Cochrane Library

Cochrane Database of Systematic Reviews

Exercise therapies for preventing or treating aromatase inhibitor-induced musculoskeletal symptoms in early breast cancer (Review)

Roberts KE, Rickett K, Feng S, Vagenas D, Woodward NE

Roberts KE, Rickett K, Feng S, Vagenas D, Woodward NE.
Exercise therapies for preventing or treating aromatase inhibitor-induced musculoskeletal symptoms in early breast cancer.
Cochrane Database of Systematic Reviews 2020, Issue 1. Art. No.: CD012988.
DOI: 10.1002/14651858.CD012988.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Exercise therapies for preventing or treating aromatase inhibitor-induced musculoskeletal symptoms in early breast cancer

Kate E Roberts[1,2], Kirsty Rickett[3], Sophie Feng[4], Dimitrios Vagenas[5], Natasha E Woodward[2,4]

[1]Department of Medical Oncology, Princess Alexandra Hospital, Woolloongabba, Australia. [2]School of Clinical Medicine, Mater Clinical Unit, Mater Hospital, University of Queensland, South Brisbane, Australia. [3]The University of Queensland Library, UQ/Mater McAuley Library, Brisbane, Australia. [4]Department of Medical Oncology, Mater Misericordiae Ltd, South Brisbane, Australia. [5]Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia

Contact: Kate E Roberts, Department of Medical Oncology, Princess Alexandra Hospital, Ipswich Road, Woolloongabba, Queensland, Australia. kate.roberts@health.qld.gov.au, katee.roberts@gmail.com.

Editorial group: Cochrane Breast Cancer Group.
Publication status and date: New, published in Issue 1, 2020.

Citation: Roberts KE, Rickett K, Feng S, Vagenas D, Woodward NE. Exercise therapies for preventing or treating aromatase inhibitor-induced musculoskeletal symptoms in early breast cancer. *Cochrane Database of Systematic Reviews* 2020, Issue 1. Art. No.: CD012988. DOI: 10.1002/14651858.CD012988.pub2.

ABSTRACT

Background

Survival for stage I to III, hormone receptor-positive, breast cancer has substantially improved over time due to advances in screening, surgery and adjuvant therapy. However many adjuvant therapies have significant treatment-related toxicities, which worsen quality of life for breast cancer survivors. Postmenopausal women with hormone receptor-positive breast cancer are now prescribed aromatase inhibitors (AI) as standard, with longer durations of therapy, up to 10 years, being considered for certain women. AI treatment is associated with a high incidence of AI-induced musculoskeletal symptoms (AIMSS), often described as symmetrical pain and soreness in the joints, musculoskeletal pain and joint stiffness. AIMSS reduces compliance with AI therapy in up to one half of women undergoing adjuvant AI therapy, potentially compromising breast cancer outcomes. Exercise has been investigated for the prevention and treatment of AIMSS but the effect of this intervention remains unclear.

Objectives

To assess the effects of exercise therapies on the prevention or management of aromatase inhibitor-induced musculoskeletal symptoms (AIMSS) in women with stage I to III hormone receptor-positive breast cancer.

Search methods

We searched Cochrane Breast Cancer's Specialised Register, CENTRAL, MEDLINE, Embase and CINAHL databases up to 13 December 2018. We also searched two conference proceedings portals and two clinical trials registries for ongoing studies or unpublished trials, or both, in August 2019. We also reviewed reference lists of the included studies.

Selection criteria

We included randomised controlled trials that compared exercise versus a comparator arm. We did not impose any restriction on the comparator arm, which could include an alternative type of exercise, no exercise or a waiting list control. Both published and non-peer-reviewed studies were eligible.

Data collection and analysis

Two review authors independently extracted data, assessed risk of bias and certainty of the evidence using the GRADE approach. The outcomes investigated were pain, joint stiffness, grip strength, health-related quality of life, cancer-specific quality of life, adherence to

AI therapy, adverse events, incidence of AIMSS, breast cancer-specific survival and overall survival. For continuous outcomes that were assessed with the same instrument, we used the mean difference (MD); for those outcomes that used different instruments, we used the standardised mean difference (SMD) for the analysis. For dichotomous outcomes, we reported outcomes as an odds ratio (OR).

