.1	6.6	18	4.7%	-0.26 [-0.90 , 0.38]
Subtotal (95% CI)			**503**			**491**	**100.0%**	**-0.30 [-0.46 , -0.14]**
Heterogeneity: Tau² = 0.04; Chi² = 25.35, df = 17 (P = 0.09); I² = 33%								
Test for overall effect: Z = 3.64 (P = 0.0003)								
8.2.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-12.7	10.9	62	-12.1	10.8	31	20.1%	-0.05 [-0.49 , 0.38]
Courneya 2003c	-8.3	7.9	24	-8.8	8.1	28	15.6%	0.06 [-0.48 , 0.61]
Fillion 2008	-2.63	0.72	43	-2.24	0.88	42	20.1%	-0.48 [-0.91 , -0.05]
Knols 2011	-42.3	9.1	51	-41.9	7.3	54	22.4%	-0.05 [-0.43 , 0.33]
Thorsen 2005	33.8	21.2	52	25.9	20.7	49	21.8%	0.37 [-0.02 , 0.77]
Subtotal (95% CI)			**232**			**204**	**100.0%**	**-0.03 [-0.31 , 0.25]**
Heterogeneity: Tau² = 0.05; Chi² = 8.37, df = 4 (P = 0.08); I² = 52%								
Test for overall effect: Z = 0.19 (P = 0.85)								
8.2.3 6 months' follow-up								
Cadmus 2009	-51.9	9	37	-50.6	10	37	67.3%	-0.14 [-0.59 , 0.32]
Rogers 2009	-13.9	11.7	19	-13.5	11.3	17	32.7%	-0.03 [-0.69 , 0.62]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**-0.10 [-0.48 , 0.27]**
Heterogeneity: Tau² = 0.00; Chi² = 0.06, df = 1 (P = 0.80); I² = 0%								
Test for overall effect: Z = 0.53 (P = 0.59)								
Test for subgroup differences: Chi² = 3.13, df = 2 (P = 0.21), I² = 36.1%								



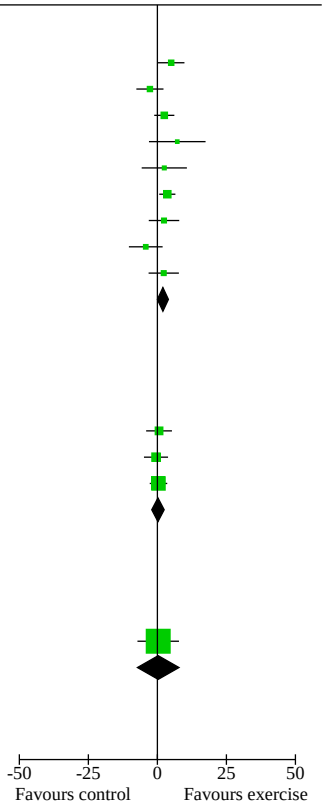
Analysis 8.3. Comparison 8: Fatigue, Outcome 3: FACT fatigue subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.3.1 Up to 12 weeks' follow-up								
Courneya 2009	4.4	10.1	60	-0.1	6.7	38	27.1%	4.50 [1.17 , 7.83]
McNeely 2008a	3.1	9	27	1.6	5.6	25	19.4%	1.50 [-2.54 , 5.54]
Oh 2010	6.34	7.2	54	0.64	5.1	54	46.6%	5.70 [3.35 , 8.05]
Rogers 2009	-1.9	9.6	20	-4.2	12.3	18	6.9%	2.30 [-4.77 , 9.37]
Subtotal (95% CI)			**161**			**135**	**100.0%**	**4.33 [2.43 , 6.22]**
Heterogeneity: Tau² = 0.56; Chi² = 3.48, df = 3 (P = 0.32); I² = 14%								
Test for overall effect: Z = 4.48 (P < 0.00001)								
8.3.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-0.4	11	62	0.2	12	31	49.7%	-0.60 [-5.63 , 4.43]
Courneya 2003c	-9.3	10.2	24	-2	7.5	28	50.3%	-7.30 [-12.24 , -2.36]
Subtotal (95% CI)			**86**			**59**	**100.0%**	**-3.97 [-10.53 , 2.60]**
Heterogeneity: Tau² = 15.97; Chi² = 3.47, df = 1 (P = 0.06); I² = 71%								
Test for overall effect: Z = 1.18 (P = 0.24)								
8.3.3 6 months' follow-up								
Rogers 2009	0.08	7.8	19	-1.2	8.6	17	100.0%	1.28 [-4.11 , 6.67]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**1.28 [-4.11 , 6.67]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.47 (P = 0.64)								
Test for subgroup differences: Chi² = 6.35, df = 2 (P = 0.04), I² = 68.5%								



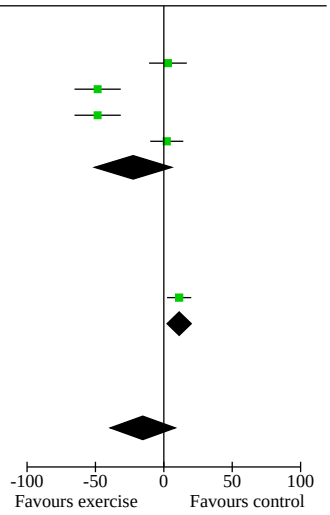
Analysis 8.4. Comparison 8: Fatigue, Outcome 4: FACT fatigue subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.4.1 Up to 12 weeks' follow-up								
Bourke 2011	48	4	8	43	6	9	12.1%	5.00 [0.20 , 9.80]
Courneya 2003a	19.67	11.31	31	22.37	9.84	45	11.8%	-2.70 [-7.61 , 2.21]
Courneya 2009	40.5	9.4	60	38	11.1	62	17.3%	2.50 [-1.15 , 6.15]
Danhauer 2009	39.8	11.5	13	32.6	15.5	14	3.5%	7.20 [-3.05 , 17.45]
Donnelly 2011	36.5	10.61	16	34	13.3	17	5.2%	2.50 [-5.69 , 10.69]
Knols 2011	43.6	8.4	57	40	7.5	57	21.8%	3.60 [0.68 , 6.52]
McNeely 2008a	36.7	9	27	34.3	11.1	25	9.9%	2.40 [-3.12 , 7.92]
Moadel 2007	30.56	10.85	45	34.77	13.58	26	8.5%	-4.21 [-10.32 , 1.90]
Rogers 2009	12.4	10.4	20	10.1	6.6	18	10.0%	2.30 [-3.18 , 7.78]
Subtotal (95% CI)			**277**			**273**	**100.0%**	**2.00 [0.00 , 3.99]**
Heterogeneity: Tau² = 2.54; Chi² = 11.20, df = 8 (P = 0.19); I² = 29%								
Test for overall effect: Z = 1.96 (P = 0.05)								
8.4.2 More than 12 weeks' up to 6 months' follow-up								
Courneya 2003b	12.7	10.9	62	12.1	10.8	31	23.1%	0.60 [-4.07 , 5.27]
Courneya 2003c	8.3	7.9	24	8.8	8.1	28	26.6%	-0.50 [-4.86 , 3.86]
Knols 2011	42.3	9.1	51	41.9	7.3	54	50.3%	0.40 [-2.77 , 3.57]
Subtotal (95% CI)			**137**			**113**	**100.0%**	**0.21 [-2.04 , 2.45]**
Heterogeneity: Tau² = 0.00; Chi² = 0.14, df = 2 (P = 0.93); I² = 0%								
Test for overall effect: Z = 0.18 (P = 0.86)								
8.4.3 6 months' follow-up								
Rogers 2009	13.8	11.7	19	13.5	11.3	17	100.0%	0.30 [-7.22 , 7.82]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**0.30 [-7.22 , 7.82]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.08 (P = 0.94)								
Test for subgroup differences: Chi² = 1.42, df = 2 (P = 0.49), I² = 0%								



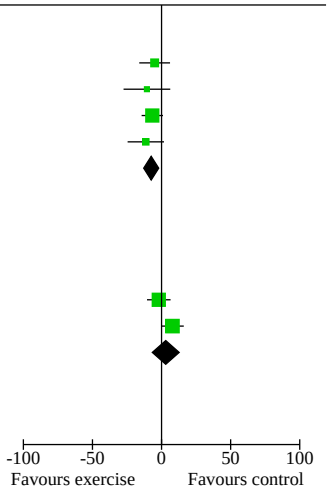
Analysis 8.5. Comparison 8: Fatigue, Outcome 5: QLQ-C30 fatigue subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.5.1 Up to 12 weeks' follow-up								
Dimeo 2004	9	25	34	6	33	35	20.0%	3.00 [-10.79 , 16.79]
Herrero 2006	-42	19.6	8	6.4	14.5	8	19.4%	-48.40 [-65.29 , -31.51]
Herrero 2006	-42	19.6	8	6.4	14.5	8	19.4%	-48.40 [-65.29 , -31.51]
Mehnert 2011	-0.74	20.73	30	-2.88	25.3	27	20.4%	2.14 [-9.95 , 14.23]
Subtotal (95% CI)			**80**			**78**	**79.1%**	**-22.45 [-50.66 , 5.77]**
Heterogeneity: Tau² = 770.07; Chi² = 44.20, df = 3 (P < 0.00001); I² = 93%								
Test for overall effect: Z = 1.56 (P = 0.12)								
8.5.2 More than 12 weeks' less than 6 months' follow-up								
Thorsen 2005	-5.8	22.3	52	-17	23	49	20.9%	11.20 [2.36 , 20.04]
Subtotal (95% CI)			**52**			**49**	**20.9%**	**11.20 [2.36 , 20.04]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.48 (P = 0.01)								
Total (95% CI)			**132**			**127**	**100.0%**	**-15.38 [-39.15 , 8.38]**
Heterogeneity: Tau² = 684.39; Chi² = 66.22, df = 4 (P < 0.00001); I² = 94%								
Test for overall effect: Z = 1.27 (P = 0.20)								
Test for subgroup differences: Chi² = 4.97, df = 1 (P = 0.03), I² = 79.9%								



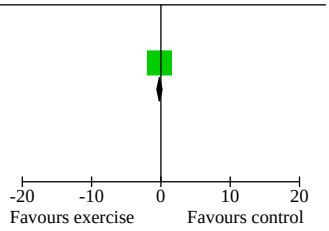
Analysis 8.6. Comparison 8: Fatigue, Outcome 6: QLQ-C30 fatigue subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.6.1 Up to 12 weeks' follow-up								
Dimeo 2004	34	21	34	39	26	35	23.7%	-5.00 [-16.14 , 6.14]
Herrero 2006	-4.2	19.6	8	6.4	14.5	8	10.3%	-10.60 [-27.49 , 6.29]
Knols 2011	31.6	22.1	57	38.3	20	57	49.1%	-6.70 [-14.44 , 1.04]
Mehnert 2011	28.52	25.58	30	39.92	25.09	27	16.9%	-11.40 [-24.57 , 1.77]
Subtotal (95% CI)			**129**			**127**	**100.0%**	**-7.50 [-12.92 , -2.07]**
Heterogeneity: Tau² = 0.00; Chi² = 0.70, df = 3 (P = 0.87); I² = 0%								
Test for overall effect: Z = 2.71 (P = 0.007)								
8.6.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	30.1	24.9	51	32.1	19.1	54	49.2%	-2.00 [-10.52 , 6.52]
Thorsen 2005	33.8	21.2	52	25.9	20.7	49	50.8%	7.90 [-0.27 , 16.07]
Subtotal (95% CI)			**103**			**103**	**100.0%**	**3.03 [-6.67 , 12.73]**
Heterogeneity: Tau² = 30.85; Chi² = 2.70, df = 1 (P = 0.10); I² = 63%								
Test for overall effect: Z = 0.61 (P = 0.54)								
Test for subgroup differences: Chi² = 3.44, df = 1 (P = 0.06), I² = 71.0%								



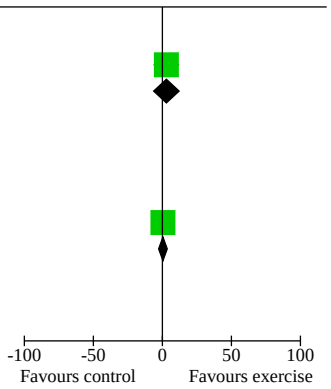
Analysis 8.7. Comparison 8: Fatigue, Outcome 7: Multidimensional Fatigue Inventory follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.7.1 Up to 12 weeks' follow-up								
Fillion 2008	2.65	0.92	44	2.86	0.83	43	100.0%	-0.21 [-0.58 , 0.16]
Subtotal (95% CI)			**44**			**43**	**100.0%**	**-0.21 [-0.58 , 0.16]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.12 (P = 0.26)								
Test for subgroup differences: Not applicable								



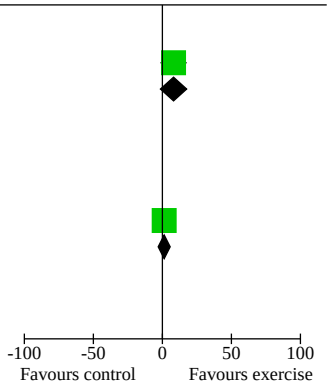
Analysis 8.8. Comparison 8: Fatigue, Outcome 8: MOS SF-36 vitality subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.8.1 Up to 12 weeks' follow-up								
Mehnert 2011	6.89	16.68	30	3.99	19	28	100.0%	2.90 [-6.33 , 12.13]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**2.90 [-6.33 , 12.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.62 (P = 0.54)								
8.8.2 6 months' follow-up								
Cadmus 2009	1.6	6.6	37	1.2	7.1	37	100.0%	0.40 [-2.72 , 3.52]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.40 [-2.72 , 3.52]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.25 (P = 0.80)								
Test for subgroup differences: Chi² = 0.25, df = 1 (P = 0.61), I² = 0%								



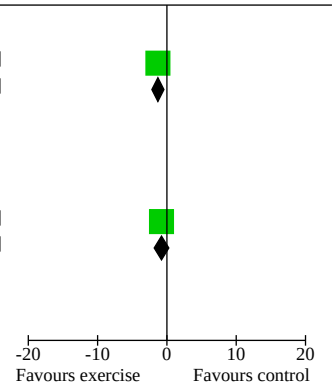
Analysis 8.9. Comparison 8: Fatigue, Outcome 9: MOS SF-36 vitality subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.9.1 Up to 12 weeks' follow-up								
Mehnert 2011	64.22	17.08	30	56.13	19.58	28	100.0%	8.09 [-1.39 , 17.57]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**8.09 [-1.39 , 17.57]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.67 (P = 0.09)								
8.9.2 6 months' follow-up								
Cadmus 2009	51.9	9	37	50.6	10	37	100.0%	1.30 [-3.03 , 5.63]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**1.30 [-3.03 , 5.63]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.59 (P = 0.56)								
Test for subgroup differences: Chi² = 1.63, df = 1 (P = 0.20), I² = 38.6%								



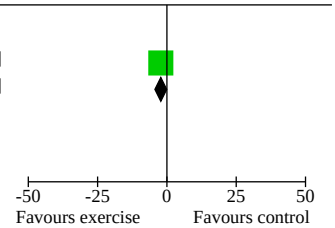
Analysis 8.10. Comparison 8: Fatigue, Outcome 10: Piper Revised Fatigue Scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.10.1 Up to 12 weeks' follow-up								
Daley 2007a	2.14	1.75	33	3.44	1.85	33	100.0%	-1.30 [-2.17 , -0.43]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-1.30 [-2.17 , -0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.93 (P = 0.003)								
8.10.2 More than 12 weeks' less than 6 months' follow-up								
Daley 2007a	2.42	1.95	28	3.2	2.18	31	100.0%	-0.78 [-1.83 , 0.27]
Subtotal (95% CI)			**28**			**31**	**100.0%**	**-0.78 [-1.83 , 0.27]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.45 (P = 0.15)								
Test for subgroup differences: Chi² = 0.56, df = 1 (P = 0.46), I² = 0%								



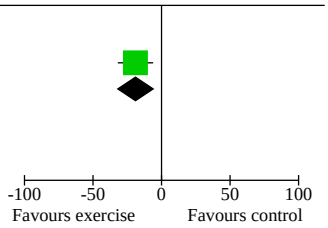
Analysis 8.11. Comparison 8: Fatigue, Outcome 11: Schwartz Cancer Fatigue scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.11.1 Up to 12 weeks' follow-up								
Milne 2008a	-2.7	3.7	29	-0.5	4.5	29	100.0%	-2.20 [-4.32 , -0.08]
Subtotal (95% CI)			**29**			**29**	**100.0%**	**-2.20 [-4.32 , -0.08]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.03 (P = 0.04)								
Test for subgroup differences: Not applicable								



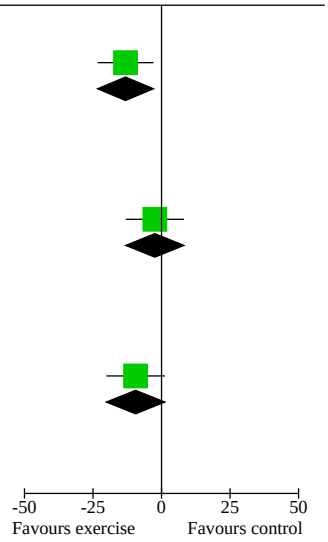
Analysis 8.12. Comparison 8: Fatigue, Outcome 12: Linear Analog Self-Assessment energy scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.12.1 Up to 12 weeks' follow-up								
Burnham 2002	19.7	15.7	14	38.7	13.4	7	100.0%	-19.00 [-31.89 , -6.11]
Subtotal (95% CI)			**14**			**7**	**100.0%**	**-19.00 [-31.89 , -6.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.89 (P = 0.004)								
Test for subgroup differences: Not applicable								



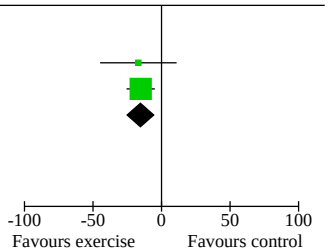
Analysis 8.13. Comparison 8: Fatigue, Outcome 13: Linear Analog Self-Assessment scale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.13.1 Up to 12 weeks' follow-up								
Pinto 2005	-14.93	23.5	39	-1.79	23.5	43	100.0%	-13.14 [-23.32 , -2.96]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**-13.14 [-23.32 , -2.96]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.53 (P = 0.01)								
8.13.2 6 months' follow-up								
Pinto 2005	-10.1	25	39	-7.66	69.7	365	100.0%	-2.44 [-13.06 , 8.18]
Subtotal (95% CI)			**39**			**365**	**100.0%**	**-2.44 [-13.06 , 8.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.45 (P = 0.65)								
8.13.3 More than 6 months' follow-up								
Pinto 2005	-13.31	25.042	39	-3.84	24.1	43	100.0%	-9.47 [-20.13 , 1.19]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**-9.47 [-20.13 , 1.19]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.74 (P = 0.08)								
Test for subgroup differences: Chi² = 2.08, df = 2 (P = 0.35), I² = 3.9%								



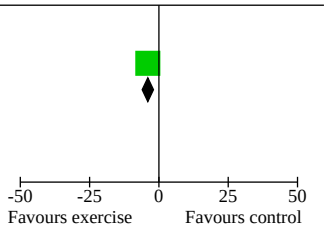
Analysis 8.14. Comparison 8: Fatigue, Outcome 14: Linear Analog Self-Assessment scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.14.1 Up to 12 weeks' follow-up								
Burnham 2002	15.3	21.4	14	32.2	34.5	7	12.0%	-16.90 [-44.81 , 11.01]
Pinto 2005	27.08	21.41	39	42.28	26.2	43	88.0%	-15.20 [-25.52 , -4.88]
Subtotal (95% CI)			**53**			**50**	**100.0%**	**-15.40 [-25.08 , -5.73]**
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.91); I² = 0%								
Test for overall effect: Z = 3.12 (P = 0.002)								
Test for subgroup differences: Not applicable								



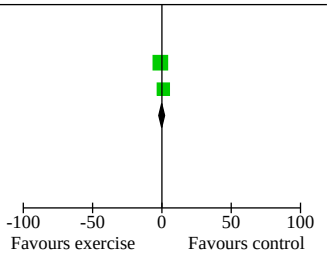
Analysis 8.15. Comparison 8: Fatigue, Outcome 15: Schwartz Cancer Fatigue scale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.15.1 Up to 12 weeks' follow-up								
Milne 2008a	12.9	3	29	16.9	4.7	29	100.0%	-4.00 [-6.03 , -1.97]
Subtotal (95% CI)			**29**			**29**	**100.0%**	**-4.00 [-6.03 , -1.97]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.86 (P = 0.0001)								
Test for subgroup differences: Not applicable								



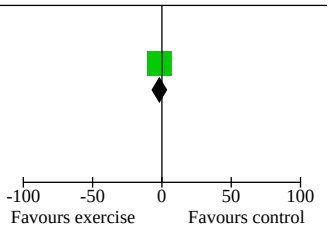
Analysis 8.16. Comparison 8: Fatigue, Outcome 16: POMS fatigue subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.16.1 Up to 12 weeks' follow-up								
Oh 2010	-2.42	5	54	-1.3	4.9	54	55.4%	-1.12 [-2.99 , 0.75]
Targ 2002	1.27	7.1	79	0.25	8.2	88	44.6%	1.02 [-1.30 , 3.34]
Subtotal (95% CI)			**133**			**142**	**100.0%**	**-0.17 [-2.25 , 1.92]**
Heterogeneity: Tau² = 1.13; Chi² = 1.98, df = 1 (P = 0.16); I² = 50%								
Test for overall effect: Z = 0.16 (P = 0.88)								
Test for subgroup differences: Not applicable								



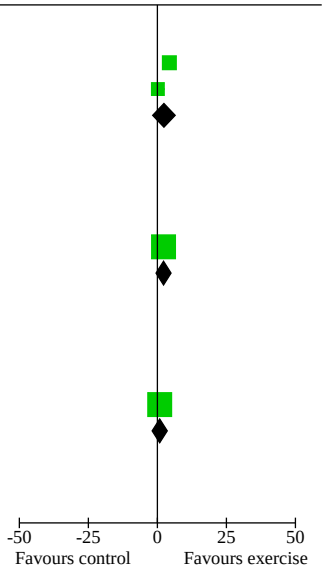
Analysis 8.17. Comparison 8: Fatigue, Outcome 17: POMS - fatigue subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.17.1 Up to 12 weeks' follow-up								
Pinto 2003	7.16	6.4	12	9	6.4	12	100.0%	-1.84 [-6.96 , 3.28]
Subtotal (95% CI)			**12**			**12**	**100.0%**	**-1.84 [-6.96 , 3.28]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.70 (P = 0.48)								
Test for subgroup differences: Not applicable								



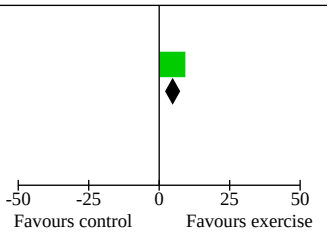
Analysis 8.18. Comparison 8: Fatigue, Outcome 18: POMS vigor subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.18.1 Up to 12 weeks' follow-up								
Oh 2010	3.81	4	54	-0.53	4.3	54	52.6%	4.34 [2.77 , 5.91]
Targ 2002	1.08	7.5	79	0.9	8.6	88	47.4%	0.18 [-2.26 , 2.62]
Subtotal (95% CI)			**133**			**142**	**100.0%**	**2.37 [-1.70 , 6.44]**
Heterogeneity: Tau² = 7.56; Chi² = 7.90, df = 1 (P = 0.005); I² = 87%								
Test for overall effect: Z = 1.14 (P = 0.25)								
8.18.2 6 months' follow-up								
Pinto 2005	0.06	6.4	39	-2.17	6.1	43	100.0%	2.23 [-0.48 , 4.94]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**2.23 [-0.48 , 4.94]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.61 (P = 0.11)								
8.18.3 More than 6 months' follow-up								
Pinto 2005	2.04	6.4	39	1.2	6.2	43	100.0%	0.84 [-1.89 , 3.57]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**0.84 [-1.89 , 3.57]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.60 (P = 0.55)								
Test for subgroup differences: Chi² = 0.63, df = 2 (P = 0.73), I² = 0%								