Main results

We included seven studies with 400 randomised participants; one study assessed exercise for preventing AIMSS and six studies assessed treating AIMSS.

For preventing AIMSS, the single study reported no difference in pain scores, grip strength or compliance to taking AI medication between groups. Data values were not provided in the study and no other outcomes were reported.

For managing AIMSS, we found that the evidence for the effect of exercise therapies on overall change in worst pain scores was very uncertain (SMD −0.23, 95% confidence interval (CI) −0.78 to 0.32; 4 studies, 284 women; very low-certainty evidence). The evidence suggested that exercise therapies result in little to no difference in overall change in stiffness scores (Western Ontario McMasters Universities Osteoarthritis Index (WOMAC) stiffness score MD −0.76, 95% CI −1.67 to 0.15 and Visual Analogues Scale (VAS) stiffness score MD −0.42, 95% CI −2.10 to 1.26; 1 study, 53 women; low-certainty evidence). The evidence was very uncertain for the outcomes of overall change in grip strength (MD 0.30, 95% CI −0.55 to 1.15; 1 study, 83 women; very low-certainty evidence); overall change in health-related quality of life (subscales of SF-36 tool ranged from least benefit of MD 1.88, 95% CI −2.69 to 6.45 to most benefit of MD 9.70, 95% CI 1.67 to 17.73; 2 studies, 123 women, very low-certainty evidence); overall change in cancer-specific quality of life (MD 4.58, 95% CI −0.61 to 9.78; 2 studies, 136 women; very low-certainty evidence); and adherence to aromatase inhibitors (OR 2.43, 95% CI 0.41 to 14.63; 2 studies, 224 women; very low-certainty evidence). There were no adverse events identified across four studies in either arm (0 events reported; 4 studies; 331 participants; low-certainty evidence). There were no data reported on incidence of AIMSS, breast cancer-specific survival or overall survival.

Authors' conclusions

Given the wide-ranging benefits of exercise for people affected by cancer, it was surprising that this review provided no clear evidence of benefit for exercise therapies in women with early breast cancer with AIMSS. This review only yielded seven eligible studies with 400 participants, which is likely to have underpowered the findings. The meta-analysis was challenging due to the considerable heterogeneity amongst the trials, with a wide range of exercise regimens and follow-up periods. Despite these inconclusive findings, exercise needs to be part of routine care for women with breast cancer due to its wide-ranging benefits. Future research in this area would be enhanced with further understanding of the mechanism of AIMSS, a single clear definition of the condition, and phase III randomised controlled trials that are adequately powered to test targeted exercise interventions on the key clinical outcomes in this condition.

PLAIN LANGUAGE SUMMARY

Exercise therapies for preventing or treating aromatase inhibitor-induced musculoskeletal symptoms in early breast cancer

What is the aim of this review?

Aromatase inhibitors (AI) are a hormonal therapy used to treat a particular type of breast cancer in post-menopausal women. However, they can cause joint and muscle pain (called aromatase inhibitor-induced musculoskeletal symptoms, or AIMSS). The aim of this Cochrane Review was to find out whether exercise therapies can reduce this pain in women undergoing treatment for early breast cancer.

Key messages

It is unclear if exercise improves, worsens, or makes no difference to pain, quality of life, grip strength, or the number of women continuing to take AI medication. Exercise likely results in little to no difference in stiffness in women suffering from AIMSS, although the certainty of this evidence was also low. Exercise is probably safe in women with AIMSS.

What was studied in the review?

Studies have shown a survival benefit for women when they take AIs for five to ten years after surgery, but unfortunately, they are associated with musculoskeletal side effects that may cause some women to stop taking their medication, which may have an impact on their survival. We looked at whether exercise could help prevent or treat the joint pains, stiffness and muscle aches from AIs that are being taken by women with breast cancer to prevent a recurrence. We looked at research studies of exercise compared to either usual care, being on a waiting list for an exercise treatment, or another exercise like walking, in women who had AIMSS. Women aged 18 years or older with early stage breast cancer being treated with AI were included. In most studies, the women had to have joint or muscle pains whilst being treated with an AI.

We studied a number of outcomes, including changes in pain, stiffness, hand strength (grip strength), the number of women continuing to take AI medication, the quality of life of women on AI medication, and the safety of the exercise programmes.

What are the main results of the review?

We collected and analysed all relevant studies to answer this question and found seven studies with 400 women. The studies included different numbers of women, ranging from 20 to 121 participants. Three studies were conducted in the USA, one study in the UK, one study in Australia, one study in Canada and one study in Japan. Overall, the certainty of the evidence for most outcomes was very low. This may have been because many of the studies did not have many participants, making it hard to find small differences. Other problems were that the women and the people assessing the results, knew which exercise therapy the woman was receiving, and this may have introduced bias. Many studies did not report all of their results, and some of the studies were not carried out to a high research standard.

Therefore it is unclear whether exercise has a positive or negative effect on pain, grip strength, the number of women continuing to take AI medication, or the quality of life of women with AIMSS, because of the very low certainty of the evidence. Exercise likely results in little to no change in stiffness in women suffering from AIMSS. Importantly, exercise is probably safe, with no harms reported, although the studies did not follow up the women for very long. There were no data available to assess the effect of exercise on survival in women with AIMSS. Despite these inconclusive findings, exercise should still be recommended as part of routine care for women with breast cancer, due to its wide-ranging benefits.

How up to date is this review?

The last search for studies in this review was performed in December 2018 and the search for ongoing studies was conducted in August 2019.

Summary of findings for the main comparison. Exercise therapies compared to standard care for the management of aromatase inhibitor-induced musculoskeletal symptoms

Exercise therapies compared to standard care for the management of aromatase inhibitor-induced musculoskeletal symptoms

Patient or population: women with aromatase inhibitor-induced musculoskeletal symptoms (AIMSS)
Setting: outpatient
Intervention: exercise therapies
Comparison: standard care

Outcomes	Anticipated absolute effects* (95% CI) with the use of exercise	Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
Overall change in worst pain scores	SMD 0.23 SD lower (0.78 lower to 0.32 higher)	-	284 (4 RCTs)	⊕⊖⊖⊖ Very low[1,2,3,4]	The evidence is very uncertain about the effect of exercise therapies on overall change in worst pain scores.
Overall change in stiffness scores	The effect in this single study ranged from MD 0.76 points lower (1.67 lower to 0.15 higher) to MD 0.42 points lower (2.10 lower to 1.26 higher)	-	53 (1 RCT)	⊕⊕⊖⊖ Low[5,6]	The evidence suggests that exercise therapies result in little to no difference in overall change in stiffness scores.
Overall change in grip strength	MD 0.30 points higher (0.55 lower to 1.15 higher)	-	83 (1 RCT)	⊕⊖⊖⊖ Very low[6,7,8]	The evidence is very uncertain about the effect of exercise therapies on overall change in grip strength.
Overall change in health-related quality of life	We could not calculate total score. Effect within subscales of HR-QoL ranged from MD 1.88 points higher (2.69 lower to 6.45 higher) to 9.70 points higher (1.67 higher to 17.73 higher)	-	123 (2 RCTs)	⊕⊖⊖⊖ Very low[1,3,9,10]	The evidence is very uncertain about the effect of exercise therapies on overall change in health-related quality of life.
Overall change in cancer-specific quality of life	MD 4.58 points higher (0.61 lower to 9.78 higher)	-	136 (2 RCTs)	⊕⊖⊖⊖ Very low[1,3,11,12]	The evidence is very uncertain about the effect of exercise therapies on overall change in disease-specific quality of life.
Adverse effects secondary to the intervention	Nil adverse events in either arm.	Not estimable	331 (4 RCTs)	⊕⊕⊖⊖ Low[1,13]	The evidence suggests that exercise therapies are low risk, with no adverse events reported across four studies

Adherence to aromatase inhibitors	-	OR 2.43 (0.41 to 14.63)	224 (2 RCTs)	⊕⊖⊖⊖ Very low[3,13,14]	The evidence is very uncertain about the effect of exercise therapies on adherence to aromatase inhibitors.