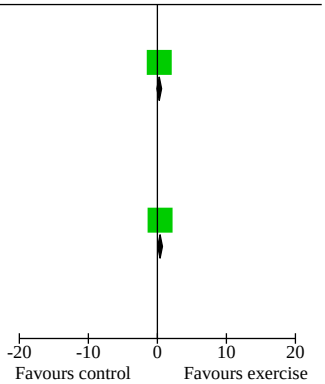
Analysis 8.19. Comparison 8: Fatigue, Outcome 19: POMS vigor subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
8.19.1 Up to 12 weeks' follow-up								
Pinto 2005	20.58	5.7	39	15.81	5.39	43	100.0%	4.77 [2.36 , 7.18]
Subtotal (95% CI)			**39**			**43**	**100.0%**	**4.77 [2.36 , 7.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.88 (P = 0.0001)								
Test for subgroup differences: Not applicable								

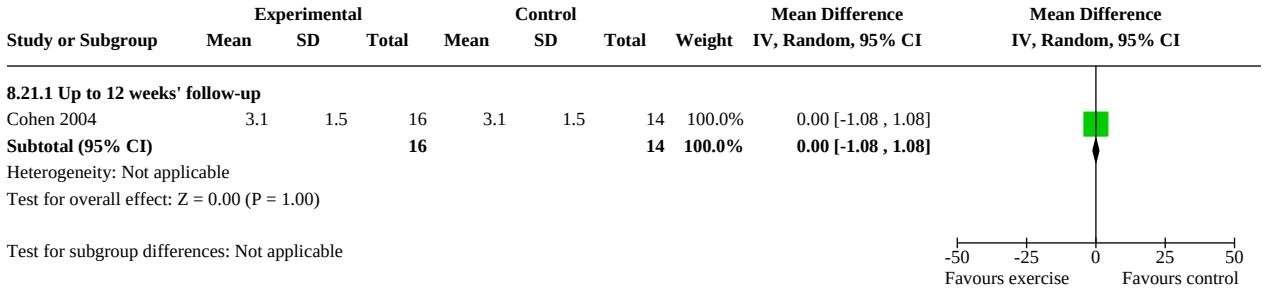


Analysis 8.20. Comparison 8: Fatigue, Outcome 20: POMS short form vigor follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
8.20.1 Up to 12 weeks' follow-up								
Fillion 2008	2.41	0.72	44	2.14	0.87	42	100.0%	0.27 [-0.07 , 0.61]
Subtotal (95% CI)			**44**			**42**	**100.0%**	**0.27 [-0.07 , 0.61]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.56 (P = 0.12)								
8.20.2 More than 12 weeks' less than 6 months' follow-up								
Fillion 2008	2.63	0.72	43	2.24	0.88	42	100.0%	0.39 [0.05 , 0.73]
Subtotal (95% CI)			**43**			**42**	**100.0%**	**0.39 [0.05 , 0.73]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.23 (P = 0.03)								
Test for subgroup differences: Chi² = 0.24, df = 1 (P = 0.63), I² = 0%								



Analysis 8.21. Comparison 8: Fatigue, Outcome 21: Brief Fatigue Inventory follow-up values



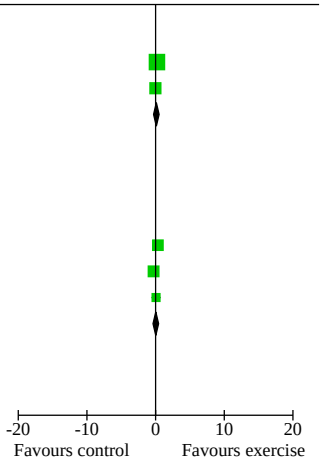
Comparison 9. General health perspective

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
9.1 Overall general health change	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
9.1.1 Up to 12 weeks' follow-up	2	96	Std. Mean Difference (IV, Random, 95% CI)	0.11 [-0.29, 0.51]
9.1.2 6 months' follow-up	3	189	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.38, 0.44]
9.2 Overall general health follow-up values	6		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
9.2.1 Up to 12 weeks' follow-up	4	249	Std. Mean Difference (IV, Random, 95% CI)	0.14 [-0.20, 0.49]
9.2.2 More than 12 weeks' less than 6 months' follow-up	1	105	Std. Mean Difference (IV, Random, 95% CI)	0.19 [-0.19, 0.58]
9.2.3 6 months' follow-up	3	189	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.30, 0.36]
9.3 QLQ-C30 subscale follow-up values	2	288	Mean Difference (IV, Random, 95% CI)	4.03 [0.14, 7.92]
9.3.1 Up to 12 weeks' follow-up	2	183	Mean Difference (IV, Random, 95% CI)	4.25 [-1.09, 9.60]
9.3.2 More than 12 weeks' less than 6 months' follow-up	1	105	Mean Difference (IV, Random, 95% CI)	3.30 [-3.25, 9.85]
9.4 CARES subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.4.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-1.30 [-3.23, 0.63]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
9.5 CARES subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.5.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-0.60 [-3.09, 1.89]
9.6 MOS general health subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.6.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	3.36 [-5.43, 12.15]
9.6.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	3.00 [-0.98, 6.98]
9.7 MOS general health subscale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.7.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	4.37 [-4.43, 13.17]
9.7.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	5.00 [-1.92, 11.92]
9.8 Single question change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.8.1 Up to 12 weeks' follow-up	1	38	Mean Difference (IV, Random, 95% CI)	-0.02 [-0.44, 0.40]
9.8.2 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	0.04 [-0.35, 0.43]
9.9 Single question follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.9.1 Up to 12 weeks' follow-up	1	38	Mean Difference (IV, Random, 95% CI)	-0.30 [-0.76, 0.16]
9.9.2 6 months' follow-up	1	36	Mean Difference (IV, Random, 95% CI)	-0.20 [-0.79, 0.39]

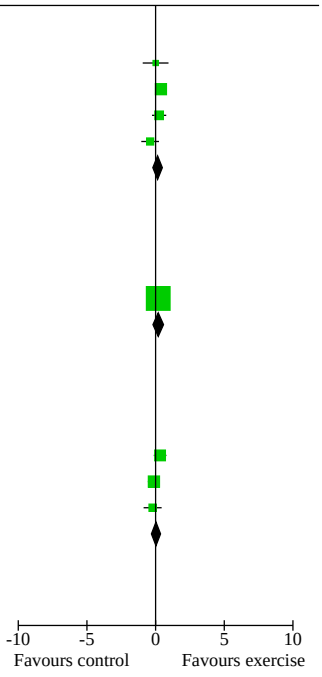
Analysis 9.1. Comparison 9: General health perspective, Outcome 1: Overall general health change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
9.1.1 Up to 12 weeks' follow-up								
Mehnert 2011	6.6	13.7	30	3.24	19.7	28	60.3%	0.20 [-0.32 , 0.71]
Rogers 2009	0.15	0.6	20	0.17	0.7	18	39.7%	-0.03 [-0.67 , 0.61]
Subtotal (95% CI)			**50**			**46**	**100.0%**	**0.11 [-0.29 , 0.51]**
Heterogeneity: Tau² = 0.00; Chi² = 0.29, df = 1 (P = 0.59); I² = 0%								
Test for overall effect: Z = 0.52 (P = 0.60)								
9.1.2 6 months' follow-up								
Cadmus 2009	0.6	7.5	37	-2.4	9.8	37	37.0%	0.34 [-0.12 , 0.80]
Ohira 2006	-0.7	4.8	39	0.6	3.9	40	38.2%	-0.29 [-0.74 , 0.15]
Rogers 2009	0.16	0.6	19	0.12	0.6	17	24.9%	0.07 [-0.59 , 0.72]
Subtotal (95% CI)			**95**			**94**	**100.0%**	**0.03 [-0.38 , 0.44]**
Heterogeneity: Tau² = 0.06; Chi² = 3.82, df = 2 (P = 0.15); I² = 48%								
Test for overall effect: Z = 0.14 (P = 0.89)								
Test for subgroup differences: Chi² = 0.07, df = 1 (P = 0.79), I² = 0%								



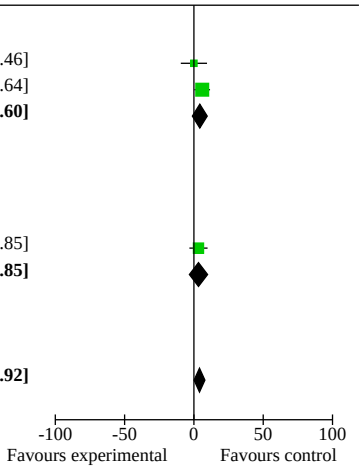
Analysis 9.2. Comparison 9: General health perspective, Outcome 2: Overall general health follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
9.2.1 Up to 12 weeks' follow-up								
Dimeo 2004	62	21	34	62	19	5	11.4%	0.00 [-0.94 , 0.94]
Knols 2011	73.7	15	57	67.7	15.7	57	40.0%	0.39 [0.02 , 0.76]
Mehnert 2011	74.38	11.84	30	70.01	20.81	28	27.8%	0.26 [-0.26 , 0.77]
Rogers 2009	3.5	0.9	20	3.8	0.5	18	20.7%	-0.40 [-1.04 , 0.25]
Subtotal (95% CI)			**141**			**108**	**100.0%**	**0.14 [-0.20 , 0.49]**
Heterogeneity: Tau² = 0.04; Chi² = 4.52, df = 3 (P = 0.21); I² = 34%								
Test for overall effect: Z = 0.82 (P = 0.41)								
9.2.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	72.9	17.2	51	69.6	17	54	100.0%	0.19 [-0.19 , 0.58]
Subtotal (95% CI)			**51**			**54**	**100.0%**	**0.19 [-0.19 , 0.58]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.98 (P = 0.33)								
9.2.3 6 months' follow-up								
Cadmus 2009	91.2	12.6	37	86.2	17.4	37	38.1%	0.33 [-0.13 , 0.78]
Ohira 2006	-53.9	5.3	39	-53.2	6	40	40.3%	-0.12 [-0.56 , 0.32]
Rogers 2009	3.5	0.9	19	3.7	0.9	17	21.6%	-0.22 [-0.87 , 0.44]
Subtotal (95% CI)			**95**			**94**	**100.0%**	**0.03 [-0.30 , 0.36]**
Heterogeneity: Tau² = 0.02; Chi² = 2.60, df = 2 (P = 0.27); I² = 23%								
Test for overall effect: Z = 0.17 (P = 0.87)								
Test for subgroup differences: Chi² = 0.45, df = 2 (P = 0.80), I² = 0%								



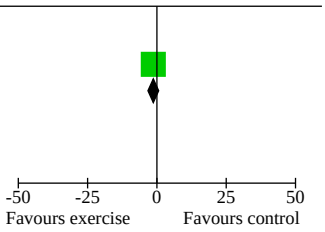
Analysis 9.3. Comparison 9: General health perspective, Outcome 3: QLQ-C30 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
9.3.1 Up to 12 weeks' follow-up								
Dimeo 2004	62	21	34	62	19	35	16.9%	0.00 [-9.46 , 9.46]
Knols 2011	73.7	15	57	67.7	15.7	57	47.7%	6.00 [0.36 , 11.64]
Subtotal (95% CI)			**91**			**92**	**64.6%**	**4.25 [-1.09 , 9.60]**
Heterogeneity: Tau² = 2.22; Chi² = 1.14, df = 1 (P = 0.29); I² = 12%								
Test for overall effect: Z = 1.56 (P = 0.12)								
9.3.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	72.9	17.2	51	69.6	17	54	35.4%	3.30 [-3.25 , 9.85]
Subtotal (95% CI)			**51**			**54**	**35.4%**	**3.30 [-3.25 , 9.85]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.99 (P = 0.32)								
Total (95% CI)			**142**			**146**	**100.0%**	**4.03 [0.14 , 7.92]**
Heterogeneity: Tau² = 0.00; Chi² = 1.21, df = 2 (P = 0.54); I² = 0%								
Test for overall effect: Z = 2.03 (P = 0.04)								
Test for subgroup differences: Chi² = 0.05, df = 1 (P = 0.83), I² = 0%								



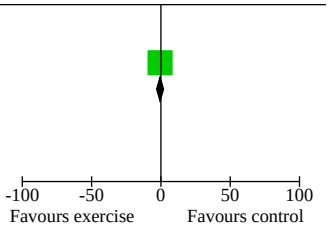
Analysis 9.4. Comparison 9: General health perspective, Outcome 4: CARES subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
9.4.1 6 months' follow-up								
Ohira 2006	-0.7	4.8	39	0.6	3.9	40	100.0%	-1.30 [-3.23 , 0.63]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-1.30 [-3.23 , 0.63]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Not applicable								



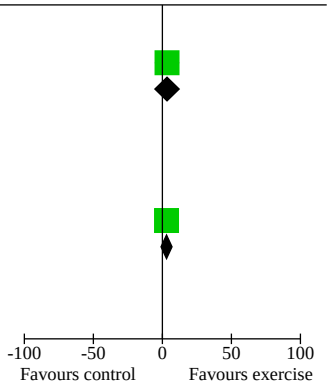
Analysis 9.5. Comparison 9: General health perspective, Outcome 5: CARES subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
9.5.1 6 months' follow-up								
Ohira 2006	-53.8	5.3	39	-53.2	6	40	100.0%	-0.60 [-3.09 , 1.89]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-0.60 [-3.09 , 1.89]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.47 (P = 0.64)								
Test for subgroup differences: Not applicable								



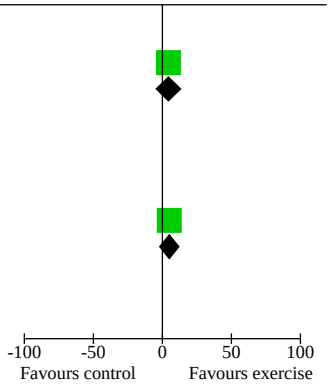
Analysis 9.6. Comparison 9: General health perspective, Outcome 6: MOS general health subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
9.6.1 Up to 12 weeks' follow-up								
Mehnert 2011	6.6	13.7	30	3.24	19.7	28	100.0%	3.36 [-5.43 , 12.15]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**3.36 [-5.43 , 12.15]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								
9.6.2 6 months' follow-up								
Cadmus 2009	0.6	7.5	37	-2.4	9.8	37	100.0%	3.00 [-0.98 , 6.98]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**3.00 [-0.98 , 6.98]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.48 (P = 0.14)								
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.94), I² = 0%								



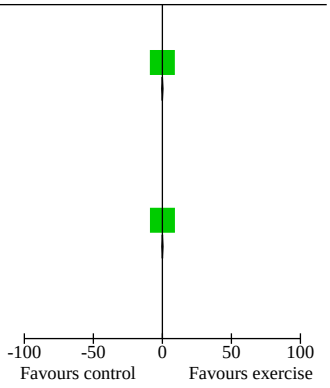
Analysis 9.7. Comparison 9: General health perspective, Outcome 7: MOS general health subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
9.7.1 Up to 12 weeks' follow-up								
Mehnert 2011	74.38	11.84	30	70.01	20.81	28	100.0%	4.37 [-4.43 , 13.17]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**4.37 [-4.43 , 13.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.97 (P = 0.33)								
9.7.2 6 months' follow-up								
Cadmus 2009	91.2	12.6	37	86.2	17.4	37	100.0%	5.00 [-1.92 , 11.92]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**5.00 [-1.92 , 11.92]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.42 (P = 0.16)								
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.91), I² = 0%								

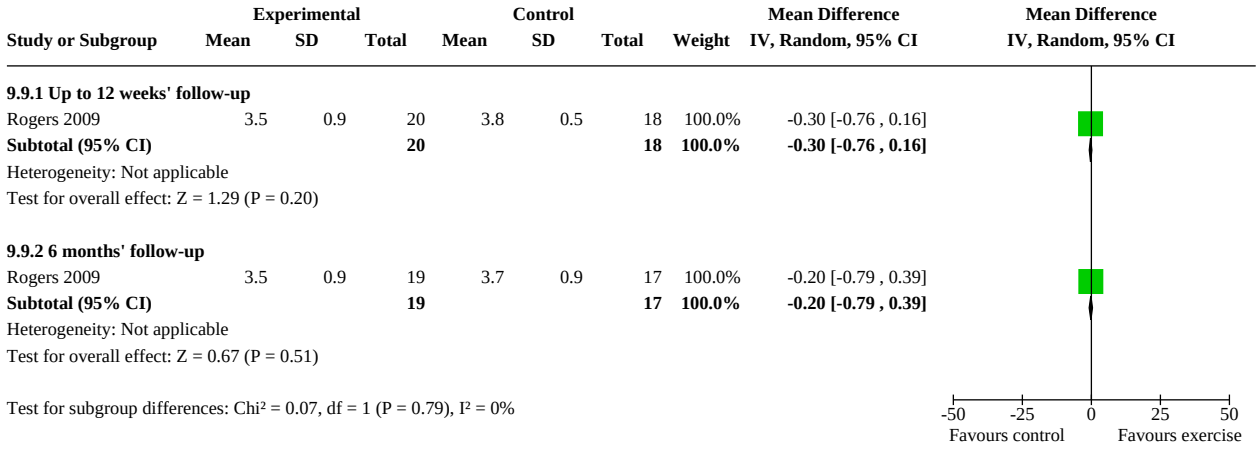


Analysis 9.8. Comparison 9: General health perspective, Outcome 8: Single question change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
9.8.1 Up to 12 weeks' follow-up								
Rogers 2009	0.15	0.6	20	0.17	0.7	18	100.0%	-0.02 [-0.44 , 0.40]
Subtotal (95% CI)			**20**			**18**	**100.0%**	**-0.02 [-0.44 , 0.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.09 (P = 0.93)								
9.8.2 6 months' follow-up								
Rogers 2009	0.16	0.6	19	0.12	0.6	17	100.0%	0.04 [-0.35 , 0.43]
Subtotal (95% CI)			**19**			**17**	**100.0%**	**0.04 [-0.35 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.20 (P = 0.84)								
Test for subgroup differences: Chi² = 0.04, df = 1 (P = 0.84), I² = 0%								



Analysis 9.9. Comparison 9: General health perspective, Outcome 9: Single question follow-up values

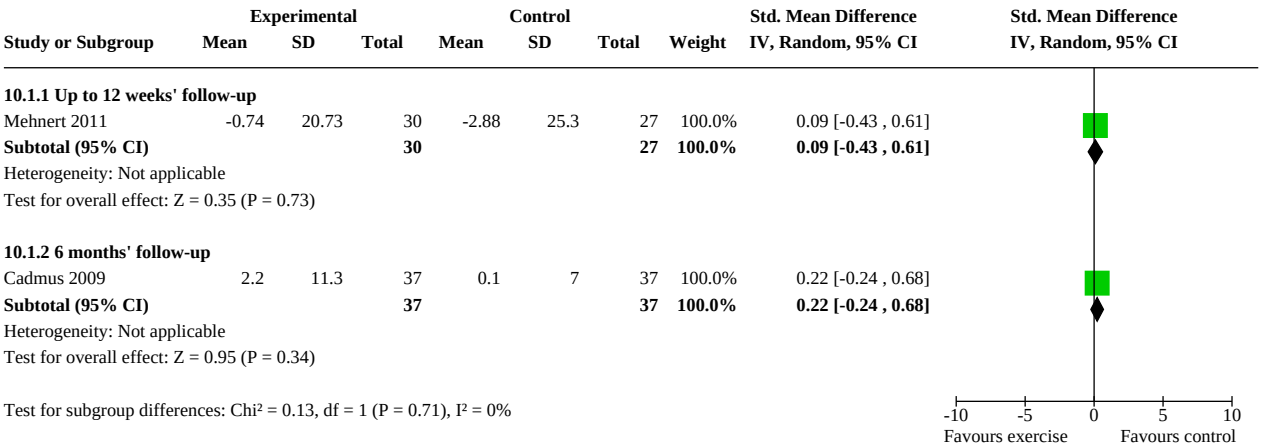


Comparison 10. Pain

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
10.1 Overall pain change	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
10.1.1 Up to 12 weeks' follow-up	1	57	Std. Mean Difference (IV, Random, 95% CI)	0.09 [-0.43, 0.61]
10.1.2 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	0.22 [-0.24, 0.68]
10.2 Overall pain follow-up values	5		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
10.2.1 Up to 12 weeks' follow-up	4	289	Std. Mean Difference (IV, Random, 95% CI)	-0.29 [-0.55, -0.04]
10.2.2 More than 12 weeks' less than 6 months' follow-up	1	105	Std. Mean Difference (IV, Random, 95% CI)	0.05 [-0.34, 0.43]
10.2.3 6 months' follow-up	1	74	Std. Mean Difference (IV, Random, 95% CI)	0.05 [-0.40, 0.51]
10.3 QLQ-C30 pain subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
10.3.1 Up to 12 weeks' follow-up	1	57	Mean Difference (IV, Random, 95% CI)	2.14 [-9.95, 14.23]
10.4 QLQ-C30 subscale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
10.4.1 Up to 12 weeks' follow-up	2	183	Mean Difference (IV, Random, 95% CI)	-3.73 [-13.52, 6.05]

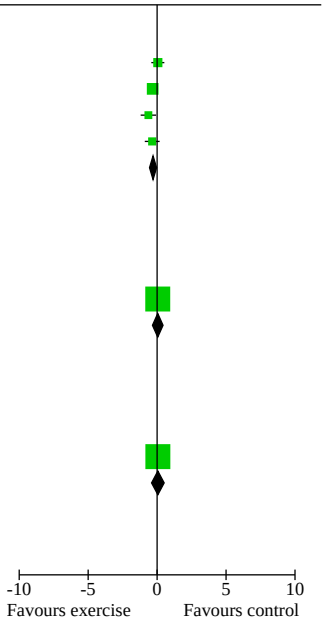
Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
10.4.2 More than 12 weeks' less than 6 months' follow-up	1	105	Mean Difference (IV, Random, 95% CI)	1.30 [-9.41, 12.01]
10.5 MOS SF-36 subscale change	2	132	Mean Difference (IV, Random, 95% CI)	1.09 [-3.86, 6.04]
10.5.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	-5.21 [-18.13, 7.71]
10.5.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	2.10 [-2.18, 6.38]
10.6 MOS SF-36 follow-up values	2	132	Mean Difference (IV, Random, 95% CI)	0.14 [-3.77, 4.05]
10.6.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	5.90 [-6.47, 18.27]
10.6.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-0.50 [-4.62, 3.62]
10.7 SPADI subscale follow-up values	1	52	Mean Difference (IV, Random, 95% CI)	-14.00 [-26.04, -1.96]
10.7.1 Up to 12 weeks' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	-14.00 [-26.04, -1.96]

Analysis 10.1. Comparison 10: Pain, Outcome 1: Overall pain change



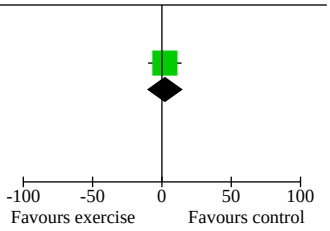
Analysis 10.2. Comparison 10: Pain, Outcome 2: Overall pain follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
10.2.1 Up to 12 weeks' follow-up								
Dimeo 2004	30	27	34	29	3	35	24.8%	0.05 [-0.42 , 0.52]
Knols 2011	17.8	27.1	57	26.8	28.4	57	36.8%	-0.32 [-0.69 , 0.05]
McNeely 2008a	15	17	27	29	26	25	18.6%	-0.63 [-1.19 , -0.07]
Mehnert 2011	21.67	25.58	27	31.48	28.62	27	19.8%	-0.36 [-0.89 , 0.18]
Subtotal (95% CI)			**145**			**144**	**100.0%**	**-0.29 [-0.55 , -0.04]**
Heterogeneity: Tau² = 0.01; Chi² = 3.55, df = 3 (P = 0.31); I² = 15%								
Test for overall effect: Z = 2.25 (P = 0.02)								
10.2.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	21.2	30.6	51	19.9	24.9	54	100.0%	0.05 [-0.34 , 0.43]
Subtotal (95% CI)			**51**			**54**	**100.0%**	**0.05 [-0.34 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
10.2.3 6 months' follow-up								
Cadmus 2009	-50.3	9.1	37	-50.8	9	37	100.0%	0.05 [-0.40 , 0.51]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.05 [-0.40 , 0.51]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
Test for subgroup differences: Chi² = 2.99, df = 2 (P = 0.22), I² = 33.2%								