*__The risk in the intervention group__ (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).
AIMSS: aromatase inhibitor-induced musculoskeletal symptoms; **CI:** confidence interval; **HRQoL:** health-related quality of life; **RR:** risk ratio; **OR:** odds ratio; **SD:** standard deviation; **SMD:** standardised mean difference; **MD:** mean difference; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Risk of bias concerns with each study, including lack of blinding of participants/personnel and outcome assessors. There were serious concerns regarding random sequence allocation and allocation concealment in one study. High risk of attrition bias in three studies. Concerns regarding exercise contamination in the control arm in two studies. Downgraded 1 point.
[2]Significant statistical heterogeneity, I^2 = 79%, resulting in downgrading 1 point for inconsistency.
[3]Multiple studies only written in abstract form, without pursuing full publication, including studies that had unpublished results relevant to this outcome. Strong suspicion of publication bias, and therefore downgraded 1 point.
[4]Downgraded 1 point for imprecision, due to a number of factors: sample sizes were too small to determine an accurate result; the width of the confidence interval is consistent with both benefit and harm; and one of the studies included skewed data.
[5]High risk of bias for this study, due to serious concerns with random sequence generation and allocation concealment, and also the lack of blinding of participants/personnel and outcome assessors. In addition, high risk of attrition bias. Selective reporting bias for one study for this outcome, which did not report results. Downgraded 1 point.
[6]Small number of participants and null effect and appreciable harm and benefit included in the confidence interval. Downgraded 1 point.
[7]Two studies did not report grip-strength results, as only published in abstract form. Downgraded 1 point for publication bias.
[8]Downgraded 1 point for risk of bias, due to inability to blind participants/personnel to the intervention, and lack of blinding for outcome assessors. Concerns regarding incomplete outcome data and exercise contamination in the control arm.
[9]Imprecision was present, due to wide range of confidence intervals, a sample size that was too small to provide accurate results, and inclusion of one study that had skewed data. Downgraded 1 point.
[10]Downgraded 1 point for risk of bias. Lack of blinding for participants/personnel, and inadequate allocation concealment. Judged as high risk of attrition bias in one study, and concerns regarding exercise contamination in the control arms of both studies. Poor adherence to exercise in one study.
[11]Downgraded 1 point for imprecision, due to small sample size and wide confidence intervals, which included both the null effect and appreciable benefit.
[12]Risk of bias concerns with each study, including lack of blinding of participants/personnel and outcome assessors. There were serious concerns regarding random sequence allocation and allocation concealment in one study. High risk of attrition bias in both studies. In addition, concerns regarding exercise contamination in the control arm in one study. Downgraded 1 point.
[13]Downgraded 1 point for imprecision, because the sample size was small, and the event rate low.
[14]Downgraded 1 point due to high risk of bias with each study, including high risk of attrition bias in both studies, and one study only being published in abstract form so limited data available. It is unclear how much lack of blinding of participants and personnel may have impacted on this outcome.

BACKGROUND

Description of the condition

Breast cancer remains a major public health problem despite advances in screening and treatment. There was an estimated 1.67 million new cases diagnosed in 2012, making breast cancer the most common non-skin cancer in women (Ferlay 2012). With 522,000 deaths, breast cancer was the fifth most common cause of cancer death globally in 2012 (Ferlay 2012). In women in high-income countries, breast cancer is second to lung cancer as the leading cause of cancer death, and in low- to middle-income countries, breast cancer remains the leading cause of cancer death (Ferlay 2012). Hormone receptor-positive breast cancer, that is, oestrogen receptor (ER)-positive, or progesterone receptor (PR)-positive, or both, accounts for about 80% of breast cancer, with women with early breast cancer usually having oestrogen or 'endocrine-sensitive' cancer (Nadji 2005). Treatment of postmenopausal women with hormone receptor-positive breast cancer with aromatase inhibitor (AI) medications is effective. Five years of AI therapy in early breast cancer improves disease-free survival (DFS) and breast cancer specific survival (BCSS) when compared to another hormonal therapy, tamoxifen (Aydiner 2013; EBCTCG 2015). Recent guidelines (Burstein 2019) now recommend consideration of 10 years of AI treatment for certain high-risk subgroups, such as node-positive patients.