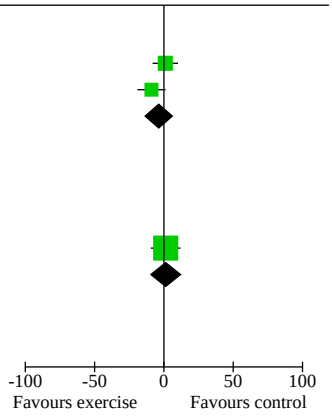
Analysis 10.3. Comparison 10: Pain, Outcome 3: QLQ-C30 pain subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
10.3.1 Up to 12 weeks' follow-up								
Mehnert 2011	-0.74	20.73	30	-2.88	25.3	27	100.0%	2.14 [-9.95 , 14.23]
Subtotal (95% CI)			**30**			**27**	**100.0%**	**2.14 [-9.95 , 14.23]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.73)								
Test for subgroup differences: Not applicable								



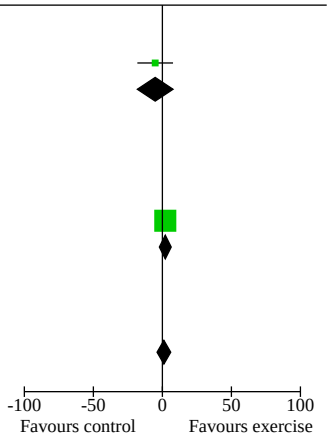
Analysis 10.4. Comparison 10: Pain, Outcome 4: QLQ-C30 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
10.4.1 Up to 12 weeks' follow-up								
Dimeo 2004	30	27	34	29	3	35	52.7%	1.00 [-8.13 , 10.13]
Knols 2011	17.8	27.1	57	26.8	28.4	57	47.3%	-9.00 [-19.19 , 1.19]
Subtotal (95% CI)			**91**			**92**	**100.0%**	**-3.73 [-13.52 , 6.05]**
Heterogeneity: Tau² = 25.63; Chi² = 2.05, df = 1 (P = 0.15); I² = 51%								
Test for overall effect: Z = 0.75 (P = 0.45)								
10.4.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	21.2	30.6	51	19.9	24.9	54	100.0%	1.30 [-9.41 , 12.01]
Subtotal (95% CI)			**51**			**54**	**100.0%**	**1.30 [-9.41 , 12.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
Test for subgroup differences: Chi² = 0.46, df = 1 (P = 0.50), I² = 0%								



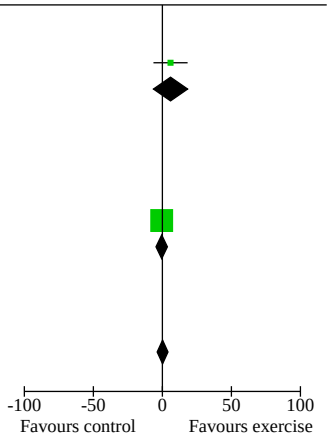
Analysis 10.5. Comparison 10: Pain, Outcome 5: MOS SF-36 subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
10.5.1 Up to 12 weeks' follow-up								
Mehnert 2011	2.4	24.95	30	7.61	25.2	28	13.8%	-5.21 [-18.13 , 7.71]
Subtotal (95% CI)			**30**			**28**	**13.8%**	**-5.21 [-18.13 , 7.71]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.79 (P = 0.43)								
10.5.2 6 months' follow-up								
Cadmus 2009	2.2	11.3	37	0.1	7	37	86.2%	2.10 [-2.18 , 6.38]
Subtotal (95% CI)			**37**			**37**	**86.2%**	**2.10 [-2.18 , 6.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.96 (P = 0.34)								
Total (95% CI)			**67**			**65**	**100.0%**	**1.09 [-3.86 , 6.04]**
Heterogeneity: Tau² = 2.62; Chi² = 1.11, df = 1 (P = 0.29); I² = 10%								
Test for overall effect: Z = 0.43 (P = 0.67)								
Test for subgroup differences: Chi² = 1.11, df = 1 (P = 0.29), I² = 9.8%								



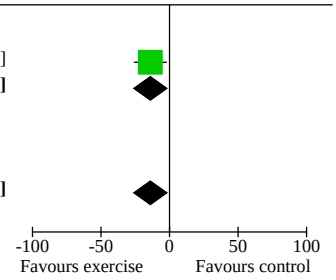
Analysis 10.6. Comparison 10: Pain, Outcome 6: MOS SF-36 follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
10.6.1 Up to 12 weeks' follow-up								
Mehnert 2011	79.4	23.1	30	73.5	24.83	28	10.0%	5.90 [-6.47 , 18.27]
Subtotal (95% CI)			**30**			**28**	**10.0%**	**5.90 [-6.47 , 18.27]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.94 (P = 0.35)								
10.6.2 6 months' follow-up								
Cadmus 2009	50.3	9.1	37	50.8	9	37	90.0%	-0.50 [-4.62 , 3.62]
Subtotal (95% CI)			**37**			**37**	**90.0%**	**-0.50 [-4.62 , 3.62]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
Total (95% CI)			**67**			**65**	**100.0%**	**0.14 [-3.77 , 4.05]**
Heterogeneity: Tau² = 0.00; Chi² = 0.93, df = 1 (P = 0.34); I² = 0%								
Test for overall effect: Z = 0.07 (P = 0.94)								
Test for subgroup differences: Chi² = 0.93, df = 1 (P = 0.34), I² = 0%								



Analysis 10.7. Comparison 10: Pain, Outcome 7: SPADI subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
10.7.1 Up to 12 weeks' follow-up								
McNeely 2008a	15	17	27	29	26	25	100.0%	-14.00 [-26.04 , -1.96]
Subtotal (95% CI)			**27**			**25**	**100.0%**	**-14.00 [-26.04 , -1.96]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.28 (P = 0.02)								
Total (95% CI)			**27**			**25**	**100.0%**	**-14.00 [-26.04 , -1.96]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.28 (P = 0.02)								
Test for subgroup differences: Not applicable								



Comparison 11. Physical functioning

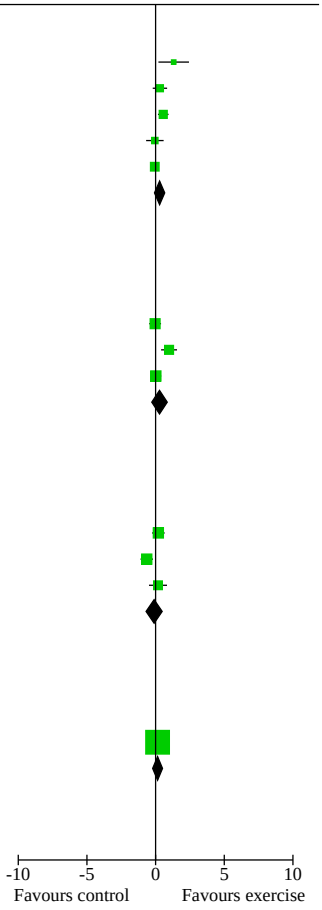
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
11.1 Overall physical functioning change	11		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
11.1.1 Up to 12 weeks' follow-up	5	386	Std. Mean Difference (IV, Random, 95% CI)	0.29 [-0.08, 0.66]
11.1.2 More than 12 weeks' less than 6 months' follow-up	3	246	Std. Mean Difference (IV, Random, 95% CI)	0.28 [-0.28, 0.85]
11.1.3 6 months' follow-up	3	189	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.69, 0.48]
11.1.4 More than 6 months' follow-up	1	120	Std. Mean Difference (IV, Random, 95% CI)	0.14 [-0.21, 0.50]
11.2 Overall physical functioning follow-up values	20		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
11.2.1 Up to 12 weeks' follow-up	15	878	Std. Mean Difference (IV, Random, 95% CI)	0.36 [0.09, 0.64]
11.2.2 More than 12 weeks' less than 6 months' follow-up	7	559	Std. Mean Difference (IV, Random, 95% CI)	0.25 [-0.05, 0.54]
11.2.3 6 months' follow-up	2	110	Std. Mean Difference (IV, Random, 95% CI)	0.19 [-0.18, 0.57]
11.2.4 More than 6 months' follow-up	1	120	Std. Mean Difference (IV, Random, 95% CI)	0.19 [-0.17, 0.55]
11.3 FACT physical function subscale change	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.3.1 Up to 12 weeks' follow-up	3	313	Mean Difference (IV, Random, 95% CI)	0.74 [-1.14, 2.62]
11.3.2 More than 12 weeks' less than 6 months' follow-up	2	145	Mean Difference (IV, Random, 95% CI)	1.22 [-0.97, 3.41]
11.3.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	0.64 [-0.60, 1.89]
11.4 FACT physical function subscale follow-up values	9	673	Mean Difference (IV, Random, 95% CI)	1.13 [-0.09, 2.35]
11.4.1 Up to 12 weeks' follow-up	6	356	Mean Difference (IV, Random, 95% CI)	1.60 [-0.71, 3.92]
11.4.2 More than 12 weeks' less than 6 months' follow-up	3	207	Mean Difference (IV, Random, 95% CI)	0.55 [-0.45, 1.55]
11.4.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	0.85 [-0.58, 2.28]
11.5 QLQ-C30 subscale change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.5.1 More than 12 weeks' less than 6 months' follow-up	2	73	Mean Difference (IV, Random, 95% CI)	6.23 [1.74, 10.72]
11.5.2 More than 12 weeks' less than 6 months' follow-up	1	101	Mean Difference (IV, Random, 95% CI)	0.30 [-6.83, 7.43]
11.6 QLQ-C30 subscale follow-up values	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.6.1 Up to 12 weeks' follow-up	4	256	Mean Difference (IV, Random, 95% CI)	2.55 [-0.29, 5.38]
11.6.2 More than 12 weeks' less than 6 months' follow-up	2	206	Mean Difference (IV, Random, 95% CI)	0.90 [-6.83, 8.64]
11.7 MOS SF-36 subscale change	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.7.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	6.11 [-1.19, 13.41]
11.7.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-0.20 [-2.37, 1.97]
11.7.3 More than 6 months' follow-up	1	120	Mean Difference (IV, Random, 95% CI)	2.50 [-3.66, 8.66]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
11.8 MOS SF-12 subscale follow-up values	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.8.1 Up to 12 weeks' follow-up	3	172	Mean Difference (IV, Random, 95% CI)	4.74 [0.31, 9.17]
11.8.2 More than 12 weeks' less than 6 months' follow-up	1	87	Mean Difference (IV, Random, 95% CI)	2.12 [-2.18, 6.42]
11.9 MOS SF-36 subscale follow-up values	3	323	Mean Difference (IV, Random, 95% CI)	6.42 [1.14, 11.71]
11.9.1 Up to 12 weeks' follow-up	1	59	Mean Difference (IV, Random, 95% CI)	9.83 [5.26, 14.40]
11.9.2 6 months' follow-up	2	144	Mean Difference (IV, Random, 95% CI)	7.93 [-4.16, 20.02]
11.9.3 More than 6 months' follow-up	1	120	Mean Difference (IV, Random, 95% CI)	1.40 [-1.25, 4.05]
11.10 CARES subscale change	1	79	Mean Difference (IV, Random, 95% CI)	-3.30 [-5.54, -1.06]
11.10.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-3.30 [-5.54, -1.06]
11.11 CARES subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.11.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-4.10 [-7.06, -1.14]
11.12 Body Esteem Scale - physical condition follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
11.12.1 Up to 12 weeks' follow-up	2	106	Mean Difference (IV, Random, 95% CI)	4.41 [0.57, 8.25]

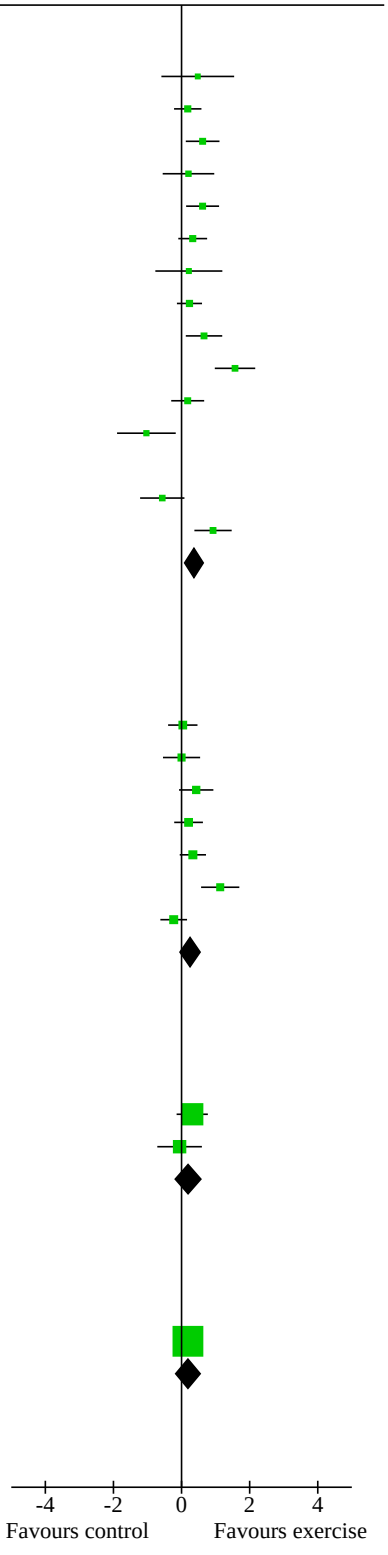
Analysis 11.1. Comparison 11: Physical functioning, Outcome 1: Overall physical functioning change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
11.1.1 Up to 12 weeks' follow-up								
Herrero 2006	6.7	5	8	-1.7	6.9	8	8.5%	1.32 [0.21 , 2.43]
Mehnert 2011	5.3	10.92	30	1.49	13	27	20.6%	0.31 [-0.21 , 0.84]
Oh 2010	3.06	4	54	0.98	3.4	54	25.4%	0.56 [0.17 , 0.94]
Rogers 2009	2.3	5.1	20	2.6	6	18	17.2%	-0.05 [-0.69 , 0.58]
Targ 2002	0.88	6.2	79	1.28	7.1	88	28.3%	-0.06 [-0.36 , 0.24]
Subtotal (95% CI)			**191**			**195**	**100.0%**	**0.29 [-0.08 , 0.66]**
Heterogeneity: Tau² = 0.10; Chi² = 10.76, df = 4 (P = 0.03); I² = 63%								
Test for overall effect: Z = 1.54 (P = 0.12)								
11.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-0.5	5.7	62	-0.3	5.7	31	34.5%	-0.03 [-0.47 , 0.40]
Courneya 2003c	1.9	2	24	-0.2	2.2	28	29.7%	0.98 [0.40 , 1.56]
Thorsen 2005	11	16.1	52	10.7	20.1	49	35.8%	0.02 [-0.37 , 0.41]
Subtotal (95% CI)			**138**			**108**	**100.0%**	**0.28 [-0.28 , 0.85]**
Heterogeneity: Tau² = 0.19; Chi² = 8.95, df = 2 (P = 0.01); I² = 78%								
Test for overall effect: Z = 0.99 (P = 0.32)								
11.1.3 6 months' follow-up								
Cadmus 2009	0.1	2.8	37	-0.5	3.1	37	35.5%	0.20 [-0.26 , 0.66]
Ohira 2006	-2.1	5.6	39	1.2	4.5	40	35.6%	-0.64 [-1.10 , -0.19]
Rogers 2009	1.8	5.4	19	0.9	4.6	17	28.9%	0.17 [-0.48 , 0.83]
Subtotal (95% CI)			**95**			**94**	**100.0%**	**-0.11 [-0.69 , 0.48]**
Heterogeneity: Tau² = 0.20; Chi² = 7.78, df = 2 (P = 0.02); I² = 74%								
Test for overall effect: Z = 0.36 (P = 0.72)								
11.1.4 More than 6 months' follow-up								
Speck 2010	6.6	17.1	58	4.1	17.3	62	100.0%	0.14 [-0.21 , 0.50]
Subtotal (95% CI)			**58**			**62**	**100.0%**	**0.14 [-0.21 , 0.50]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.79 (P = 0.43)								
Test for subgroup differences: Chi² = 1.45, df = 3 (P = 0.69), I² = 0%								



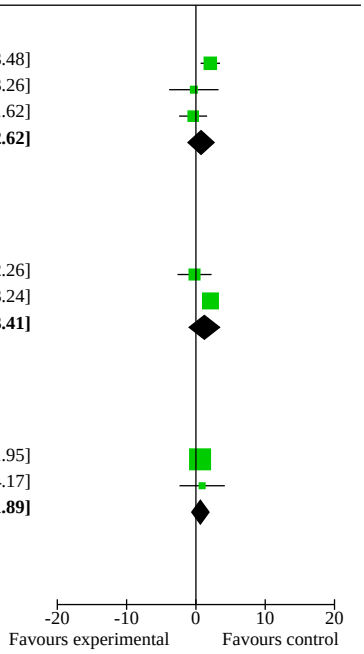
Analysis 11.2. Comparison 11: Physical functioning, Outcome 2: Overall physical functioning follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
11.2.1 Up to 12 weeks' follow-up								
Banasik 2011	0.77	0.94	7	0.42	0.25	7	4.1%	0.48 [-0.59 , 1.54]
Courneya 2003a	20.3	5.14	51	19.36	5.11	45	8.8%	0.18 [-0.22 , 0.58]
Daley 2007a	25.61	2.15	33	23.61	3.97	33	8.0%	0.62 [0.12 , 1.11]
Danhauer 2009	22.5	7.6	13	21.1	5.7	14	5.9%	0.20 [-0.55 , 0.96]
Dimeo 2004	76	1.6	34	75	1.6	35	8.1%	0.62 [0.13 , 1.10]
Fillion 2008	45.1	10.42	44	41.76	9.76	43	8.7%	0.33 [-0.10 , 0.75]
Herrero 2006	94.1	7.5	8	92.5	6.6	8	4.5%	0.21 [-0.77 , 1.20]
Knols 2011	83.7	14.2	57	80.4	14	57	9.1%	0.23 [-0.14 , 0.60]
Mehnert 2011	94	9.32	30	85.95	14.48	27	7.7%	0.66 [0.12 , 1.19]
Milne 2008a	21.9	3	29	16	4.3	29	7.2%	1.57 [0.98 , 2.16]
Moadel 2007	21.03	6.05	45	19.91	6.3	26	8.1%	0.18 [-0.30 , 0.66]
Pinto 2003	-32.3	4.3	12	-25.3	8.2	12	5.2%	-1.03 [-1.89 , -0.17]
Pinto 2005	-2.21	0	39	-0.18	0	43		Not estimable
Rogers 2009	23.3	4.5	20	25.4	2.3	18	6.7%	-0.57 [-1.22 , 0.08]
Tang 2010	47.26	2.78	24	37.43	13.39	35	7.6%	0.92 [0.38 , 1.47]
Subtotal (95% CI)			**446**			**432**	**100.0%**	**0.36 [0.09 , 0.64]**
Heterogeneity: Tau² = 0.18; Chi² = 43.16, df = 13 (P < 0.0001); I² = 70%								
Test for overall effect: Z = 2.61 (P = 0.009)								
11.2.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	22.8	5.2	62	22.6	5.9	31	14.9%	0.04 [-0.39 , 0.47]
Courneya 2003c	25.3	2.5	24	25.3	2.8	28	12.5%	0.00 [-0.55 , 0.55]
Daley 2007a	25.22	2.79	31	23.74	3.91	31	13.4%	0.43 [-0.07 , 0.93]
Fillion 2008	46.76	9.24	44	44.64	11.1	43	15.2%	0.21 [-0.22 , 0.63]
Knols 2011	87.4	14.1	51	82.4	15.7	54	16.0%	0.33 [-0.05 , 0.72]
Tang 2010	50.78	16.6	24	35.03	11.3	35	12.2%	1.14 [0.57 , 1.70]
Thorsen 2005	85.9	13.1	52	88.8	11.8	49	15.8%	-0.23 [-0.62 , 0.16]
Subtotal (95% CI)			**288**			**271**	**100.0%**	**0.25 [-0.05 , 0.54]**
Heterogeneity: Tau² = 0.10; Chi² = 17.62, df = 6 (P = 0.007); I² = 66%								
Test for overall effect: Z = 1.66 (P = 0.10)								
11.2.3 6 months' follow-up								
Cadmus 2009	25.1	2.7	37	24	4.1	37	67.1%	0.31 [-0.15 , 0.77]
Rogers 2009	23.2	4.5	19	23.5	5.7	17	32.9%	-0.06 [-0.71 , 0.60]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.19 [-0.18 , 0.57]**
Heterogeneity: Tau² = 0.00; Chi² = 0.83, df = 1 (P = 0.36); I² = 0%								
Test for overall effect: Z = 1.00 (P = 0.32)								
11.2.4 More than 6 months' follow-up								
Speck 2010	52.4	7	58	51	7.8	62	100.0%	0.19 [-0.17 , 0.55]
Subtotal (95% CI)			**58**			**62**	**100.0%**	**0.19 [-0.17 , 0.55]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.02 (P = 0.31)								
Test for subgroup differences: Chi² = 0.84, df = 3 (P = 0.84), I² = 0%								



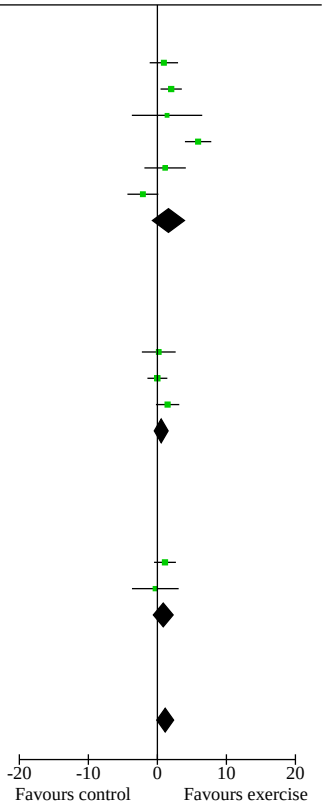
Analysis 11.3. Comparison 11: Physical functioning, Outcome 3: FACT physical function subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.3.1 Up to 12 weeks' follow-up								
Oh 2010	3.06	4	54	0.98	3.4	54	45.3%	2.08 [0.68 , 3.48]
Rogers 2009	2.3	5.1	20	2.6	6	18	19.1%	-0.30 [-3.86 , 3.26]
Targ 2002	0.88	6.2	79	1.28	7.1	88	35.6%	-0.40 [-2.42 , 1.62]
Subtotal (95% CI)			**153**			**160**	**100.0%**	**0.74 [-1.14 , 2.62]**
Heterogeneity: Tau² = 1.53; Chi² = 4.60, df = 2 (P = 0.10); I² = 56%								
Test for overall effect: Z = 0.77 (P = 0.44)								
11.3.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	-0.5	5.7	62	-0.3	5.7	31	38.3%	-0.20 [-2.66 , 2.26]
Courneya 2003c	1.9	2	24	-0.2	2.2	28	61.7%	2.10 [0.96 , 3.24]
Subtotal (95% CI)			**86**			**59**	**100.0%**	**1.22 [-0.97 , 3.41]**
Heterogeneity: Tau² = 1.69; Chi² = 2.77, df = 1 (P = 0.10); I² = 64%								
Test for overall effect: Z = 1.09 (P = 0.28)								
11.3.3 6 months' follow-up								
Cadmus 2009	0.1	2.8	37	-0.5	3.1	37	85.5%	0.60 [-0.75 , 1.95]
Rogers 2009	1.8	5.4	19	0.9	4.6	17	14.5%	0.90 [-2.37 , 4.17]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.64 [-0.60 , 1.89]**
Heterogeneity: Tau² = 0.00; Chi² = 0.03, df = 1 (P = 0.87); I² = 0%								
Test for overall effect: Z = 1.01 (P = 0.31)								
Test for subgroup differences: Chi² = 0.20, df = 2 (P = 0.90), I² = 0%								



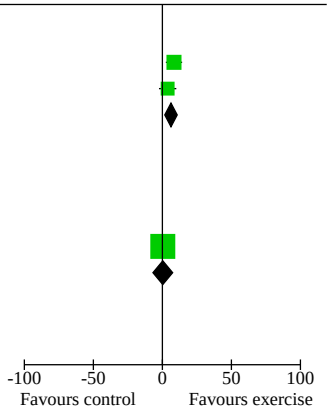
Analysis 11.4. Comparison 11: Physical functioning, Outcome 4: FACT physical function subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.4.1 Up to 12 weeks' follow-up								
Courneya 2003a	20.3	5.14	51	19.36	5.11	45	9.7%	0.94 [-1.11 , 2.99]
Daley 2007a	25.61	2.15	33	23.61	3.97	33	11.1%	2.00 [0.46 , 3.54]
Danhauer 2009	22.5	7.6	13	21.1	5.7	14	4.0%	1.40 [-3.70 , 6.50]
Milne 2008a	21.9	3	29	16	4.3	29	10.1%	5.90 [3.99 , 7.81]
Moadel 2007	21.03	6.05	45	19.91	6.3	26	7.4%	1.12 [-1.88 , 4.12]
Rogers 2009	23.3	4.5	20	25.4	2.3	18	9.3%	-2.10 [-4.34 , 0.14]
Subtotal (95% CI)			**191**			**165**	**51.6%**	**1.60 [-0.71 , 3.92]**
Heterogeneity: Tau² = 6.59; Chi² = 30.28, df = 5 (P < 0.0001); I² = 83%								
Test for overall effect: Z = 1.36 (P = 0.17)								
11.4.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	22.8	5.2	62	22.6	5.9	31	8.7%	0.20 [-2.25 , 2.65]
Courneya 2003c	25.3	2.5	24	25.3	2.8	28	11.3%	0.00 [-1.44 , 1.44]
Daley 2007a	25.22	2.79	31	23.74	3.91	31	10.7%	1.48 [-0.21 , 3.17]
Subtotal (95% CI)			**117**			**90**	**30.8%**	**0.55 [-0.45 , 1.55]**
Heterogeneity: Tau² = 0.00; Chi² = 1.80, df = 2 (P = 0.41); I² = 0%								
Test for overall effect: Z = 1.08 (P = 0.28)								
11.4.3 6 months' follow-up								
Cadmus 2009	25.1	2.7	37	24	4.1	37	11.0%	1.10 [-0.48 , 2.68]
Rogers 2009	23.2	4.5	19	23.5	5.7	17	6.6%	-0.30 [-3.68 , 3.08]
Subtotal (95% CI)			**56**			**54**	**17.6%**	**0.85 [-0.58 , 2.28]**
Heterogeneity: Tau² = 0.00; Chi² = 0.54, df = 1 (P = 0.46); I² = 0%								
Test for overall effect: Z = 1.16 (P = 0.25)								
Total (95% CI)			**364**			**309**	**100.0%**	**1.13 [-0.09 , 2.35]**
Heterogeneity: Tau² = 2.87; Chi² = 36.92, df = 10 (P < 0.0001); I² = 73%								
Test for overall effect: Z = 1.82 (P = 0.07)								
Test for subgroup differences: Chi² = 0.69, df = 2 (P = 0.71), I² = 0%								



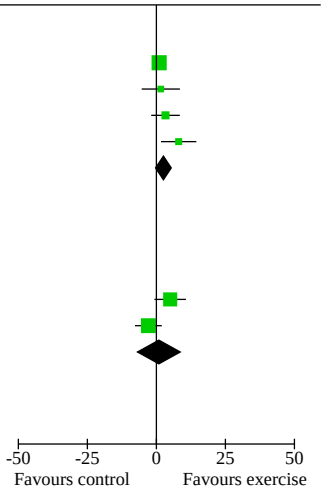
Analysis 11.5. Comparison 11: Physical functioning, Outcome 5: QLQ-C30 subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.5.1 More than 12 weeks' less than 6 months' follow-up								
Herrero 2006	6.7	5	8	-1.7	6.9	8	52.7%	8.40 [2.50 , 14.30]
Mehnert 2011	5.3	10.92	30	1.49	13	27	47.3%	3.81 [-2.46 , 10.08]
Subtotal (95% CI)			**38**			**35**	**100.0%**	**6.23 [1.74 , 10.72]**
Heterogeneity: Tau² = 0.88; Chi² = 1.09, df = 1 (P = 0.30); I² = 8%								
Test for overall effect: Z = 2.72 (P = 0.007)								
11.5.2 More than 12 weeks' less than 6 months' follow-up								
Thorsen 2005	11	16.1	52	10.7	20.1	49	100.0%	0.30 [-6.83 , 7.43]
Subtotal (95% CI)			**52**			**49**	**100.0%**	**0.30 [-6.83 , 7.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.08 (P = 0.93)								
Test for subgroup differences: Chi² = 1.90, df = 1 (P = 0.17), I² = 47.5%								



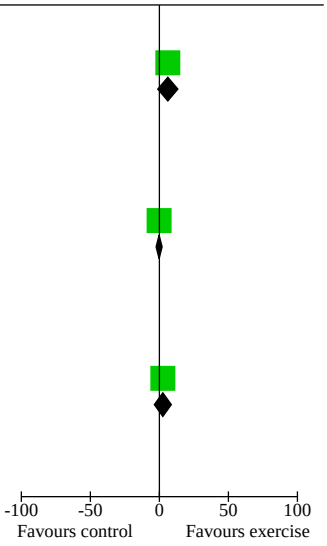
Analysis 11.6. Comparison 11: Physical functioning, Outcome 6: QLQ-C30 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.6.1 Up to 12 weeks' follow-up								
Dimeo 2004	76	1.6	34	75	1.6	35	53.2%	1.00 [0.24 , 1.76]
Herrero 2006	94.1	7.5	8	92.5	6.6	8	12.9%	1.60 [-5.32 , 8.52]
Knols 2011	83.7	14.2	57	80.4	14	57	19.4%	3.30 [-1.88 , 8.48]
Mehnert 2011	94	9.32	30	85.93	14.48	27	14.5%	8.07 [1.67 , 14.47]
Subtotal (95% CI)			**129**			**127**	**100.0%**	**2.55 [-0.29 , 5.38]**
Heterogeneity: Tau² = 3.77; Chi² = 5.32, df = 3 (P = 0.15); I² = 44%								
Test for overall effect: Z = 1.76 (P = 0.08)								
11.6.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	87.4	14.1	51	82.4	15.7	54	48.1%	5.00 [-0.70 , 10.70]
Thorsen 2005	85.9	13.1	52	88.8	11.8	49	51.9%	-2.90 [-7.76 , 1.96]
Subtotal (95% CI)			**103**			**103**	**100.0%**	**0.90 [-6.83 , 8.64]**
Heterogeneity: Tau² = 23.90; Chi² = 4.27, df = 1 (P = 0.04); I² = 77%								
Test for overall effect: Z = 0.23 (P = 0.82)								
Test for subgroup differences: Chi² = 0.15, df = 1 (P = 0.70), I² = 0%								



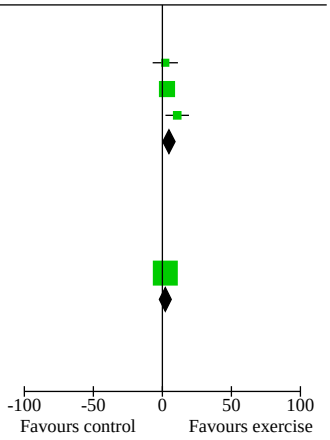
Analysis 11.7. Comparison 11: Physical functioning, Outcome 7: MOS SF-36 subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.7.1 Up to 12 weeks' follow-up								
Mehnert 2011	3.96	10.83	30	-2.15	16.7	28	100.0%	6.11 [-1.19 , 13.41]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**6.11 [-1.19 , 13.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.64 (P = 0.10)								
11.7.2 6 months' follow-up								
Cadmus 2009	-0.2	5.1	37	0	4.4	37	100.0%	-0.20 [-2.37 , 1.97]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.20 [-2.37 , 1.97]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.18 (P = 0.86)								
11.7.3 More than 6 months' follow-up								
Speck 2010	6.6	17.1	58	4.1	17.3	62	100.0%	2.50 [-3.66 , 8.66]
Subtotal (95% CI)			**58**			**62**	**100.0%**	**2.50 [-3.66 , 8.66]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.80 (P = 0.43)								
Test for subgroup differences: Chi² = 3.07, df = 2 (P = 0.22), I² = 34.9%								



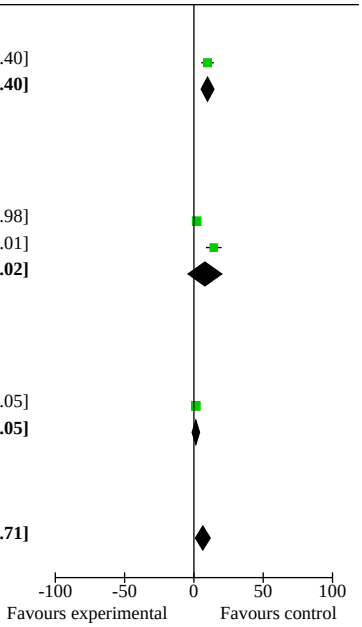
Analysis 11.8. Comparison 11: Physical functioning, Outcome 8: MOS SF-12 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.8.1 Up to 12 weeks' follow-up								
Danhauer 2009	44.8	12.4	13	42.7	11.8	14	19.7%	2.10 [-7.05 , 11.25]
Fillion 2008	45.1	10.42	44	41.76	9.76	43	58.0%	3.34 [-0.90 , 7.58]
Mehnert 2011	88.96	9.13	30	78.21	21.22	28	22.2%	10.75 [2.24 , 19.26]
Subtotal (95% CI)			**87**			**85**	**100.0%**	**4.74 [0.31 , 9.17]**
Heterogeneity: Tau² = 4.12; Chi² = 2.62, df = 2 (P = 0.27); I² = 24%								
Test for overall effect: Z = 2.10 (P = 0.04)								
11.8.2 More than 12 weeks' less than 6 months' follow-up								
Fillion 2008	46.76	9.24	44	44.64	11.1	43	100.0%	2.12 [-2.18 , 6.42]
Subtotal (95% CI)			**44**			**43**	**100.0%**	**2.12 [-2.18 , 6.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.97 (P = 0.33)								
Test for subgroup differences: Chi² = 0.69, df = 1 (P = 0.40), I² = 0%								



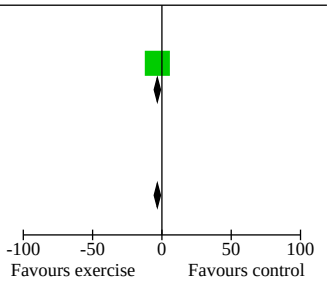
Analysis 11.9. Comparison 11: Physical functioning, Outcome 9: MOS SF-36 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.9.1 Up to 12 weeks' follow-up								
Tang 2010	47.26	2.78	24	37.43	13.39	35	24.0%	9.83 [5.26 , 14.40]
Subtotal (95% CI)			**24**			**35**	**24.0%**	**9.83 [5.26 , 14.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.21 (P < 0.0001)								
11.9.2 6 months' follow-up								
Cadmus 2009	50	6.4	37	48	7.6	48	26.8%	2.00 [-0.98 , 4.98]
Tang 2010	50.28	10.6	24	35.93	11.3	35	21.9%	14.35 [8.69 , 20.01]
Subtotal (95% CI)			**61**			**83**	**48.7%**	**7.93 [-4.16 , 20.02]**
Heterogeneity: Tau² = 70.94; Chi² = 14.33, df = 1 (P = 0.0002); I² = 93%								
Test for overall effect: Z = 1.29 (P = 0.20)								
11.9.3 More than 6 months' follow-up								
Speck 2010	52.4	7	58	51	7.8	62	27.3%	1.40 [-1.25 , 4.05]
Subtotal (95% CI)			**58**			**62**	**27.3%**	**1.40 [-1.25 , 4.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.04 (P = 0.30)								
Total (95% CI)			**143**			**180**	**100.0%**	**6.42 [1.14 , 11.71]**
Heterogeneity: Tau² = 24.83; Chi² = 24.54, df = 3 (P < 0.0001); I² = 88%								
Test for overall effect: Z = 2.38 (P = 0.02)								
Test for subgroup differences: Chi² = 10.27, df = 2 (P = 0.006), I² = 80.5%								



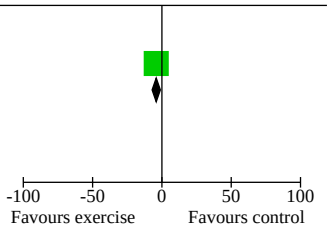
Analysis 11.10. Comparison 11: Physical functioning, Outcome 10: CARES subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.10.1 6 months' follow-up								
Ohira 2006	-2.1	5.6	39	1.2	4.5	40	100.0%	-3.30 [-5.54 , -1.06]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-3.30 [-5.54 , -1.06]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.88 (P = 0.004)								
Total (95% CI)			**39**			**40**	**100.0%**	**-3.30 [-5.54 , -1.06]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.88 (P = 0.004)								
Test for subgroup differences: Not applicable								



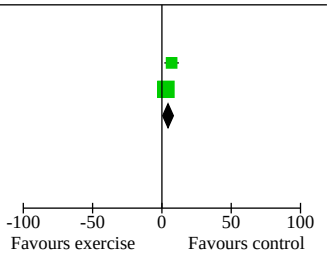
Analysis 11.11. Comparison 11: Physical functioning, Outcome 11: CARES subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.11.1 6 months' follow-up								
Ohira 2006	44.2	5.6	39	48.3	7.7	40	100.0%	-4.10 [-7.06 , -1.14]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-4.10 [-7.06 , -1.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.71 (P = 0.007)								
Test for subgroup differences: Not applicable								



Analysis 11.12. Comparison 11: Physical functioning, Outcome 12: Body Esteem Scale - physical condition follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
11.12.1 Up to 12 weeks' follow-up								
Pinto 2003	32.3	4.3	12	25.3	8.2	12	36.4%	7.00 [1.76 , 12.24]
Pinto 2005	30.26	7.88	39	27.33	6.76	43	63.6%	2.93 [-0.26 , 6.12]
Subtotal (95% CI)			**51**			**55**	**100.0%**	**4.41 [0.57 , 8.25]**
Heterogeneity: $Tau^2 = 3.38$; $Chi^2 = 1.69$, df = 1 (P = 0.19); $I^2 = 41\%$								
Test for overall effect: Z = 2.25 (P = 0.02)								
Test for subgroup differences: Not applicable								



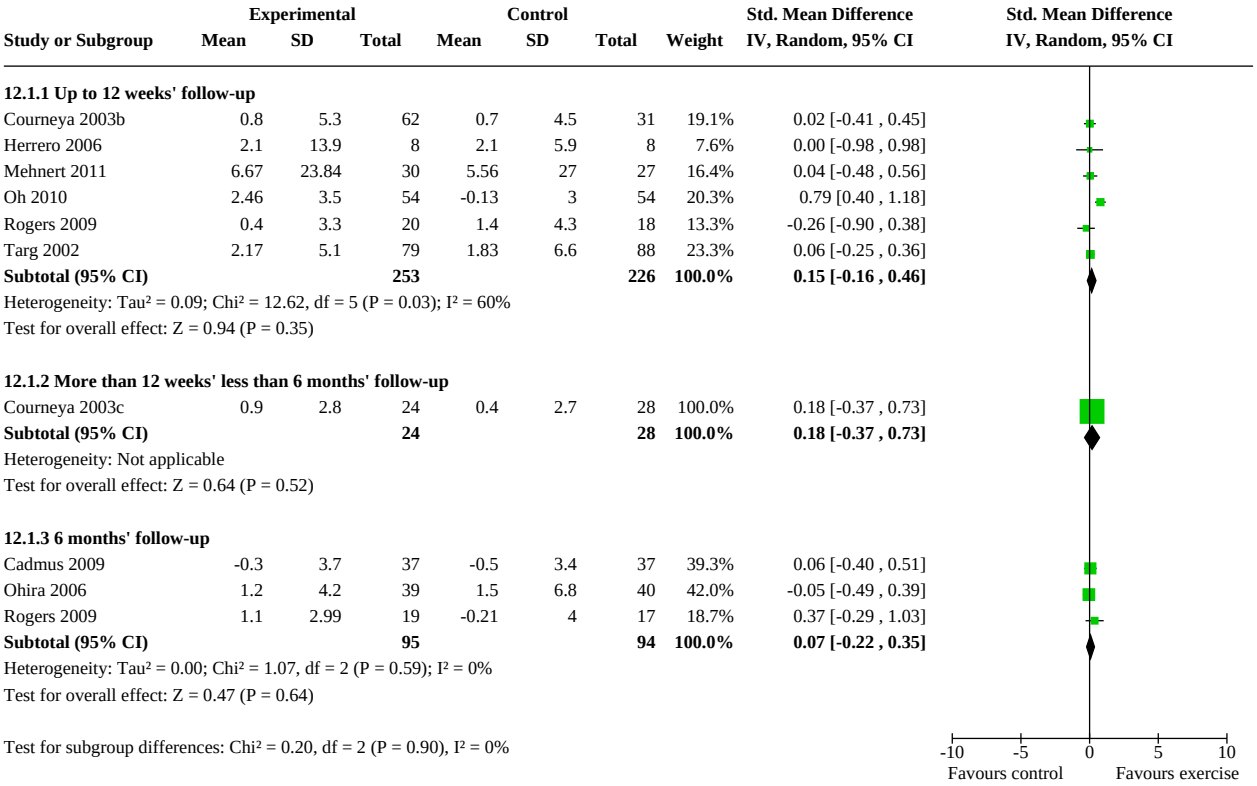
Comparison 12. Role function

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
12.1 Overall role function change	9		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
12.1.1 Up to 12 weeks' follow-up	6	479	Std. Mean Difference (IV, Random, 95% CI)	0.15 [-0.16, 0.46]
12.1.2 More than 12 weeks' less than 6 months' follow-up	1	52	Std. Mean Difference (IV, Random, 95% CI)	0.18 [-0.37, 0.73]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
12.1.3 6 months' follow-up	3	189	Std. Mean Difference (IV, Random, 95% CI)	0.07 [-0.22, 0.35]
12.2 Overall role function follow-up values	16		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
12.2.1 Up to 12 weeks' follow-up	12	719	Std. Mean Difference (IV, Random, 95% CI)	0.27 [0.04, 0.50]
12.2.2 More than 12 weeks' less than 6 months' follow-up	4	263	Std. Mean Difference (IV, Random, 95% CI)	0.18 [-0.06, 0.42]
12.2.3 6 months' follow-up	3	189	Std. Mean Difference (IV, Random, 95% CI)	0.16 [-0.13, 0.44]
12.3 FACT functional well-being subscale change	6		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.3.1 Up to 12 weeks' follow-up	4	406	Mean Difference (IV, Random, 95% CI)	0.70 [-0.96, 2.36]
12.3.2 More than 12 weeks' less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	0.50 [-1.00, 2.00]
12.3.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	0.56 [-0.77, 1.89]
12.4 FACT functional well-being follow-up values	9		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.4.1 Up to 12 weeks' follow-up	7	449	Mean Difference (IV, Random, 95% CI)	1.63 [-0.43, 3.70]
12.4.2 More than 12 weeks' less than 6 months' follow-up	2	113	Mean Difference (IV, Random, 95% CI)	0.74 [-1.03, 2.51]
12.4.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	0.91 [-1.02, 2.84]
12.5 QLQ-C30 role functioning change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.5.1 Up to 12 weeks' follow-up	2	73	Mean Difference (IV, Random, 95% CI)	0.42 [-7.80, 8.65]
12.6 QLQ-C30 role functioning subscale follow-up values	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.6.1 Up to 12 weeks' follow-up	4	256	Mean Difference (IV, Random, 95% CI)	7.23 [0.59, 13.87]
12.6.2 More than 12 weeks' less than 6 months' follow-up	2	150	Mean Difference (IV, Random, 95% CI)	4.63 [-3.55, 12.80]
12.7 MOS SF-36 role function subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.7.1 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-0.40 [-5.23, 4.43]
12.8 MOS SF-36 role function subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.8.1 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	0.20 [-3.71, 4.11]

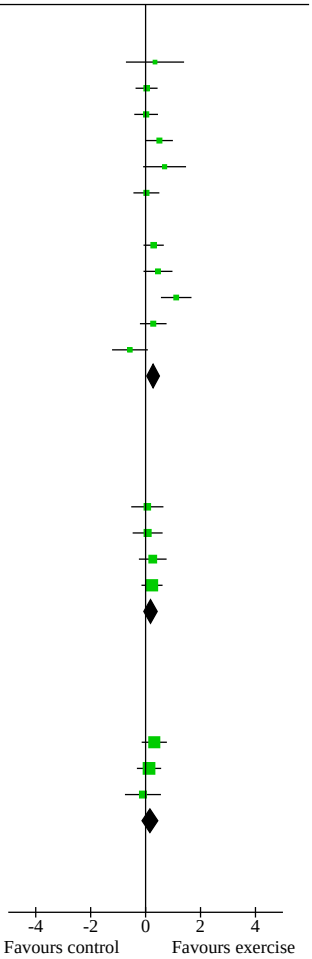
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
12.9 CARES marital subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.9.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	0.30 [-2.19, 2.79]
12.10 CARES marital subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.10.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	0.80 [-2.00, 3.60]

Analysis 12.1. Comparison 12: Role function, Outcome 1: Overall role function change