However, AIs are commonly associated with joint and muscular symptoms, referred to as aromatase inhibitor-induced musculoskeletal symptoms (AIMSS; Lintermans 2013). Nearly half of all women on AIs experience these side effects (Beckwee 2017). AIMSS often presents as symmetrical pain or soreness in multiple joints, and is also often associated with early morning stiffness (Burstein 2007). Despite the survival advantage of AIs, these side effects are causing a quarter to half of all women on this treatment to discontinue (Chim 2013; Henry 2012; Kadakia 2016). An association between switching AIs and the development of new musculoskeletal pain has been identified (Kemp-Casey 2017). If AIMSS can be managed, then quality of life and adherence to treatment may improve, and the survival advantage from using AI therapy may not be compromised.

Description of the intervention

Exercise can be defined as "a subset of physical activity that is planned, structured, repetitive, and has as a final or an intermediate objective of the improvement or maintenance of physical fitness" (Caspersen 1985). The definition of therapy in the Merriam-Webster dictionary is the "therapeutic treatment especially of bodily, mental, or behavioral disorder" (Merriam-Webster). Exercise therapies investigated in this review involve a variety of therapeutic interventions intended to improve or maintain fitness. These include, but are not restricted to, cardiovascular and resistance exercises, yoga, tai-chi, aquatic exercise, walking and pilates.

How the intervention might work

The cause of AIMSS is unknown, and therefore the mechanism for the effectiveness of exercise therapies on AIMSS cannot be ascertained. There has been a growing interest in conducting research into the effect of exercise on a wide variety of conditions, such as the effect on cancer mortality, recurrence and treatment-related adverse effects (Cormie 2017), cancer-related fatigue and mobility (Dennett 2016), quality of life in cancer survivors (Mishra 2012), the immune system (Szlezak 2016), and rheumatological conditions, such as osteoarthritis (Fransen 2014; Osteras 2017). There has been a large phase III randomised controlled trial (RCT) investigating the intervention of cardiovascular and resistance exercise in the treatment of AIMSS, which reported a clinically significant benefit with the use of exercise (Irwin 2015). Therefore, even though the mechanism of any potential benefit of exercise in this area is largely unknown, a positive result from a large phase III RCT, plus multiple other smaller studies in this field, warrants a comprehensive review of these therapies.

Why it is important to do this review

AIMSS has a clinical impact on the management of women with breast cancer, as studies have shown substantial rates of suboptimal adherence to AIs (Brier 2017; Hadji 2014; Henry 2012; Hershman 2011; Partridge 2008; Presant 2007). Non-compliance with endocrine therapies in the adjuvant setting may impact on women's survival (Hershman 2011). To date, there is limited evidence regarding the best management options for AIMSS. With a growing emphasis on the management of survivorship issues for women with early breast cancer, this area of research is very topical, and of increasing importance. It has been identified that oncologists may feel ill-equipped to prescribe exercise to women with early breast cancer (Smaradottir 2017), and providing a stronger evidence base for the role of exercise in managing symptoms may assist with this issue.

OBJECTIVES

To assess the effects of exercise therapies on the prevention or management of aromatase inhibitor-induced musculoskeletal symptoms (AIMSS) in women with stage I to III hormone receptor-positive breast cancer.

METHODS

Criteria for considering studies for this review

Types of studies

All RCTs looking at the prevention or management of AIMSS in women with stage I to III hormone receptor-positive breast cancer. AIMSS was defined by the study authors of each trial. We excluded animal and in vitro studies. We considered studies in all languages for inclusion.

Types of participants

Women aged 18 years and older with stage I to III ER-positive, or PR-positive breast cancer, or both, who were being treated adjuvantly with AIs.