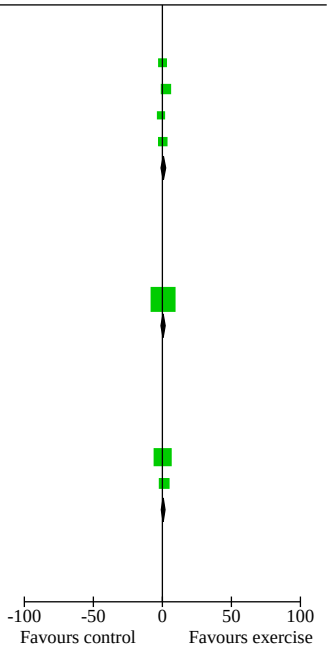
Analysis 12.2. Comparison 12: Role function, Outcome 2: Overall role function follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.2.1 Up to 12 weeks' follow-up								
Banasik 2011	3.46	0.64	7	3.24	0.56	7	3.7%	0.34 [-0.72 , 1.40]
Courneya 2003a	16.58	5.3	51	16.39	5.45	45	11.6%	0.04 [-0.37 , 0.44]
Courneya 2003b	21.5	5.1	62	21.4	5.3	31	11.0%	0.02 [-0.41 , 0.45]
Daley 2007a	22.88	4.52	33	20.3	5.57	33	9.9%	0.50 [0.01 , 0.99]
Danhauer 2009	21.9	4.7	13	17.4	7.5	14	5.8%	0.69 [-0.09 , 1.47]
Dimeo 2004	62	32	34	61	36	35	10.2%	0.03 [-0.44 , 0.50]
Herrero 2006	97.9	5.9	8	100	0	8		Not estimable
Knols 2011	68.7	24.2	57	61.3	26	57	12.3%	0.29 [-0.08 , 0.66]
Mehnert 2011	90	20.34	30	79.63	25.04	27	9.2%	0.45 [-0.08 , 0.98]
Milne 2008a	19.5	4.5	29	13.7	5.7	29	8.8%	1.11 [0.56 , 1.67]
Moadel 2007	18.17	6.74	45	16.21	7.34	26	10.0%	0.28 [-0.21 , 0.76]
Rogers 2009	21.4	4.8	20	23.8	3.2	18	7.4%	-0.57 [-1.22 , 0.08]
Subtotal (95% CI)			**389**			**330**	**100.0%**	**0.27 [0.04 , 0.50]**
Heterogeneity: Tau² = 0.08; Chi² = 21.32, df = 10 (P = 0.02); I² = 53%								
Test for overall effect: Z = 2.33 (P = 0.02)								
12.2.2 More than 12 weeks' less than 6 months' follow-up								
Bai 2004	62.28	23.9	24	60.71	25	21	17.2%	0.06 [-0.52 , 0.65]
Courneya 2003c	23.4	4	24	23.1	4	28	19.8%	0.07 [-0.47 , 0.62]
Daley 2007a	22.29	5.71	31	20.71	6.27	30	23.2%	0.26 [-0.24 , 0.76]
Knols 2011	76.3	26.9	51	70.2	25	54	39.9%	0.23 [-0.15 , 0.62]
Subtotal (95% CI)			**130**			**133**	**100.0%**	**0.18 [-0.06 , 0.42]**
Heterogeneity: Tau² = 0.00; Chi² = 0.47, df = 3 (P = 0.93); I² = 0%								
Test for overall effect: Z = 1.45 (P = 0.15)								
12.2.3 6 months' follow-up								
Cadmus 2009	23.1	4.7	37	21.5	5.3	37	38.9%	0.32 [-0.14 , 0.77]
Ohira 2006	-49.2	6.5	39	-50	6.2	40	42.0%	0.12 [-0.32 , 0.57]
Rogers 2009	21.9	4.7	19	22.4	5.4	17	19.1%	-0.10 [-0.75 , 0.56]
Subtotal (95% CI)			**95**			**94**	**100.0%**	**0.16 [-0.13 , 0.44]**
Heterogeneity: Tau² = 0.00; Chi² = 1.06, df = 2 (P = 0.59); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 0.49, df = 2 (P = 0.78), I² = 0%								



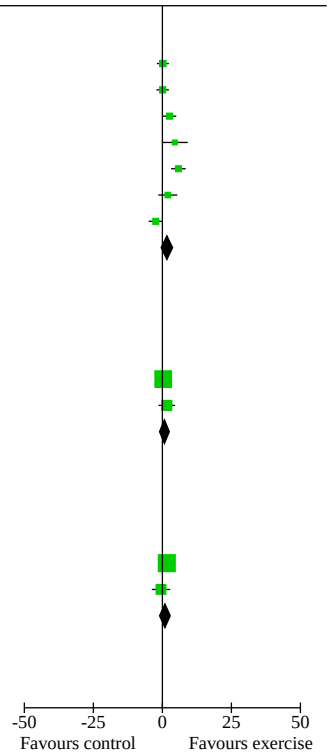
Analysis 12.3. Comparison 12: Role function, Outcome 3: FACT functional well-being subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.3.1 Up to 12 weeks' follow-up								
Courneya 2003b	0.8	5.3	62	0.7	4.5	31	23.4%	0.10 [-1.96 , 2.16]
Oh 2010	2.46	3.5	54	-0.13	3	54	30.5%	2.59 [1.36 , 3.82]
Rogers 2009	0.4	3.3	20	1.4	4.3	18	20.3%	-1.00 [-3.46 , 1.46]
Targ 2002	2.17	5.1	79	1.83	6.6	88	25.8%	0.34 [-1.44 , 2.12]
Subtotal (95% CI)			**215**			**191**	**100.0%**	**0.70 [-0.96 , 2.36]**
Heterogeneity: Tau² = 1.96; Chi² = 9.98, df = 3 (P = 0.02); I² = 70%								
Test for overall effect: Z = 0.82 (P = 0.41)								
12.3.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	0.9	2.8	24	0.4	2.7	28	100.0%	0.50 [-1.00 , 2.00]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**0.50 [-1.00 , 2.00]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.65 (P = 0.51)								
12.3.3 6 months' follow-up								
Cadmus 2009	-0.3	3.7	37	-0.5	3.4	37	67.4%	0.20 [-1.42 , 1.82]
Rogers 2009	1.1	2.99	19	-0.21	4	17	32.6%	1.31 [-1.02 , 3.64]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.56 [-0.77 , 1.89]**
Heterogeneity: Tau² = 0.00; Chi² = 0.59, df = 1 (P = 0.44); I² = 0%								
Test for overall effect: Z = 0.83 (P = 0.41)								
Test for subgroup differences: Chi² = 0.03, df = 2 (P = 0.98), I² = 0%								



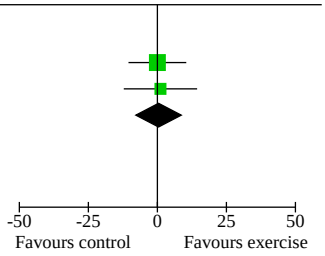
Analysis 12.4. Comparison 12: Role function, Outcome 4: FACT functional well-being follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.4.1 Up to 12 weeks' follow-up								
Courneya 2003a	16.58	5.3	51	16.39	5.45	45	16.2%	0.19 [-1.97 , 2.35]
Courneya 2003b	21.5	5.1	62	21.4	5.3	31	15.9%	0.10 [-2.16 , 2.36]
Daley 2007a	22.88	4.52	33	20.3	5.57	33	15.4%	2.58 [0.13 , 5.03]
Danhauer 2009	21.9	4.7	13	17.4	7.5	14	9.8%	4.50 [-0.19 , 9.19]
Milne 2008a	19.5	4.5	29	13.7	5.7	29	14.9%	5.80 [3.16 , 8.44]
Moadel 2007	18.17	6.74	45	16.21	7.34	26	12.7%	1.96 [-1.48 , 5.40]
Rogers 2009	21.4	4.8	20	23.8	3.2	18	15.1%	-2.40 [-4.97 , 0.17]
Subtotal (95% CI)			**253**			**196**	**100.0%**	**1.63 [-0.43 , 3.70]**
Heterogeneity: Tau² = 5.65; Chi² = 24.17, df = 6 (P = 0.0005); I² = 75%								
Test for overall effect: Z = 1.55 (P = 0.12)								
12.4.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	23.4	4	24	23.1	4	28	65.6%	0.30 [-1.88 , 2.48]
Daley 2007a	22.29	5.71	31	20.71	6.27	30	34.4%	1.58 [-1.43 , 4.59]
Subtotal (95% CI)			**55**			**58**	**100.0%**	**0.74 [-1.03 , 2.51]**
Heterogeneity: Tau² = 0.00; Chi² = 0.46, df = 1 (P = 0.50); I² = 0%								
Test for overall effect: Z = 0.82 (P = 0.41)								
12.4.3 6 months' follow-up								
Cadmus 2009	23.1	4.7	37	21.5	5.3	37	67.3%	1.60 [-0.68 , 3.88]
Rogers 2009	21.9	4.7	19	22.4	5.4	17	32.7%	-0.50 [-3.82 , 2.82]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.91 [-1.02 , 2.84]**
Heterogeneity: Tau² = 0.09; Chi² = 1.04, df = 1 (P = 0.31); I² = 4%								
Test for overall effect: Z = 0.93 (P = 0.35)								
Test for subgroup differences: Chi² = 0.45, df = 2 (P = 0.80), I² = 0%								



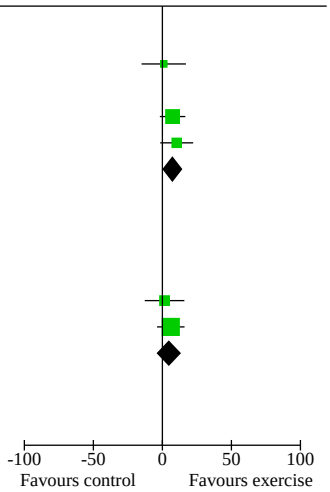
Analysis 12.5. Comparison 12: Role function, Outcome 5: QLQ-C30 role functioning change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.5.1 Up to 12 weeks' follow-up								
Herrero 2006	2.1	13.9	8	2.1	5.9	8	61.7%	0.00 [-10.46 , 10.46]
Mehnert 2011	6.67	23.84	30	5.56	27	27	38.3%	1.11 [-12.18 , 14.40]
Subtotal (95% CI)			**38**			**35**	**100.0%**	**0.42 [-7.80 , 8.65]**
Heterogeneity: Tau² = 0.00; Chi² = 0.02, df = 1 (P = 0.90); I² = 0%								
Test for overall effect: Z = 0.10 (P = 0.92)								
Test for subgroup differences: Not applicable								



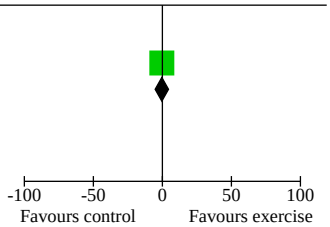
Analysis 12.6. Comparison 12: Role function, Outcome 6: QLQ-C30 role functioning subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.6.1 Up to 12 weeks' follow-up								
Dimeo 2004	62	32	34	61	36	35	17.1%	1.00 [-15.06 , 17.06]
Herrero 2006	97.9	5.9	8	100	0	8		Not estimable
Knols 2011	68.7	24.2	57	61.3	26	57	51.9%	7.40 [-1.82 , 16.62]
Mehnert 2011	90	20.34	30	79.63	25.04	27	31.0%	10.37 [-1.55 , 22.29]
Subtotal (95% CI)			**129**			**127**	**100.0%**	**7.23 [0.59 , 13.87]**
Heterogeneity: Tau² = 0.00; Chi² = 0.85, df = 2 (P = 0.66); I² = 0%								
Test for overall effect: Z = 2.13 (P = 0.03)								
12.6.2 More than 12 weeks' less than 6 months' follow-up								
Bai 2004	62.28	23.9	24	60.71	25	21	32.5%	1.57 [-12.77 , 15.91]
Knols 2011	76.3	26.9	51	70.2	25	54	67.5%	6.10 [-3.85 , 16.05]
Subtotal (95% CI)			**75**			**75**	**100.0%**	**4.63 [-3.55 , 12.80]**
Heterogeneity: Tau² = 0.00; Chi² = 0.26, df = 1 (P = 0.61); I² = 0%								
Test for overall effect: Z = 1.11 (P = 0.27)								
Test for subgroup differences: Chi² = 0.23, df = 1 (P = 0.63), I² = 0%								



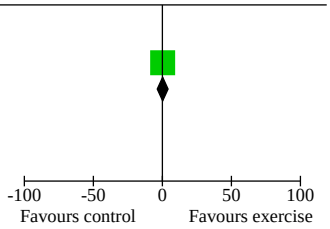
Analysis 12.7. Comparison 12: Role function, Outcome 7: MOS SF-36 role function subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.7.1 6 months' follow-up								
Cadmus 2009	1.5	11.9	37	1.9	9.1	37	100.0%	-0.40 [-5.23 , 4.43]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-0.40 [-5.23 , 4.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.16 (P = 0.87)								
Test for subgroup differences: Not applicable								



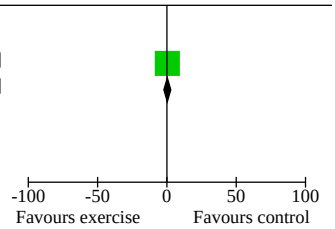
Analysis 12.8. Comparison 12: Role function, Outcome 8: MOS SF-36 role function subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.8.1 6 months' follow-up								
Cadmus 2009	49.7	9.3	37	49.5	7.8	37	100.0%	0.20 [-3.71 , 4.11]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.20 [-3.71 , 4.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.10 (P = 0.92)								
Test for subgroup differences: Not applicable								



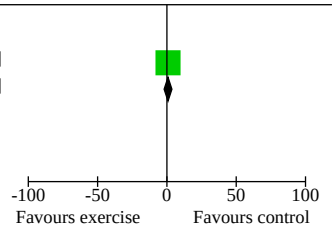
Analysis 12.9. Comparison 12: Role function, Outcome 9: CARES marital subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.9.1 6 months' follow-up								
Ohira 2006	-1.2	4.2	39	-1.5	6.8	40	100.0%	0.30 [-2.19 , 2.79]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**0.30 [-2.19 , 2.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
Test for subgroup differences: Not applicable								



Analysis 12.10. Comparison 12: Role function, Outcome 10: CARES marital subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
12.10.1 6 months' follow-up								
Ohira 2006	-49.2	6.5	39	-50	6.2	40	100.0%	0.80 [-2.00 , 3.60]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**0.80 [-2.00 , 3.60]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.56 (P = 0.58)								
Test for subgroup differences: Not applicable								

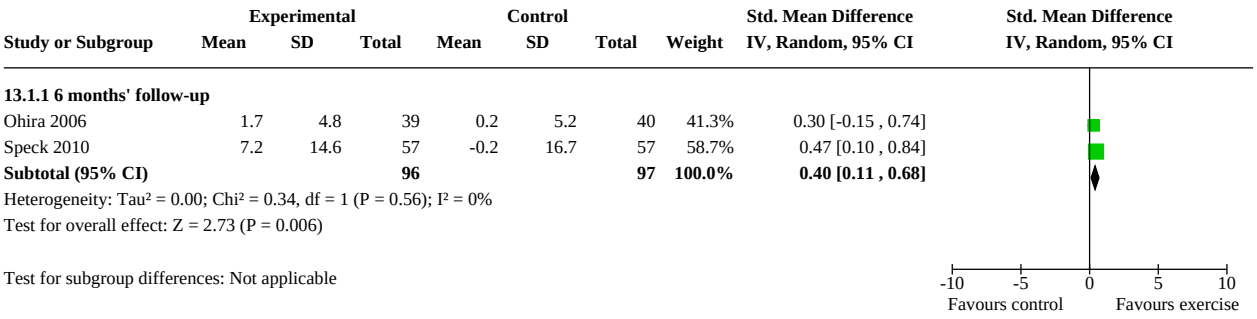


Comparison 13. Sexuality

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
13.1 Overall sexuality change	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
13.1.1 6 months' follow-up	2	193	Std. Mean Difference (IV, Random, 95% CI)	0.40 [0.11, 0.68]
13.2 Overall sexuality follow-up values	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only

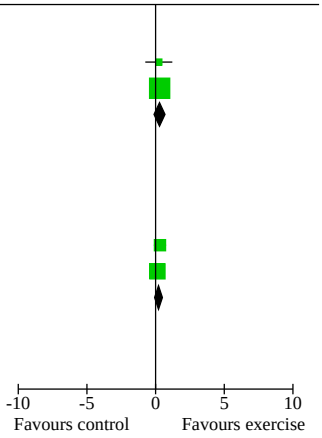
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
13.2.1 Up to 12 weeks' follow-up	2	100	Std. Mean Difference (IV, Random, 95% CI)	0.28 [-0.11, 0.68]
13.2.2 6 months' follow-up	2	193	Std. Mean Difference (IV, Random, 95% CI)	0.21 [-0.07, 0.49]
13.3 Body Esteem Scale - sexual attractiveness follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.3.1 Up to 12 weeks' follow-up	2	100	Mean Difference (IV, Random, 95% CI)	2.28 [-0.83, 5.39]
13.4 CARE sexuality change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.4.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-1.50 [-3.71, 0.71]
13.5 CARE sexuality follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.5.1 6 months' follow-up	1	79	Mean Difference (IV, Random, 95% CI)	-2.50 [-5.92, 0.92]
13.6 BIRS appearance and sexuality subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.6.1 More than 6 months' follow-up	1	114	Mean Difference (IV, Random, 95% CI)	7.40 [1.64, 13.16]
13.7 BIRS appearance and sexuality follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.7.1 More than 6 months' follow-up	1	114	Mean Difference (IV, Random, 95% CI)	-0.80 [-2.92, 1.32]

Analysis 13.1. Comparison 13: Sexuality, Outcome 1: Overall sexuality change



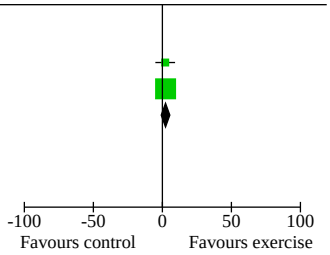
Analysis 13.2. Comparison 13: Sexuality, Outcome 2: Overall sexuality follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
13.2.1 Up to 12 weeks' follow-up								
Pinto 2003	45.75	9.2	12	43.7	6.1	6	16.4%	0.23 [-0.75 , 1.22]
Pinto 2005	42.45	9.24	39	40.12	6.3	43	83.6%	0.29 [-0.14 , 0.73]
Subtotal (95% CI)			**51**			**49**	**100.0%**	**0.28 [-0.11 , 0.68]**
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.91); I² = 0%								
Test for overall effect: Z = 1.40 (P = 0.16)								
13.2.2 6 months' follow-up								
Ohira 2006	-51	7.5	39	-53.5	8	40	40.7%	0.32 [-0.12 , 0.76]
Speck 2010	-27.3	5.3	57	-28.1	6.2	57	59.3%	0.14 [-0.23 , 0.51]
Subtotal (95% CI)			**96**			**97**	**100.0%**	**0.21 [-0.07 , 0.49]**
Heterogeneity: Tau² = 0.00; Chi² = 0.38, df = 1 (P = 0.54); I² = 0%								
Test for overall effect: Z = 1.46 (P = 0.14)								
Test for subgroup differences: Chi² = 0.09, df = 1 (P = 0.77), I² = 0%								



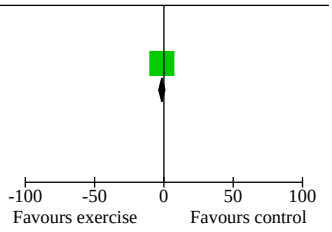
Analysis 13.3. Comparison 13: Sexuality, Outcome 3: Body Esteem Scale - sexual attractiveness follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
13.3.1 Up to 12 weeks' follow-up								
Pinto 2003	45.75	9.2	12	43.7	6.1	6	19.0%	2.05 [-5.09 , 9.19]
Pinto 2005	42.45	9.24	39	40.12	6.3	43	81.0%	2.33 [-1.13 , 5.79]
Subtotal (95% CI)			**51**			**49**	**100.0%**	**2.28 [-0.83 , 5.39]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.94); I² = 0%								
Test for overall effect: Z = 1.43 (P = 0.15)								
Test for subgroup differences: Not applicable								



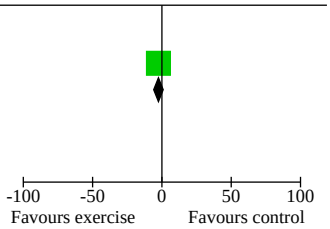
Analysis 13.4. Comparison 13: Sexuality, Outcome 4: CARE sexuality change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
13.4.1 6 months' follow-up								
Ohira 2006	-1.7	4.8	39	-0.2	5.2	40	100.0%	-1.50 [-3.71 , 0.71]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-1.50 [-3.71 , 0.71]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.33 (P = 0.18)								
Test for subgroup differences: Not applicable								



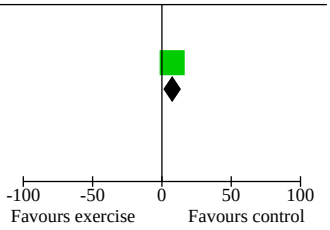
Analysis 13.5. Comparison 13: Sexuality, Outcome 5: CARE sexuality follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
13.5.1 6 months' follow-up								
Ohira 2006	51	7.5	39	53.5	8	40	100.0%	-2.50 [-5.92 , 0.92]
Subtotal (95% CI)			**39**			**40**	**100.0%**	**-2.50 [-5.92 , 0.92]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.43 (P = 0.15)								
Test for subgroup differences: Not applicable								



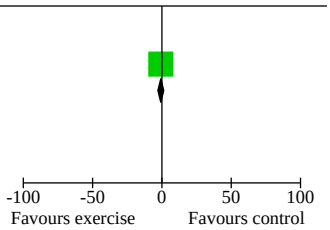
Analysis 13.6. Comparison 13: Sexuality, Outcome 6: BIRS appearance and sexuality subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
13.6.1 More than 6 months' follow-up								
Speck 2010	7.2	14.6	57	-0.2	16.7	57	100.0%	7.40 [1.64 , 13.16]
Subtotal (95% CI)			**57**			**57**	**100.0%**	**7.40 [1.64 , 13.16]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.52 (P = 0.01)								
Test for subgroup differences: Not applicable								



Analysis 13.7. Comparison 13: Sexuality, Outcome 7: BIRS appearance and sexuality follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
13.7.1 More than 6 months' follow-up								
Speck 2010	27.3	5.3	57	28.1	6.2	57	100.0%	-0.80 [-2.92 , 1.32]
Subtotal (95% CI)			**57**			**57**	**100.0%**	**-0.80 [-2.92 , 1.32]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.74 (P = 0.46)								
Test for subgroup differences: Not applicable								



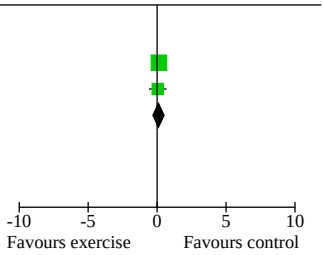
Comparison 14. Sleep disturbances

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
14.1 Overall sleep disturbance change	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
14.1.1 Up to 12 weeks' follow-up	2	95	Std. Mean Difference (IV, Random, 95% CI)	0.10 [-0.30, 0.50]
14.2 Overall sleep disturbance follow-up values	8		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
14.2.1 Up to 12 weeks' follow-up	8	438	Std. Mean Difference (IV, Random, 95% CI)	-0.46 [-0.72, -0.20]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
14.2.2 More than 12 weeks' less than 6 months' follow-up	2	164	Std. Mean Difference (IV, Random, 95% CI)	-0.52 [-1.64, 0.61]
14.3 Pittsburgh Sleep Quality Index change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
14.3.1 Up to 12 weeks' follow-up	1	38	Mean Difference (IV, Random, 95% CI)	0.17 [-2.19, 2.53]
14.4 Pittsburg Sleep Quality Index follow-up values	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
14.4.1 Up to 12 weeks' follow-up	5	199	Mean Difference (IV, Random, 95% CI)	-1.55 [-3.12, 0.02]
14.4.2 More than 12 weeks' less than 6 months' follow-up	1	59	Mean Difference (IV, Random, 95% CI)	-3.33 [-4.88, -1.78]
14.5 QLQ-C30 subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
14.5.1 Up to 12 weeks' follow-up	1	57	Mean Difference (IV, Random, 95% CI)	4.32 [-12.18, 20.82]
14.6 QLQ-C30 subscale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
14.6.1 Up to 12 weeks' follow-up	2	183	Mean Difference (IV, Random, 95% CI)	-3.11 [-4.66, -1.57]
14.6.2 More than 12 weeks' less than 6 months' follow-up	1	105	Mean Difference (IV, Random, 95% CI)	1.30 [-10.72, 13.32]

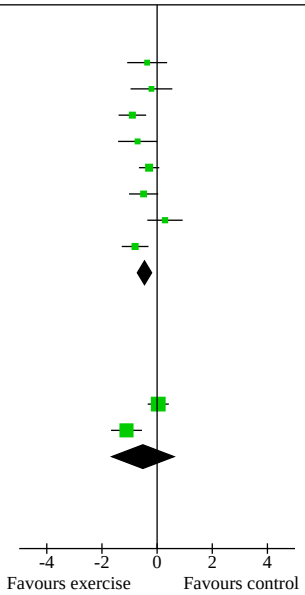
Analysis 14.1. Comparison 14: Sleep disturbances, Outcome 1: Overall sleep disturbance change

	Experimental			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
14.1.1 Up to 12 weeks' follow-up									
Mehnert 2011	-5.56	29.88	30	-9.88	33.3	27	59.9%	0.14 [-0.39 , 0.66]	
Rogers 2009	0.45	2.1	20	0.28	4.7	18	40.1%	0.05 [-0.59 , 0.68]	
Subtotal (95% CI)			**50**			**45**	**100.0%**	**0.10 [-0.30 , 0.50]**	
Heterogeneity: Tau² = 0.00; Chi² = 0.04, df = 1 (P = 0.83); I² = 0%									
Test for overall effect: Z = 0.48 (P = 0.63)									
Test for subgroup differences: Not applicable									



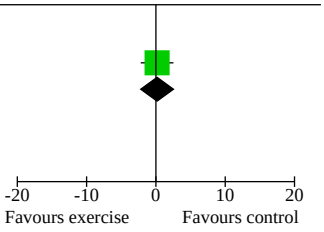
Analysis 14.2. Comparison 14: Sleep disturbances, Outcome 2: Overall sleep disturbance follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
14.2.1 Up to 12 weeks' follow-up								
Cohen 2004	5.82	8.1	16	8.1	2.4	14	9.1%	-0.36 [-1.08 , 0.36]
Danhauer 2009	6.1	4.3	13	7	4.2	14	8.5%	-0.21 [-0.96 , 0.55]
Dimeo 2004	31	3.1	34	34	3.5	34	14.5%	-0.90 [-1.40 , -0.40]
Donnelly 2011	6.88	3.96	16	10.06	4.75	17	9.4%	-0.71 [-1.41 , -0.00]
Knols 2011	23	27.4	57	31.5	30.8	57	19.2%	-0.29 [-0.66 , 0.08]
Mehnert 2011	23.73	27	30	38.27	31.63	27	13.6%	-0.49 [-1.02 , 0.04]
Rogers 2009	6.7	4.2	20	5.5	4	18	10.7%	0.29 [-0.35 , 0.93]
Tang 2010	10.67	3.23	36	13.09	2.74	35	15.0%	-0.80 [-1.28 , -0.31]
Subtotal (95% CI)			**222**			**216**	**100.0%**	**-0.46 [-0.72 , -0.20]**
Heterogeneity: Tau² = 0.05; Chi² = 11.83, df = 7 (P = 0.11); I² = 41%								
Test for overall effect: Z = 3.48 (P = 0.0005)								
14.2.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	26.3	31.9	51	25	30.9	54	51.6%	0.04 [-0.34 , 0.42]
Tang 2010	9.78	3.06	24	13.11	2.89	35	48.4%	-1.11 [-1.67 , -0.55]
Subtotal (95% CI)			**75**			**89**	**100.0%**	**-0.52 [-1.64 , 0.61]**
Heterogeneity: Tau² = 0.60; Chi² = 11.09, df = 1 (P = 0.0009); I² = 91%								
Test for overall effect: Z = 0.90 (P = 0.37)								
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.92), I² = 0%								



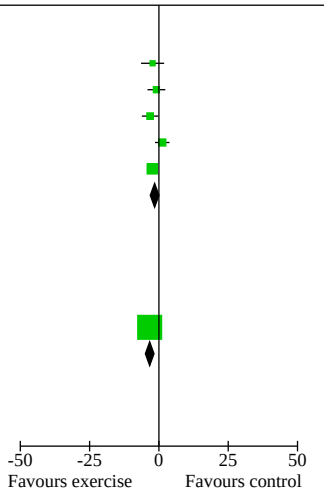
Analysis 14.3. Comparison 14: Sleep disturbances, Outcome 3: Pittsburgh Sleep Quality Index change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
14.3.1 Up to 12 weeks' follow-up								
Rogers 2009	0.45	2.1	20	0.28	4.7	18	100.0%	0.17 [-2.19 , 2.53]
Subtotal (95% CI)			**20**			**18**	**100.0%**	**0.17 [-2.19 , 2.53]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.14 (P = 0.89)								
Test for subgroup differences: Not applicable								



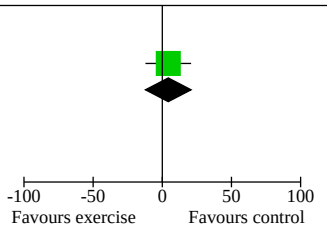
Analysis 14.4. Comparison 14: Sleep disturbances, Outcome 4: Pittsburg Sleep Quality Index follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
14.4.1 Up to 12 weeks' follow-up								
Cohen 2004	5.82	8.1	16	8.1	2.4	14	10.9%	-2.28 [-6.44 , 1.88]
Danhauer 2009	6.1	4.3	13	7	4.2	14	15.9%	-0.90 [-4.11 , 2.31]
Donnelly 2011	6.88	3.96	16	10.06	4.75	17	17.6%	-3.18 [-6.16 , -0.20]
Rogers 2009	6.7	4.2	20	5.5	4	18	20.6%	1.20 [-1.41 , 3.81]
Tang 2010	10.67	3.23	36	13.09	2.74	35	34.9%	-2.42 [-3.81 , -1.03]
Subtotal (95% CI)			**101**			**98**	**100.0%**	**-1.55 [-3.12 , 0.02]**
Heterogeneity: Tau² = 1.32; Chi² = 7.02, df = 4 (P = 0.13); I² = 43%								
Test for overall effect: Z = 1.94 (P = 0.05)								
14.4.2 More than 12 weeks' less than 6 months' follow-up								
Tang 2010	9.78	3.06	24	13.11	2.89	35	100.0%	-3.33 [-4.88 , -1.78]
Subtotal (95% CI)			**24**			**35**	**100.0%**	**-3.33 [-4.88 , -1.78]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.20 (P < 0.0001)								
Test for subgroup differences: Chi² = 2.50, df = 1 (P = 0.11), I² = 60.0%								



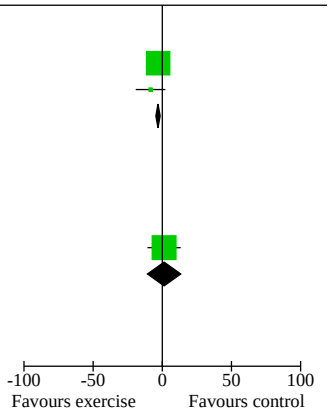
Analysis 14.5. Comparison 14: Sleep disturbances, Outcome 5: QLQ-C30 subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
14.5.1 Up to 12 weeks' follow-up								
Mehnert 2011	-5.56	29.88	30	-9.88	33.3	27	100.0%	4.32 [-12.18 , 20.82]
Subtotal (95% CI)			**30**			**27**	**100.0%**	**4.32 [-12.18 , 20.82]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.51 (P = 0.61)								
Test for subgroup differences: Not applicable								



Analysis 14.6. Comparison 14: Sleep disturbances, Outcome 6: QLQ-C30 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
14.6.1 Up to 12 weeks' follow-up								
Dimeo 2004	31	3.1	34	34	3.5	35	97.9%	-3.00 [-4.56 , -1.44]
Knols 2011	23	27.4	57	31.5	30.8	57	2.1%	-8.50 [-19.20 , 2.20]
Subtotal (95% CI)			**91**			**92**	**100.0%**	**-3.11 [-4.66 , -1.57]**
Heterogeneity: Tau² = 0.00; Chi² = 0.99, df = 1 (P = 0.32); I² = 0%								
Test for overall effect: Z = 3.96 (P < 0.0001)								
14.6.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	26.3	31.9	51	25	30.9	54	100.0%	1.30 [-10.72 , 13.32]
Subtotal (95% CI)			**51**			**54**	**100.0%**	**1.30 [-10.72 , 13.32]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.21 (P = 0.83)								
Test for subgroup differences: Chi² = 0.51, df = 1 (P = 0.48), I² = 0%								



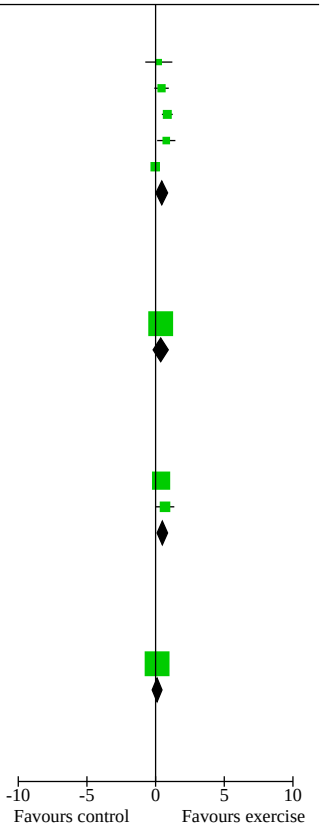
Comparison 15. Social functioning

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
15.1 Overall social functioning change	8		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
15.1.1 Up to 12 weeks' follow-up	5	386	Std. Mean Difference (IV, Random, 95% CI)	0.45 [0.02, 0.87]
15.1.2 More than 12 weeks' less than 6 months' follow-up	1	52	Std. Mean Difference (IV, Random, 95% CI)	0.37 [-0.18, 0.92]
15.1.3 6 months' follow-up	2	110	Std. Mean Difference (IV, Random, 95% CI)	0.49 [0.11, 0.87]
15.1.4 More than 6 months' follow-up	1	121	Std. Mean Difference (IV, Random, 95% CI)	0.11 [-0.25, 0.47]
15.2 Overall social functioning follow-up values	14		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
15.2.1 Up to 12 weeks' follow-up	10	530	Std. Mean Difference (IV, Random, 95% CI)	0.23 [-0.02, 0.48]
15.2.2 More than 12 weeks' less than 6 months' follow-up	4	312	Std. Mean Difference (IV, Random, 95% CI)	0.08 [-0.16, 0.31]
15.2.3 6 months' follow-up	2	110	Std. Mean Difference (IV, Random, 95% CI)	0.33 [-0.04, 0.71]
15.2.4 More than 6 months' follow-up	1	121	Std. Mean Difference (IV, Random, 95% CI)	0.14 [-0.21, 0.50]
15.3 FACT social functioning subscale change	5		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.3.1 Up to 12 weeks' follow-up	3	313	Mean Difference (IV, Random, 95% CI)	1.73 [-0.33, 3.79]
15.3.2 More than 12 weeks' less than 6 months' follow-up	1	52	Mean Difference (IV, Random, 95% CI)	1.10 [-0.45, 2.65]
15.3.3 6 months' follow-up	2	110	Mean Difference (IV, Random, 95% CI)	2.14 [0.25, 4.02]
15.4 FACT social functioning subscale follow-up values	8		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.4.1 Up to 12 weeks' follow-up	5	261	Mean Difference (IV, Random, 95% CI)	1.77 [0.56, 2.97]
15.4.2 More than 12 weeks' less than 6 months' follow-up	3	207	Mean Difference (IV, Random, 95% CI)	0.19 [-1.20, 1.58]
15.4.3 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	1.70 [-1.11, 4.51]
15.5 QLQ-C30 social functioning subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.5.1 Up to 12 weeks' follow-up	2	73	Mean Difference (IV, Random, 95% CI)	8.56 [-2.04, 19.16]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
15.6 QLQ-C30 social function subscale follow-up values	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.6.1 Up to 12 weeks' follow-up	4	256	Mean Difference (IV, Random, 95% CI)	0.41 [-10.30, 11.11]
15.6.2 More than 12 weeks' less than 6 months' follow-up	1	105	Mean Difference (IV, Random, 95% CI)	3.50 [-6.42, 13.42]
15.7 MOS SF-36 subscale change	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.7.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	4.21 [-7.83, 16.25]
15.7.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	0.60 [-3.57, 4.77]
15.8 MOS SF-36 subscale follow-up values	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.8.1 Up to 12 weeks' follow-up	1	58	Mean Difference (IV, Random, 95% CI)	5.44 [-5.79, 16.67]
15.8.2 6 months' follow-up	1	74	Mean Difference (IV, Random, 95% CI)	-1.00 [-5.95, 3.95]
15.9 Social barriers change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.9.1 More than 6 months' follow-up	1	121	Mean Difference (IV, Random, 95% CI)	3.70 [-8.38, 15.78]
15.10 Social barriers follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.10.1 More than 6 months' follow-up	1	121	Mean Difference (IV, Random, 95% CI)	0.80 [-1.16, 2.76]
15.11 BIQ social body image subscale follow-up values	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
15.11.1 Up to 12 weeks' follow-up	1	57	Mean Difference (IV, Random, 95% CI)	-1.35 [-2.50, -0.20]

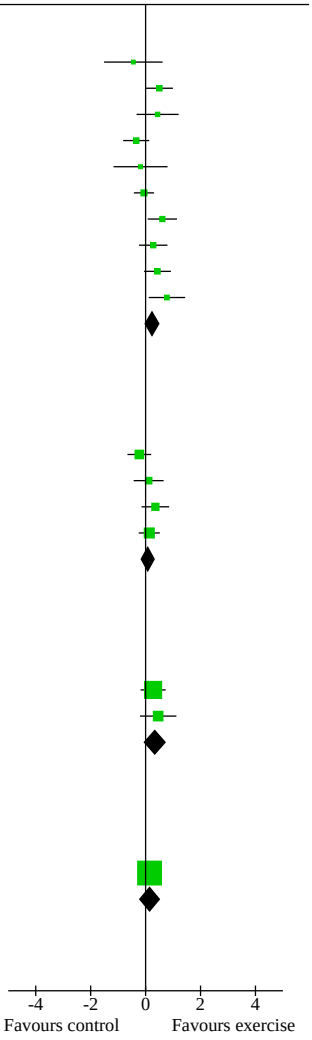
Analysis 15.1. Comparison 15: Social functioning, Outcome 1: Overall social functioning change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
15.1.1 Up to 12 weeks' follow-up								
Herrero 2006	0	15.4	8	-4.2	18.4	8	11.5%	0.23 [-0.75 , 1.22]
Mehnert 2011	12.78	24.93	30	1.23	27.8	27	20.7%	0.43 [-0.09 , 0.96]
Oh 2010	2.29	3.8	54	-0.97	3.8	54	24.1%	0.85 [0.46 , 1.25]
Rogers 2009	0.9	2.6	20	-1.2	2.7	18	17.4%	0.78 [0.11 , 1.44]
Targ 2002	0.43	4	79	0.58	5.5	88	26.3%	-0.03 [-0.33 , 0.27]
Subtotal (95% CI)			**191**			**195**	**100.0%**	**0.45 [0.02 , 0.87]**
Heterogeneity: Tau² = 0.15; Chi² = 13.98, df = 4 (P = 0.007); I² = 71%								
Test for overall effect: Z = 2.07 (P = 0.04)								
15.1.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	0.6	2.4	24	-0.5	3.3	28	100.0%	0.37 [-0.18 , 0.92]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**0.37 [-0.18 , 0.92]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.32 (P = 0.19)								
15.1.3 6 months' follow-up								
Cadmus 2009	0.6	4.6	37	-1	3.2	37	68.3%	0.40 [-0.06 , 0.86]
Rogers 2009	-0.04	4.3	19	-3.9	6.6	17	31.7%	0.69 [0.01 , 1.36]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.49 [0.11 , 0.87]**
Heterogeneity: Tau² = 0.00; Chi² = 0.47, df = 1 (P = 0.49); I² = 0%								
Test for overall effect: Z = 2.53 (P = 0.01)								
15.1.4 More than 6 months' follow-up								
Speck 2010	4.4	33.8	59	0.7	34	62	100.0%	0.11 [-0.25 , 0.47]
Subtotal (95% CI)			**59**			**62**	**100.0%**	**0.11 [-0.25 , 0.47]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.60 (P = 0.55)								
Test for subgroup differences: Chi² = 2.48, df = 3 (P = 0.48), I² = 0%								



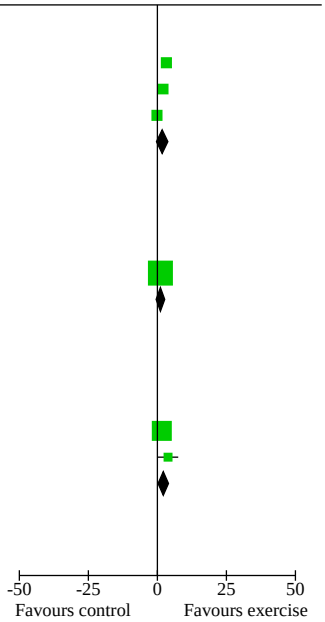
Analysis 15.2. Comparison 15: Social functioning, Outcome 2: Overall social functioning follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
15.2.1 Up to 12 weeks' follow-up								
Banasik 2011	2.8	0.67	7	3.16	0.83	7	4.4%	-0.45 [-1.51 , 0.62]
Daley 2007a	23.3	4.78	33	20.78	5.11	33	12.1%	0.50 [0.01 , 0.99]
Danhauer 2009	23.1	5	13	20.4	6.8	14	7.2%	0.44 [-0.33 , 1.20]
Dimeo 2004	68	28	34	77	24	35	12.5%	-0.34 [-0.82 , 0.13]
Herrero 2006	87.5	19.4	8	91.7	23.6	8	5.0%	-0.18 [-1.17 , 0.80]
Knols 2011	71.3	28.6	57	72.9	24.9	57	15.3%	-0.06 [-0.43 , 0.31]
Mehnert 2011	85.56	23.06	30	69.75	28.13	27	11.2%	0.61 [0.08 , 1.14]
Milne 2008a	20.6	4	29	19.5	3.8	29	11.5%	0.28 [-0.24 , 0.80]
Moadel 2007	20.68	5.19	45	18.17	6.59	26	12.2%	0.43 [-0.06 , 0.92]
Rogers 2009	0.9	2.6	20	-1.2	2.7	18	8.7%	0.78 [0.11 , 1.44]
Subtotal (95% CI)			**276**			**254**	**100.0%**	**0.23 [-0.02 , 0.48]**
Heterogeneity: Tau² = 0.07; Chi² = 16.85, df = 9 (P = 0.05); I² = 47%								
Test for overall effect: Z = 1.83 (P = 0.07)								
15.2.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	21	3.7	62	21.8	3	31	27.4%	-0.23 [-0.66 , 0.20]
Courneya 2003c	21.1	3.5	24	20.7	3.6	28	17.7%	0.11 [-0.43 , 0.66]
Daley 2007a	23.23	4.34	31	21.56	4.9	31	20.7%	0.36 [-0.15 , 0.86]
Knols 2011	78.5	26.3	51	75	25.5	54	34.3%	0.13 [-0.25 , 0.52]
Subtotal (95% CI)			**168**			**144**	**100.0%**	**0.08 [-0.16 , 0.31]**
Heterogeneity: Tau² = 0.00; Chi² = 3.20, df = 3 (P = 0.36); I² = 6%								
Test for overall effect: Z = 0.64 (P = 0.52)								
15.2.3 6 months' follow-up								
Cadmus 2009	22.3	4.9	37	20.6	7.2	37	67.8%	0.27 [-0.18 , 0.73]
Rogers 2009	21.4	4.7	19	18.8	6.4	17	32.2%	0.46 [-0.21 , 1.12]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**0.33 [-0.04 , 0.71]**
Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 1 (P = 0.66); I² = 0%								
Test for overall effect: Z = 1.73 (P = 0.08)								
15.2.4 More than 6 months' follow-up								
Speck 2010	14.9	5.2	59	14.1	5.8	62	100.0%	0.14 [-0.21 , 0.50]
Subtotal (95% CI)			**59**			**62**	**100.0%**	**0.14 [-0.21 , 0.50]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.79 (P = 0.43)								
Test for subgroup differences: Chi² = 1.57, df = 3 (P = 0.67), I² = 0%								



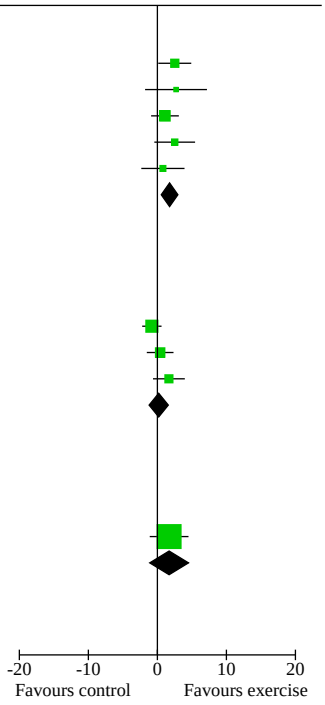
Analysis 15.3. Comparison 15: Social functioning, Outcome 3: FACT social functioning subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.3.1 Up to 12 weeks' follow-up								
Oh 2010	2.29	3.8	54	-0.97	3.8	54	34.1%	3.26 [1.83 , 4.69]
Rogers 2009	0.9	2.6	20	-1.2	2.7	18	32.0%	2.10 [0.41 , 3.79]
Targ 2002	0.43	4	79	0.58	5.5	88	33.9%	-0.15 [-1.60 , 1.30]
Subtotal (95% CI)			**153**			**160**	**100.0%**	**1.73 [-0.33 , 3.79]**
Heterogeneity: Tau² = 2.71; Chi² = 11.03, df = 2 (P = 0.004); I² = 82%								
Test for overall effect: Z = 1.65 (P = 0.10)								
15.3.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003c	0.6	2.4	24	-0.5	3.3	28	100.0%	1.10 [-0.45 , 2.65]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**1.10 [-0.45 , 2.65]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.39 (P = 0.17)								
15.3.3 6 months' follow-up								
Cadmus 2009	0.6	4.6	37	-1	3.2	37	76.3%	1.60 [-0.21 , 3.41]
Rogers 2009	-0.04	4.3	19	-3.9	6.6	17	23.7%	3.86 [0.17 , 7.55]
Subtotal (95% CI)			**56**			**54**	**100.0%**	**2.14 [0.25 , 4.02]**
Heterogeneity: Tau² = 0.36; Chi² = 1.17, df = 1 (P = 0.28); I² = 14%								
Test for overall effect: Z = 2.22 (P = 0.03)								
Test for subgroup differences: Chi² = 0.72, df = 2 (P = 0.70), I² = 0%								



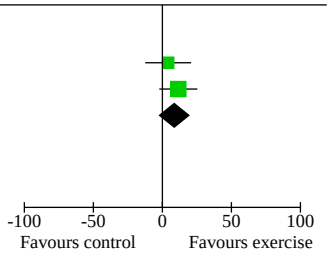
Analysis 15.4. Comparison 15: Social functioning, Outcome 4: FACT social functioning subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.4.1 Up to 12 weeks' follow-up								
Daley 2007a	23.3	4.78	33	20.78	5.11	33	25.4%	2.52 [0.13 , 4.91]
Danhauer 2009	23.1	5	13	20.4	6.8	14	7.2%	2.70 [-1.78 , 7.18]
Milne 2008a	20.6	4	29	19.5	3.8	29	35.9%	1.10 [-0.91 , 3.11]
Moadel 2007	20.68	5.19	45	18.17	6.59	26	16.6%	2.51 [-0.44 , 5.46]
Rogers 2009	22.6	4.1	20	21.8	5.7	19	14.8%	0.80 [-2.33 , 3.93]
Subtotal (95% CI)			**140**			**121**	**100.0%**	**1.77 [0.56 , 2.97]**
Heterogeneity: Tau² = 0.00; Chi² = 1.58, df = 4 (P = 0.81); I² = 0%								
Test for overall effect: Z = 2.88 (P = 0.004)								
15.4.2 More than 12 weeks' less than 6 months' follow-up								
Courneya 2003b	21	3.7	62	21.8	3	31	43.8%	-0.80 [-2.20 , 0.60]
Courneya 2003c	21.1	3.5	24	20.7	3.6	28	31.3%	0.40 [-1.53 , 2.33]
Daley 2007a	23.23	4.34	31	21.56	4.9	31	24.9%	1.67 [-0.63 , 3.97]
Subtotal (95% CI)			**117**			**90**	**100.0%**	**0.19 [-1.20 , 1.58]**
Heterogeneity: Tau² = 0.64; Chi² = 3.43, df = 2 (P = 0.18); I² = 42%								
Test for overall effect: Z = 0.27 (P = 0.79)								
15.4.3 6 months' follow-up								
Cadmus 2009	22.3	4.9	37	20.6	7.2	37	100.0%	1.70 [-1.11 , 4.51]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**1.70 [-1.11 , 4.51]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.19 (P = 0.24)								
Test for subgroup differences: Chi² = 2.98, df = 2 (P = 0.22), I² = 33.0%								



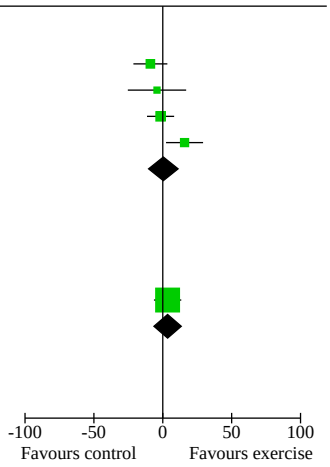
Analysis 15.5. Comparison 15: Social functioning, Outcome 5: QLQ-C30 social functioning subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.5.1 Up to 12 weeks' follow-up								
Herrero 2006	0	15.4	8	-4.2	18.4	8	40.7%	4.20 [-12.43 , 20.83]
Mehnert 2011	12.78	24.93	30	1.23	27.8	27	59.3%	11.55 [-2.22 , 25.32]
Subtotal (95% CI)			**38**			**35**	**100.0%**	**8.56 [-2.04 , 19.16]**
Heterogeneity: Tau² = 0.00; Chi² = 0.45, df = 1 (P = 0.50); I² = 0%								
Test for overall effect: Z = 1.58 (P = 0.11)								
Test for subgroup differences: Not applicable								



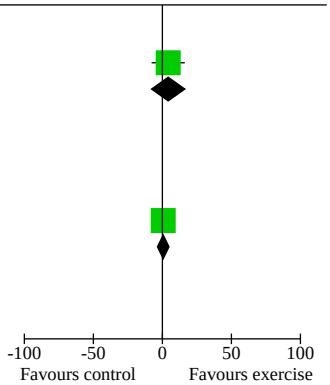
Analysis 15.6. Comparison 15: Social functioning, Outcome 6: QLQ-C30 social function subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.6.1 Up to 12 weeks' follow-up								
Dimeo 2004	68	28	34	77	24	35	27.2%	-9.00 [-21.32 , 3.32]
Herrero 2006	87.5	19.4	8	91.7	23.6	8	16.0%	-4.20 [-25.37 , 16.97]
Knols 2011	71.3	28.6	57	72.9	24.9	57	31.3%	-1.60 [-11.44 , 8.24]
Mehnert 2011	85.56	23.05	30	69.75	28.13	27	25.5%	15.81 [2.37 , 29.25]
Subtotal (95% CI)			**129**			**127**	**100.0%**	**0.41 [-10.30 , 11.11]**
Heterogeneity: Tau² = 70.07; Chi² = 7.62, df = 3 (P = 0.05); I² = 61%								
Test for overall effect: Z = 0.07 (P = 0.94)								
15.6.2 More than 12 weeks' less than 6 months' follow-up								
Knols 2011	78.5	26.3	51	75	25.5	54	100.0%	3.50 [-6.42 , 13.42]
Subtotal (95% CI)			**51**			**54**	**100.0%**	**3.50 [-6.42 , 13.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.69 (P = 0.49)								
Test for subgroup differences: Chi² = 0.17, df = 1 (P = 0.68), I² = 0%								



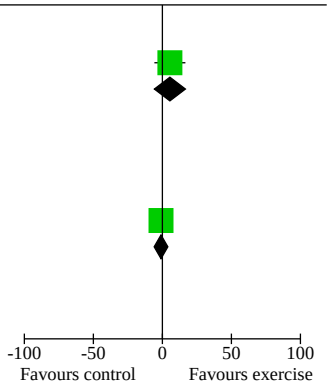
Analysis 15.7. Comparison 15: Social functioning, Outcome 7: MOS SF-36 subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.7.1 Up to 12 weeks' follow-up								
Mehnert 2011	7.33	22.33	30	3.12	24.3	28	100.0%	4.21 [-7.83 , 16.25]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**4.21 [-7.83 , 16.25]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.69 (P = 0.49)								
15.7.2 6 months' follow-up								
Cadmus 2009	0.7	11	37	0.1	6.8	37	100.0%	0.60 [-3.57 , 4.77]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**0.60 [-3.57 , 4.77]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.28 (P = 0.78)								
Test for subgroup differences: Chi² = 0.31, df = 1 (P = 0.58), I² = 0%								



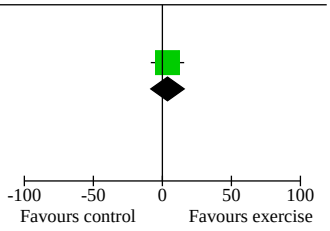
Analysis 15.8. Comparison 15: Social functioning, Outcome 8: MOS SF-36 subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.8.1 Up to 12 weeks' follow-up								
Mehnert 2011	85.35	20.88	30	79.91	22.65	28	100.0%	5.44 [-5.79 , 16.67]
Subtotal (95% CI)			**30**			**28**	**100.0%**	**5.44 [-5.79 , 16.67]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.95 (P = 0.34)								
15.8.2 6 months' follow-up								
Cadmus 2009	49.5	12	37	50.5	9.6	37	100.0%	-1.00 [-5.95 , 3.95]
Subtotal (95% CI)			**37**			**37**	**100.0%**	**-1.00 [-5.95 , 3.95]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.40 (P = 0.69)								
Test for subgroup differences: Chi² = 1.06, df = 1 (P = 0.30), I² = 5.4%								



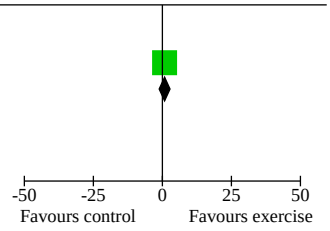
Analysis 15.9. Comparison 15: Social functioning, Outcome 9: Social barriers change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.9.1 More than 6 months' follow-up								
Speck 2010	4.4	33.8	59	0.7	34	62	100.0%	3.70 [-8.38 , 15.78]
Subtotal (95% CI)			**59**			**62**	**100.0%**	**3.70 [-8.38 , 15.78]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.60 (P = 0.55)								
Test for subgroup differences: Not applicable								



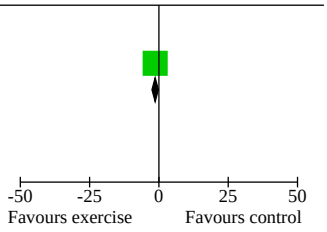
Analysis 15.10. Comparison 15: Social functioning, Outcome 10: Social barriers follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.10.1 More than 6 months' follow-up								
Speck 2010	14.9	5.2	59	14.1	5.8	62	100.0%	0.80 [-1.16 , 2.76]
Subtotal (95% CI)			**59**			**62**	**100.0%**	**0.80 [-1.16 , 2.76]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.80 (P = 0.42)								
Test for subgroup differences: Not applicable								



Analysis 15.11. Comparison 15: Social functioning, Outcome 11: BIQ social body image subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
15.11.1 Up to 12 weeks' follow-up								
Mehnert 2011	5.47	1.66	30	6.82	2.6	27	100.0%	-1.35 [-2.50 , -0.20]
Subtotal (95% CI)			**30**			**27**	**100.0%**	**-1.35 [-2.50 , -0.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.31 (P = 0.02)								
Test for subgroup differences: Not applicable								

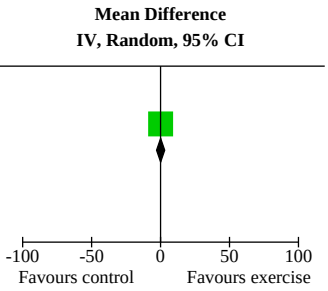


Comparison 16. Spirituality

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
16.1 FACT spirituality subscale change	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
16.1.1 Up to 12 months' follow-up	1	167	Mean Difference (IV, Random, 95% CI)	0.05 [-2.89, 2.99]
16.2 FACT spirituality subscale follow-up values	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
16.2.1 Up to 12 weeks' follow-up	3	194	Mean Difference (IV, Random, 95% CI)	1.51 [-3.64, 6.65]

Analysis 16.1. Comparison 16: Spirituality, Outcome 1: FACT spirituality subscale change

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
16.1.1 Up to 12 months' follow-up								
Targ 2002	3.4	9.2862	79	3.35	10.1001	88	100.0%	0.05 [-2.89 , 2.99]
Subtotal (95% CI)			**79**			**88**	**100.0%**	**0.05 [-2.89 , 2.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.03 (P = 0.97)								
Test for subgroup differences: Not applicable								

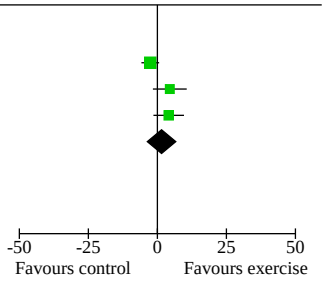


Analysis 16.2. Comparison 16: Spirituality, Outcome 2: FACT spirituality subscale follow-up values

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
16.2.1 Up to 12 weeks' follow-up								
Courneya 2003a	30.2	9	51	32.78	7	45	40.5%	-2.58 [-5.79 , 0.63]
Danhauer 2009	26	6.7	13	21.5	9.4	14	28.6%	4.50 [-1.62 , 10.62]
Moadel 2007	36.85	8.23	45	32.75	12.96	26	30.9%	4.10 [-1.43 , 9.63]
Subtotal (95% CI)			**109**			**85**	**100.0%**	**1.51 [-3.64 , 6.65]**

Heterogeneity: Tau² = 14.33; Chi² = 6.67, df = 2 (P = 0.04); I² = 70%
Test for overall effect: Z = 0.57 (P = 0.57)

Test for subgroup differences: Not applicable



ADDITIONAL TABLES

Table 1. HRQoL instruments used by investigators

Instrument name	Abbreviation	Overall domain or subscale	Direction of response	Trials using this scale
Health-related quality of life				
Cancer Rehabilitation Evaluation System Short Form	CARES-SF	HRQoL	Higher score indicates worse status	Ohira 2006
Chae and Cho	Cho VAS	HRQoL	Higher score indicates better status	Cho 2006
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	HRQoL	Higher score indicates better status	Culos-Reed 2006; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011; Thorsen 2005
Functional Assessment of Cancer Therapy - Anemia	FACT-An	HRQoL	Higher score indicates better status	Courneya 2009; McNeely 2008a
Functional Assessment of Cancer Therapy - Breast	FACT-B	HRQoL	Higher score indicates better status	Banasik 2011; Cadmus 2009; Courneya 2003c; Daley 2007a; Danhauer 2009 Milne 2008a; Rogers 2009
Functional Assessment of Cancer Therapy - Colorectal	FACT-C	HRQoL	Higher score indicates better status	Bourke 2011; Courneya 2003b
Functional Assessment of Cancer Therapy - General	FACT-G	HRQoL	Higher score indicates better status	Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Daley 2007a; Donnelly 2011; Heim 2007; McNeely 2008a; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009

Table 1. HRQoL instruments used by investigators *(Continued)*

Functional Assessment of Cancer Therapy - Fatigue	FACT-F	HRQoL	Higher score indicates better status	Heim 2007
Quality of Life Index for Cancer Patients	QoL Index	HRQoL	Higher score indicates better status	Burnham 2002
Medical Outcomes Study Short Form-36	MOS SF-36	HRQoL	Higher score indicates better status	Cadmus 2009; Mehnert 2011
Functional Assessment of Chronic Illness Therapy - Fatigue	FACIT-F	HRQoL	Higher score indicates better status	Mustian 2004; Targ 2002
Condition-specific HRQoL				
Functional Assessment of Cancer Therapy - Breast	FACT-B	Additional breast cancer concerns	Higher score indicates better status	Banasik 2011; Cadmus 2009; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Rogers 2009
Functional Assessment of Cancer Therapy - Colorectal	FACT-C	Colorectal cancer scale	Higher score indicates better status	Courneya 2003b
Functional Assessment of Cancer Therapy	FACT	Lymphoma cancer concerns	Higher score indicates better status	Courneya 2009
Functional Assessment of Chronic Illness Therapy - Fatigue	FACIT-F	General cancer concerns	Higher score indicates better status	Targ 2002
Neck Dissection Impairment Index	NDII	Head and neck specific concerns	Higher score indicates better status	McNeely 2008a
Anxiety				
Hospital Anxiety and Depression Scale	HADS	Anxiety	Higher score indicates worse status	Berglund 1994; Heim 2007; Mehnert 2011; Thorsen 2005
State-Trait Anxiety Scale	STAI	State anxiety	Higher score indicates worse status	Cadmus 2009; Cohen 2004; Segar 1998
Linear Analog Self-Assessment	LASA	Anxiety	Higher score indicates worse status	Burnham 2002
Profile of Mood Scale	POMS	Tension-anxiety	Higher score indicates worse status	Culos-Reed 2006; Moadel 2007; Oh 2010; Pinto 2003; Targ 2002

Table 1. HRQoL instruments used by investigators *(Continued)*

Body image/self-esteem				
Body Esteem Scale	BES	Weight concern	Higher score indicates better status	Pinto 2003; Pinto 2005
Physical Self-Perception Profile	PSPP	Body image	Higher score indicates better status	Daley 2007a
Body Image Questionnaire	BIQ	Individual body image	Higher score indicates worse status	Mehnert 2011
Body Image and Relationships Scale	BIRS	Self-perception	Higher score indicates worse status	Speck 2010
Social Physique Anxiety Scale	SPAS	Anxiety about body	Higher score indicates worse status	Milne 2008a
Rosenberg Self-Esteem		Self-esteem	Higher score indicates better status	Cadmus 2009; Courneya 2003c; Mustian 2004; Segar 1998
Cognitive function				
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Cognitive functioning	Higher score indicates better status	Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011
Symptoms of Stress Inventory	SOSI	Cognitive disorganization	Higher score indicates worse status	Culos-Reed 2006
Functional Assessment of Cancer Therapy - Cognitive	FACT-Cog	Cognitive functioning	Higher score indicates worse status	Oh 2010; Rogers 2009
Profile of Mood Scale	POMS	Confusion-bewilderment	Higher score indicates worse status	Moadel 2007; Oh 2010; Pinto 2003; Targ 2002
Profile of Mood Scale	POMS	Concentration	Higher score indicates worse status	Culos-Reed 2006
Linear Analog Self-Assessment	LASA	Confusion	Higher score indicates worse status	Burnham 2002
Depression				
Centers for Epidemiologic Studies - Depression Scale	CES-D	Depression	Higher score indicates worse status	Cadmus 2009; Cohen 2004; Courneya 2003a; Courneya 2003b; Courneya 2009; Dan-

Table 1. HRQoL instruments used by investigators *(Continued)*

				hauer 2009; Dodd 2010; Payne 2008
Hospital Anxiety and Depression Scale	HADS	Depression	Higher score indicates worse status	Berglund 1994; Mehnert 2011; Thorsen 2005
Beck Depression Inventory-II	BDI or BECK	Depression	Higher score indicates worse status	Daley 2007a; Donnelly 2011; Segar 1998
Finnish version of Beck	BECK	Depression	Higher score indicates worse status	Penttinen 2011
Linear Analog Self-Assessment	LASA	Depression	Higher score indicates worse status	Burnham 2002
Profile of Mood Scale	POMS	Depression-dejection	Higher score indicates worse status	Culos-Reed 2006; Oh 2010; Pinto 2003; Targ 2002
Emotional function/mental health				
Functional Assessment of Cancer Therapy	FACT subscale	Emotional well-being	Higher score indicates better status	Banasik 2011; Cadmus 2009; Courneya 2003b; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009
Functional Assessment of Chronic Illness - Fatigue	FACIT-F	Emotional well-being	Higher score indicates better status	Targ 2002
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Emotional functioning	Higher score indicates better status	Bai 2004; Culos-Reed 2006; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011; Thorsen 2005
Symptoms of Stress Inventory	SOSI	Emotional irritability	Higher score indicates worse status	Culos-Reed 2006
Profile of Mood Scale	POMS	Mood	Higher score indicates worse status	Culos-Reed 2006; Moadel 2007; Oh 2010; Pinto 2003; Pinto 2005; Targ 2002
Profile of Mood Scale	POMS	Anger-hostility	Higher score indicates worse status	Oh 2010; Pinto 2003; Targ 2002
Profile of Mood Scale	POMS	Anxiety + depression scales	Higher score indicates worse status	Fillion 2008
Profile of Mood Scale	POMS	Irritability	Higher score indicates worse status	Moadel 2007; Oh 2010

Table 1. HRQoL instruments used by investigators *(Continued)*

Linear Analog Self-Assessment	LASA	Anger	Higher score indicates worse status	Burnham 2002
Fordyce Happiness Measure	FORDYCE	Happiness	Higher score indicates better status	Cadmus 2009
Happiness Measure	HM	Happiness	Higher score indicates better status	Courneya 2003c; Courneya 2009
Medical Outcomes Study Short Form-12	MOS SF-12	Mental health component	Higher score indicates better status	Danhauer 2009; Fillion 2008
Medical Outcomes Study Short Form-36	MOS SF-36	Mental health component	Higher score indicates better status	Cadmus 2009; Speck 2010; Tang 2010
Positive and Negative Affect Scale	PANAS	Positivity and negativity	Higher score indicates better status	Danhauer 2009; Donnelly 2011; Pinto 2003
Cancer Rehabilitation Evaluation System Short Form	CARES-SF	Psychological functioning	reported as % change, higher score indicates worse status	Ohira 2006
Functional Assessment of Cancer Therapy - Breast	FACT-B	Psychological functioning	Higher score indicates better status	Banasik 2011; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Rogers 2009
Psychosocial Adjustment Scale (Lee)		Psychological functioning	Higher score indicates better status	Cho 2006
Satisfaction with Life Scale	SWLS	Satisfaction	Higher score indicates better status	Courneya 2003a; Courneya 2003b; Daley 2007a
Cohen's Perceived Stress Scale		Emotional function	Higher score indicates worse status	Cadmus 2009
Symptoms of Stress Inventory	SOSI	Stress	Higher score indicates better status	Berglund 1994; Mehnert 2011; Thorsen 2005
Symptom Checklist-90 Revised	SCL-90R	Psychological symptom burden		Mehnert 2011
Fatigue				

Table 1. HRQoL instruments used by investigators *(Continued)*

European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Fatigue	Higher score indicates worse status	Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Thorsen 2005
Functional Assessment of Cancer Therapy	FACT subscale	Fatigue	Higher score indicates better status	Bourke 2011; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Danhauer 2009; Donnelly 2011; Heim 2007; Knols 2011; McNeely 2008a; Rogers 2009
Profile of Mood Scale	POMS	Fatigue-inertia	Higher score indicates worse status	Oh 2010; Pinto 2003; Targ 2002
Profile of Mood Scale	POMS	Vigor-activity	Higher score indicates better status	Fillion 2008; Oh 2010; Pinto 2005; Targ 2002
Linear Analog Self-Assessment	LASA	Fatigue	Higher score indicates worse status	Burnham 2002
Schwartz Cancer Fatigue Scale	SCFS	Fatigue	Higher score indicates worse status	Milne 2008a
Multidimensional Fatigue Inventory	MFI	Fatigue	Higher score indicates worse status	Donnelly 2011; Fillion 2008; Heim 2007
Revised Piper Fatigue Scale	PFS	Fatigue	Higher score indicates worse status	Daley 2007a; Dodd 2010; Payne 2008
Functional Assessment of Chronic Illness Therapy - Fatigue	FACIT-F	Fatigue	Higher score indicates better status	Penttinen 2011
Linear Analog Scale for fatigue		Fatigue	Higher score indicates worse status	Pinto 2005
Medical Outcomes Study Short Form-36	MOS SF-36	Vitality	Higher score indicates better status	Cadmus 2009
Linear Analog Self-Assessment Scale	LASA	Vitality	Higher score indicates better status	Burnham 2002
Brief Fatigue Inventory	BFI	Fatigue	Higher score indicates better status	Cohen 2004
General health perspective				
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	General health score	Higher score indicates better status	Dimeo 2004; Donnelly 2011; Knols 2011; Mehnert 2011
Medical Outcomes Study Short Form-12	MOS SF-12	Item on health	Higher score indicates better status	Courneya 2009

Table 1. HRQoL instruments used by investigators *(Continued)*

Single question on health		Perceived health	Higher score indicates better status	Rogers 2009
Pain				
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Pain	Higher score indicates worse status	Dimeo 2004; Knols 2011; Mehnert 2011
Medical Outcomes Study Short Form-36	MOS SF-36	Bodily pain	Higher score indicates better status	Cadmus 2009
Shoulder Pain and Disability Index	SPADI	Pain	Higher score indicates worse status	McNeely 2008a
Worst Pain Intensity Scale	WPIS	Pain	Higher score indicates worse status	Dodd 2010
Brief Pain Inventory	BPI	Pain	Higher score indicates worse status	Fillion 2008
Physical well-being				
Cancer Rehabilitation Evaluation System Short Form	CARES-SF	Physical function	reported as % change, Higher score indicates worse status	Ohira 2006
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Physical function	Higher score indicates better status	Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Thorsen 2005
Functional Assessment of Cancer Therapy - Breast	FACT-B	Physical well-being	Higher score indicates better status	Banasik 2011; Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Courneya 2009; Daley 2007a; Danhauer 2009; Heim 2007; Mehnert 2011; Milne 2008a; Moadel 2007; Rogers 2009
Functional Assessment of Chronic Illness Fatigue	FACIT-F	Physical well-being	Higher score indicates better status	Targ 2002
Medical Outcomes Study Short Form-12	MOS SF-12	Physical function	Higher score indicates better status	Danhauer 2009; Fillion 2008
Body Image and Relationships Scale	BIRS	Strength and health	Higher score indicates worse status	Speck 2010
Medical Outcomes Study Short Form-36	MOS SF-36	Physical component	Higher score indicates better status	Cadmus 2009; Speck 2010; Tang 2010

Table 1. HRQoL instruments used by investigators *(Continued)*

Body Esteem Scale	BES	Physical condition	Higher score indicates better status	Pinto 2003; Pinto 2005
Role function				
Functional Assessment of Cancer Therapy	FACT subscale	Functional well-being	Higher score indicates better status	Banasik 2011; Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Daley 2007a; Danhauer 2009; Heim 2007; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Functional well-being	Higher score indicates better status	Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011
Functional Assessment of Chronic Illness Fatigue	FACIT-F	Functional well-being	Higher score indicates better status	Targ 2002
Medical Outcomes Study Short Form-36	MOS SF-36	Physical role	Higher score indicates better status	Cadmus 2009
Cancer Rehabilitation Evaluation System Short Form	CARES-SF	Marital role	Reported as % change, higher score indicates worse status	Ohira 2006
Sexuality				
Body Esteem Scale	BES	Sexual attractiveness	Higher score indicates better status	Pinto 2003; Pinto 2005
Cancer Rehabilitation Evaluation System Short Form	CARES-SF	Sexual	Reported at % change, higher score indicates worse status	Ohira 2006
Body Image and Relationships Scale	BIRS	Appearance and sexuality	Higher score indicates worse status	Speck 2010
Sleep				
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Insomnia	Higher score indicates worse status	Dimeo 2004; Knols 2011; Mehnert 2011; Penttinen 2011
Pittsburgh Sleep Quality Index	PSQI	Sleep disturbance	Higher score indicates worse status	Cohen 2004; Danhauer 2009; Donnelly 2011; Payne 2008; Rogers 2009
General Sleep Disturbance Scale	GSDS	Sleep disturbance	Higher score indicates worse status	Dodd 2010

Table 1. HRQoL instruments used by investigators *(Continued)*

Taiwanese Pittsburgh Sleep Quality Index	PSQI	Sleep disturbance	Higher score indicates worse status	Tang 2010
Social functioning				
European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30	QLQ-C30	Social functioning	Higher score indicates better status	Bai 2004; Dimeo 2004; Herrero 2006; Knols 2011; Mehnert 2011; Penttinen 2011
Functional Assessment of Cancer Therapy	FACT subscale	Social/family well-being	Higher score indicates better status	Banasik 2011; Cadmus 2009; Courneya 2003a; Courneya 2003b; Courneya 2003c; Daley 2007a; Danhauer 2009; Milne 2008a; Moadel 2007; Oh 2010; Rogers 2009
Functional Assessment of Chronic Illness - Fatigue	FACIT-F	Social/family well-being	Higher score indicates better status	Targ 2002
Medical Outcomes Study Short Form-36	MOS SF-36	Social/family well-being	Higher score indicates better status	Cadmus 2009
Social Barriers		Social function	Higher score indicates better status	Mehnert 2011
Body Image and Relationships Scale	BIRS	Social barriers	Higher score indicates worse status	Speck 2010
Body Image Questionnaire	BIQ	Social body image	Higher score indicates worse status	Mehnert 2011
Spiritual function				
Functional Assessment of Cancer Therapy - Breast	FACT-B	Spiritual	Higher score indicates better status	Courneya 2003a; Rogers 2009; Targ 2002
FACIT - Spirituality	FACIT-SP	Peace	Higher score indicates better status	Danhauer 2009; Moadel 2007
Principles of Living Survey	PLS	Spiritual	Higher score indicates better status	Targ 2002

APPENDICES

Appendix 1. MEDLINE search strategy

[inception to June 2009; 419 hits] [January 2009 to September 2011; 205 hits]

1. exp exercise/
2. exercise tolerance/
3. exp exertion/
4. Pliability/
5. physical fitness/
6. "Physical Education and Training"/
7. exp physical endurance/
8. exercise therapy/
9. exercising.mp.
10. physical condition$.mp.
11. stamina.mp.
12. motor activity/
13. exercise test/
14. exp Sports/
15. tai chi.mp. or tai ji/
16. yoga/
17. muscle stretching exercises/
18. exp "range of motion, articular"/
19. pilates.mp.
20. qigong.mp.
21. chi kung.mp.
22. resistance training.mp.
23. mind body therap$.mp.
24. exp complementary therapies/
25. Bad Ragaz.mp.
26. Ai Chi.mp.
27. Halliwick.mp.
28. hippotherapy.mp.
29. Hydrotherapy/
30. balance exercise$.mp.
31. aquatic exercise$.mp.
32. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31
33. "quality of life"/
34. exp health status/
35. "activities of daily living"/
36. life qualit$.mp.
37. exp self concept/
38. health level.mp.
39. level of health.mp.
40. wellness.mp.
41. well being.mp.
42. (activities of daily life or daily living activities).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
43. functional ability.mp.
44. good health.mp.
45. healthiness.mp.
46. patient reported outcomes.mp.
47. social adjustment/

48.physical limitations.mp.
49.psychiatric status.mp.
50.pain measurement/
51.functional assessment.mp.
52.fact questionnaire.mp.
53.fact survey.mp.
54.qlc-c30.mp.
55.facit.mp.
56.toi.mp.
57.(flic or sf-36 or ces-d or bdi or sta1 or bfi or hads or lasa or poms or qli or rsci or pais or bpi or msas or mos or ptgi or panas).mp.
58.sense of coherence.mp.
59.randomized.ab.
60.placebo.ab.
61.randomly.ab.
62.trial.ab.
63.randomized controlled trial.pt.
64.controlled clinical trial.pt.
65.survivors/
66.exp neoplasms/
67.cancer survivor$.mp.
68.cancer.mp.
69.post treatment.mp.
70.after treatment.mp.
71.35 or 33 or 53 or 48 or 42 or 46 or 44 or 55 or 50 or 39 or 57 or 36 or 40 or 51 or 58 or 41 or 47 or 52 or 38 or 34 or 56 or 49 or 37 or 45 or 43 or 54
72.67 or 66 or 68 or 65
73.59 or 60 or 63 or 64 or 61 or 62
74.random$.ab.
75.74 or 73
76.32 and 75 and 72

Appendix 2. CENTRAL search strategy

[inception to August 2009; 113 hits] [January 2009 to September 2011; 136 hits]

Searched via Ovid EBM Reviews

[mp = ti, ot, ab, tx, kw, ct, sh, hw]

1. exp exercise/
2. exercise tolerance/
3. exp exertion/
4. Pliability/
5. physical fitness/
6. "Physical Education and Training"/
7. exp physical endurance/
8. exercise therapy/
9. exercising.mp.
10.physical condition$.mp.
11.stamina.mp.
12.motor activity/
13.exercise test/
14.exp Sports/
15.tai chi.mp. or tai ji/
16.yoga/
17.muscle stretching exercises/

18.exp "range of motion, articular"/
19.pilates.mp.
20.qigong.mp.
21.chi kung.mp.
22.resistance training.mp.
23.mind body therap$.mp.
24.exp complementary therapies/
25.Bad Ragaz.mp.
26.Ai Chi.mp.
27.Halliwick.mp.
28.hippotherapy.mp.
29.Hydrotherapy/
30.balance exercise$.mp.
31.aquatic exercise$.mp.
32.1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31
33."quality of life"/
34.exp health status/
35."activities of daily living"/
36.life qualit$.mp.
37.exp self concept/
38.health level.mp.
39.level of health.mp.
40.wellness.mp.
41.well being.mp.
42.(activities of daily life or daily living activities).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
43.functional ability.mp.
44.good health.mp.
45.healthiness.mp.
46.patient reported outcomes.mp.
47.social adjustment/
48.physical limitations.mp.
49.psychiatric status.mp.
50.pain measurement/
51.functional assessment.mp.
52.fact questionnaire.mp.
53.fact survey.mp.
54.qlc-c30.mp.
55.facit.mp.
56.toi.mp.
57.(flic or sf-36 or ces-d or bdi or sta1 or bfi or hads or lasa or poms or qli or rsci or pais or bpi or msas or mos or ptgi or panas).mp.
58.sense of coherence.mp.
59.randomized.ab.
60.placebo.ab.
61.randomly.ab.
62.trial.ab.
63.randomized controlled trial.pt.
64.controlled clinical trial.pt.
65.survivors/
66.exp neoplasms/
67.cancer survivor$.mp.
68.cancer.mp.

69.post treatment.mp.
70.after treatment.mp.
71.35 or 33 or 53 or 48 or 42 or 46 or 44 or 55 or 50 or 39 or 57 or 36 or 40 or 51 or 58 or 41 or 47 or 52 or 38 or 34 or 56 or 49 or 37 or 45 or 43 or 54
72.67 or 66 or 68 or 65
73.59 or 60 or 63 or 64 or 61 or 62
74.random$.ab.
75.74 or 73
76.32 and 75 and 72 and 71
77.from 76 keep 1-418
78.from 77 keep 1-10
79.Physical Exertion/
80.32 or 79
81.80 and 75 and 72 and 71

Appendix 3. EMBASE search strategy

[inception to August 2009; 492 hits] [January 2009 to September 2011; 483 hits]

1. exp exercise/
2. exertion.mp.
3. pliability/
4. fitness/
5. (physical education and training).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
6. physical endurance.mp. or endurance/
7. kinesiotherapy/
8. exercising.mp.
9. "physical condition$".mp.
10. stamina.mp.
11. exp motor activity/
12. exp sports/
13. exercise test/
14. tai chi.mp.
15. tai ji.mp.
16. yoga/
17. stretching exercise/
18. "range of motion"/
19. pilates.mp.
20. qigong.mp.
21. chi kung.mp.
22. muscle strength/ or muscle training/ or resistance training.mp.
23. mind body therapy.mp.
24. alternative medicine/
25. bad ragaz.mp.
26. ai chi.mp.
27. halliwick.mp.
28. hippotherapy.mp.
29. hydrotherapy/
30. balance exercises.mp.
31. aquatic exercise/
32. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31
33. "quality of life"/
34. exp health status/

35.daily life activity/
36.life qualit$.mp.
37.exp self concept/
38.health level.mp.
39."level of health".mp.
40.wellbeing/
41.wellness.mp.
42.good health.mp.
43.functional ability.mp.
44.healthiness.mp.
45."patient reported outcomes".mp.
46.social adaptation/
47.physical limitation$.mp.
48.psychiatric status.mp.
49.pain assessment/
50.functional assessment/
51.questionnaire/ or fact questionnaire.mp.
52.fact survey.mp.
53.health survey/
54.qlc-c30.mp.
55.facit.mp.
56.toi.mp.
57.sense of coherence.mp.
58.(flic or sf-36 or ces-d or bdi or stal or bfi or hads or lasa or poms or qli or rsci or pais or bpi or msas or mos or ptgi or panas).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
59.35 or 33 or 53 or 48 or 42 or 46 or 44 or 55 or 50 or 39 or 57 or 36 or 40 or 51 or 58 or 41 or 47 or 52 or 38 or 34 or 56 or 49 or 37 or 45 or 43 or 54
60.exp neoplasm/
61.survivor/
62.cancer survivor/
63.cancer.mp.
64.60 or 63 or 61 or 62
65.59 and 32 and 64
66.("randomized controlled trial" or "clinical trial" or placebo or trial or random$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
67.66 and 65
68.limit 67 to article

Appendix 4. CINAHL search strategy

[inception to July 2009; 101 hits] [January 2009 to September 2011; 63 hits]

Search ID#	Search Terms
S68	S66 and S67
S67	(random* or placebo or trial)
S66	S33 and S62 and S65
S65	S63 or S64
S64	(MH "Cancer Survivors") OR (MH "Survivors")

(Continued)

S63	(MH "Neoplasms+")
S62	S34 or S35 or S36 or S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45 or S46 or S47 or S48 or S49 or S50 or S51 or S52 or S53 or S54 or S55 or S58 or S59 or S60 or S61
S61	"sense of coherence"
S60	(flic or sf-37 or ces-d of bdi or stal or bfi or hads or lasa or poms or qli or rsci or pals or bpi or msas or mos or ptgi or panas)
S59	toi
S58	facit
S57	qlc-c30
S56	"fact survey"
S55	fact questionnaire
S54	(MH "Questionnaires")
S53	(MH "Functional Assessment")
S52	(MH "Pain Measurement")
S51	(MH "Mental Status")
S50	"psychiatric status"
S49	"physical limitations"
S48	(MH "Social Adjustment")
S47	"patient reported outcomes"
S46	healthiness
S45	good health
S44	(MH "Functional Status")
S43	(MH "Activities of Daily Living") OR "daily activities"
S42	(MH "Psychological Well-Being")
S41	(MH "Wellness")
S40	level of health
S39	"health level"
S38	(MH "Self Concept")
S37	"life quality"

(Continued)

S36	(MH "Activities of Daily Living")
S35	(MH "Health Status")
S34	(MH "Quality of Life")
S33	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30 or S31 or S32
S32	(MH "Aquatic Exercises")
S31	(MH "Swimming")
S30	"balance exercise"
S29	(MH "Hydrotherapy")
S28	"hippotherapy" OR (MH "Horseback Riding")
S27	"bad ragaz"
S26	(MH "Alternative Therapies+")
S25	(MH "Mind Body Techniques") OR "mind body therapy"
S24	(MH "Muscle Strengthening") OR "resistance training"
S23	"chi kung"
S22	(MH "Qigong")
S21	(MH "Pilates")
S20	(MH "Range of Motion")
S19	"muscle strengthening exercises"
S18	(MH "Yoga")
S17	tai ji
S16	(MH "Tai Chi")
S15	(MH "Exercise Test")
S14	(MH "Sports+")
S13	(MH "Motor Activity")
S12	"motor processes"
S11	"stamina"
S10	"physical condition*"

(Continued)

S9	"exercising"
S8	(MH "Therapeutic Exercise")
S7	(MH "Physical Endurance")
S6	(MH "Physical Education and Training+")
S5	(MH "Physical Fitness+")
S4	(MH "Pliability")
S3	(MH "Exertion")
S2	(MH "Exercise Tolerance")
S1	(MH "Exercise+")

Appendix 5. PsycINFO search strategy

[inception to July 2009; 120 hits] [January 2009 to September 2011; 45 hits]

1. exp exercise/
2. exercise tolerance.mp.
3. exertion.mp.
4. pliability.mp.
5. exp physical fitness/
6. (physical education and training).mp. [mp=title, abstract, heading word, table of contents, key concepts]
7. exp Physical Endurance/
8. exercise therapy.mp.
9. exercising.mp.
10. physical condition$.mp.
11. stamina.mp.
12. Motor Processes/
13. motor activity.mp.
14. exp Sports/
15. exercise test.mp.
16. (tai chi or tai ji).mp. [mp=title, abstract, heading word, table of contents, key concepts]
17. exp Yoga/
18. muscle strengthening exercises.mp.
19. "Range of Motion"/
20. pilates.mp.
21. qigong.mp.
22. chi kung.mp.
23. resistance training.mp.
24. mind body therapy.mp.
25. exp Alternative Medicine/
26. bad ragaz.mp.
27. ai chi.mp.
28. halliwick.mp.
29. hippotherapy.mp.

30.hydrotherapy.mp.
31.balance exercise$.mp.
32.Swimming/ or aquatic exercise.mp.
33.1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
34.exp "quality of life"/
35.health status.mp.
36.exp "Activities of Daily Living"/
37.life quality.mp.
38.exp Self Concept/
39.health level.mp.
40.level of health.mp.
41.wellness.mp. or exp Health/
42.Well Being/
43.Daily Activities/ or activities of daily life.mp.
44.functional ability.mp.
45.good health.mp.
46.healthiness.mp.
47.patient reported outcomes.mp.
48.exp Social Adjustment/
49.physical limitations.mp.
50.psychiatric status.mp.
51.exp Pain Measurement/
52.functional assessment.mp.
53.Questionnaires/ or fact questionnaire.mp.
54.fact survey.mp.
55.qlc-c30.mp.
56.facit.mp.
57.toi.mp.
58.(flic or sf-36 or ces-d or bdi or stal or bfi or hads or lasa or poms or qli or rsci or pais or bpi or msas or mos or ptgi or panas).mp. [mp=title, abstract, heading word, table of contents, key concepts] (8685)
59."Sense of Coherence"/
60.35 or 53 or 48 or 42 or 46 or 44 or 55 or 50 or 39 or 57 or 36 or 40 or 51 or 58 or 41 or 47 or 59 or 52 or 38 or 34 or 56 or 49 or 37 or 45 or 43 or 54
61.exp Neoplasms/
62.Survivors/
63.cancer survivors.mp.
64.63 or 61 or 62
65.60 and 33 and 64
66.66 limit 65 to (("0400 empirical study" or "0830 systematic review" or 1200 meta analysis or "2000 treatment outcome/randomized clinical trial") and "0100 journal")
67.randomized controlled trial.mp.
68.clincal trial.mp.
69.random$.ab.
70.placebo.ab.
71.trial.ab.
72.69 or 67 or 71 or 70 or 68
73.72 and 65
74.73 or 66

Appendix 6. Other search strategies

LILACS search strategy [inception to August 2009; 15 hits] [January 2009 to September 2011; 0 hits]

(Neoplasms OR cancer) AND exercise AND "quality of life"

OT Seeker search strategy [inception to August 2009; 45 hits] [January 2009 to October 2011; 13 hits]

Database Note returned with search: A precise search did not find any articles. A less precise search has been done and the results are shown below.

(exercise OR exertion OR pliability OR "physical fitness" OR "physical endurance" OR "exercise therapy" OR "motor activity" OR sports) AND cancer AND "quality of life" AND survivor*

Limits: Method: clinical trial and Diagnosis/Subdiscipline: Oncology/palliative care

PEDro search strategy [inception to August 2009; 67 hits] [January 2009 to September 2011; 15 hits]

exercise AND cancer AND "quality of life" AND survivor*

SIGLE search strategy (now OpenGrey) [inception to November 2009; 0 hits] [January 2009 to October 2011; 4 hits]

exercise AND (cancer OR neoplasms) AND "quality of life AND (survivors OR post treatment)

Sociological Abstracts (SocINDEX) search strategy [inception to November 2009; 29 hits] [January 2009 to September 2011; 10 hits]

Search ID#	Search Terms
S73	S68 and S72
S72	survivors
S68	S66 and S67
S67	(random* or placebo or trial)
S66	S33 and S62 and S65
S65	S63 or S64
S64	(MH "Cancer Survivors") OR (MH "Survivors")
S63	(MH "Neoplasms+")
S62	S34 or S35 or S36 or S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45 or S46 or S47 or S48 or S49 or S50 or S51 or S52 or S53 or S54 or S55 or S58 or S59 or S60 or S61
S61	"sense of coherence"
S60	(flic or sf-37 or ces-d of bdi or stal or bfi or hads or lasa or poms or qli or rsci or pals or bpi or msas or mos or ptgi or panas)
S59	toi
S58	facit
S57	qlc-c30
S56	"fact survey"
S55	fact questionnaire
S54	(MH "Questionnaires")

(Continued)

S53	(MH "Functional Assessment")
S52	(MH "Pain Measurement")
S51	(MH "Mental Status")
S50	"psychiatric status"
S49	"physical limitations"
S48	(MH "Social Adjustment")
S47	"patient reported outcomes"
S46	healthiness
S45	good health
S44	(MH "Functional Status")
S43	(MH "Activities of Daily Living") OR "daily activities"
S42	(MH "Psychological Well-Being")
S41	(MH "Wellness")
S40	level of health
S39	"health level"
S38	(MH "Self Concept")
S37	"life quality"
S36	(MH "Activities of Daily Living")
S35	(MH "Health Status")
S34	(MH "Quality of Life")
S33	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30 or S31 or S32
S32	(MH "Aquatic Exercises")
S31	(MH "Swimming")
S30	"balance exercise"
S29	(MH "Hydrotherapy")
S28	"hippotherapy" OR (MH "Horseback Riding")
S27	"bad ragaz"

(Continued)

S26	(MH "Alternative Therapies+")
S25	(MH "Mind Body Techniques") OR "mind body therapy"
S24	(MH "Muscle Strengthening") OR "resistance training"
S23	"chi kung"
S22	(MH "Qigong")
S21	(MH "Pilates")
S20	(MH "Range of Motion")
S19	"muscle strengthening exercises"
S18	(MH "Yoga")
S17	tai ji
S16	(MH "Tai Chi")
S15	(MH "Exercise Test")
S14	(MH "Sports+")
S13	(MH "Motor Activity")
S12	"motor processes"
S11	"stamina"
S10	"physical condition*"
S9	"exercising"
S8	(MH "Therapeutic Exercise")
S7	(MH "Physical Endurance")
S6	(MH "Physical Education and Training+")
S5	(MH "Physical Fitness+")
S4	(MH "Pliability")
S3	(MH "Exertion")
S2	(MH "Exercise Tolerance")
S1	(MH "Exercise+")

SportDiscus search strategy [inception to August 2009; 31 hits] [January 2009 to September 2011; 28 hits]

Search ID#	Search Terms
S73	S33 and S58 and S71 and S72
S72	S64 or S65 or S66 or S67 or S68 or S69
S71	S59 or S60 or S61 or S62 or S63 or S70
S70	random*
S69	after treatment
S68	post treatment
S67	cancer
S66	cancer survivor*
S65	neoplasms
S64	survivors
S63	controlled clinical trial
S62	randomized controlled trial
S61	trial
S60	placebo
S59	randomized
S58	S34 or S35 or S36 or S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45 or S46 or S47 or S48 or S49 or S50 or S51 or S52 or S53 or S54 or S55 or S56 or S57
S57	sense of coherence
S56	flic or sf-36 or ces-d or bdi or sta1 or bfi or hads or lasa or poms or qli or rsci or pais or bpi or msas or mos or ptgi or panas
S55	toi
S54	facit
S53	qlc-c30
S52	fact survey
S51	fact questionnaire
S50	functional assessment
S49	pain measurement
S48	psychiatric status

(Continued)

S47	physical limitations
S46	social adjustment
S45	patient reported outcomes
S44	healthiness
S43	good health
S42	functional ability
S41	activities of daily living OR daily living activities
S40	wellness
S39	level of health
S38	health level
S37	self concept
S36	life qualit*
S35	DE "ACTIVITIES of daily living"
S34	DE "QUALITY of life" OR DE "HEALTH status indicators" OR DE "LIFESTYLES" OR DE "WELL-being"
S33	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30 or S31 or S32
S32	DE aquatic exercises
S31	balance exercise*
S30	DE hydrotherapy
S29	hippotherapy
S28	halliwick
S27	ai chi
S26	bad ragaz
S25	complementary therap*
S24	mind body therap*
S23	resistance training
S22	DE "WEIGHT training" OR DE "BENCH press" OR DE "DEAD lift (Weight lifting)" OR DE "POWER-LIFTING" OR DE "SQUAT (Weight lifting)" OR DE "STONE lifting" OR DE "WEIGHT lifting"
S21	chi kung

(Continued)

S20	qigong
S19	DE pilates
S18	DE "JOINTS -- Range of motion"
S17	muscle stretching exercise*
S16	DE yoga
S15	DE tai chi
S14	DE sports
S13	exercise test
S12	motor activity
S11	stamina
S10	physical condition*
S9	exercising
S8	DE "Exercise Therapy"
S7	physical endurance
S6	DE "physical education & training"
S5	DE "physical fitness"
S4	pliability
S3	exertion
S2	exercise tolerance
S1	DE Exercise

WHAT'S NEW

Date	Event	Description
12 May 2020	Amended	The Editors are looking for contributors to update and maintain this Cochrane Review. **Contact ruh-tr.gnoc-cochrane@nhs.net for further information**. The searches have been updated to May 2019 and potentially relevant studies added to 'Other references; Classification pending'.

HISTORY

Protocol first published: Issue 1, 2009
Review first published: Issue 8, 2012

CONTRIBUTIONS OF AUTHORS

Shiraz I. Mishra: content expert; contributed by conceptualization of the review, identifying trials eligible for the review, extracting data for trials meeting the inclusion criteria, preparing the data tables, analyzing the data, and writing the review.

Roberta W. Scherer: methodologic expert; contributed by identifying trials eligible for the review, extracting data for trials meeting the inclusion criteria, preparing the data tables, analyzing the data, and contributing to writing the review.

Paula M. Geigle: content expert; contributed by extracting data for trials meeting the inclusion criteria and providing editorial input.

Debra R. Berlanstein: information specialist; contributed to development of the search strategy, conducting all the electronic database searches, and retrieving potentially eligible trials.

Ozlem Topaloglu: contributed by extracting data for trials meeting the inclusion criteria.

Carolyn C. Gotay: content expert; contributed by conceptualization of the review, providing editorial input, data collection, design and implementing of the search strategy, extraction of data for trials meeting the inclusion criteria.

Claire Snyder: content expert; contributing by providing editorial input and expertise on analysis and interpretation of HRQoL measures.

DECLARATIONS OF INTEREST

The authors declare no conflict of interest.

SOURCES OF SUPPORT

Internal sources

- None, Other

External sources

- Cancer Restitution Funds to the State of Maryland, USA

  Restitution funds from tobacco companies awarded to the University of Maryland
- National Institute for Health Research (NIHR) Health Technology Assessment programme, UK

  HTA Project: 10/81/01 - Exercise interventions for the management of health related quality of life and fatigue in cancer survivors during and after treatment

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

This review included trials that included both participants who were undergoing active cancer treatment for their primary or recurrent cancer and those who had completed active treatment for their cancer. This revised inclusion criteria was applied to studies that did not have the majority of participants undergoing active treatment for their primary or recurrent cancer. As discussed previously, for only two outcomes (depression and social functioning) were there more than one trial with both participants who were currently undergoing active cancer treatment and those who had completed active treatment for their cancer. We observed no difference in findings when trials including participants currently undergoing active treatment for their cancer were excluded from the analyses.

We have included a 'Summary of findings' table instead of calculating numbers needed to treat for a beneficial effect (NNTB) for the review findings.

INDEX TERMS

Medical Subject Headings (MeSH)

*Exercise; *Health Status; Neoplasms [*rehabilitation]; *Quality of Life; Randomized Controlled Trials as Topic; *Survivors

MeSH check words

Humans