Types of interventions

We included all exercise therapy interventions, such as aerobic and resistance exercise, tai chi, yoga and aqua aerobics. We excluded musculoskeletal manipulation therapies, such as massage and kinesiology. We did not impose any restriction on the type of comparator arm; comparator arms could include an alternative type of exercise, no exercise, or a waiting list control.

Types of outcome measures

Primary outcomes

- Prevention and treatment of symptoms of AIMSS (pain, stiffness, and grip strength) from baseline. These were preferably assessed by validated questionnaires, such as the Visual Analogue Scale (VAS), Brief Pain Inventory (BPI), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Functional Assessment of Cancer Therapy – General (FACT-G), Medical Outcome Study Short Form 36 (SF-36), and the Modified Score for the Assessment of Chronic Rheumatoid Affections of the Hands (M-SACRAH)
- Safety of the intervention, including adverse effects, such as injury

Secondary outcomes

- Incidence of AIMSS
- Persistence and compliance of women continuing to take their AI medication due to the intervention
- Participant health-related quality of life, which was also preferably assessed by validated patient/participant-reported outcome questionnaires
- Participant cancer-specific quality of life
- Breast cancer-specific survival
- Overall survival

Search methods for identification of studies

Electronic searches

The Information Specialist (KR) designed and conducted systematic searches in the selected databases and trial registries without language, publication year or publication status restrictions. Cochrane Breast Cancer's Information Specialist conducted the search of the group's Specialised Register. Where appropriate, the search strategies also included adaptations of the highly sensitive search strategy designed by the Cochrane Collaboration (Lefebvre 2011), and the search filter for CINAHL (EBSCO) created by Mark Clowes at SIGN for identifying RCTs and controlled clinical trials.

We searched the following databases and trials registries.

- Cochrane Breast Cancer's Specialised Register. We extracted and considered for inclusion in the review trials with the key words "breast cancer" and related terms, "aromatase inhibitors", "exemestane", "anastrozole", "letrozole", "exercise", "physical activity", "resistance training", "yoga", "walking", "T'ai chi"; searched on 16 April 2018.
- Cochrane Central Register of Controlled Trials (CENTRAL; 2018, Issue 11) in the Cochrane Library (searched 13 December 2018). See Appendix 1
- MEDLINE (via PubMed) from 1946 to December 2018 (searched 13 December 2018). See Appendix 2
- Embase (via Embase.com) from 1947 to December 2018 (searched 13 December 2018). See Appendix 3
- CINAHL (via EBSCO) from 1937 to present. (Last search 13 December 2018). See Appendix 4
- World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) search portal (apps.who.int/trialsearch) for all prospectively registered and ongoing trials (searched on 18 August 2019). See Appendix 5
- Clinicaltrials.gov (clinicaltrials.gov; searched on 18 August 2019). See Appendix 6

Searching other resources

Bibliographic searching

We searched reference and citation lists of identified relevant trials and reviews to try and identify further studies. We attempted to obtain a copy of the full article for each reference reporting a potentially eligible trial. Where this was not possible, such as with the inclusion of conference abstracts, we sourced additional information from clinical trials databases, and we attempted to contact study authors to provide additional information.

Grey searching

We screened conference abstracts from major conferences such as the San Antonio Breast Cancer Symposium (SABCS) and American Society of Clinical Oncology (ASCO) up to December 2018 and any additional papers identified during the attendance at the 2019 San Antonio Breast Cancer Symposium (NW) were reported and added for inclusion, where relevant.

Data collection and analysis

Selection of studies

Review authors (SF, NW, KER and KR) screened retrieved abstracts from the literature search and assessed whether the abstracts met the specified selection criteria. Subsequently, we reviewed the full texts of all remaining studies to ensure that they still met the selection criteria. At least two review authors reviewed each study to ensure that they met the selection criteria. We resolved any disagreements on study selection by involving a separate review author (NW or KER). We recorded the selection process in a PRISMA flow diagram (Figure 1; Moher 2009). We documented the reason for excluding any studies in the Characteristics of excluded studies tables. There were no studies reported in languages other than English identified during this search, and therefore no translation was required.

Figure 1. Study flow diagram

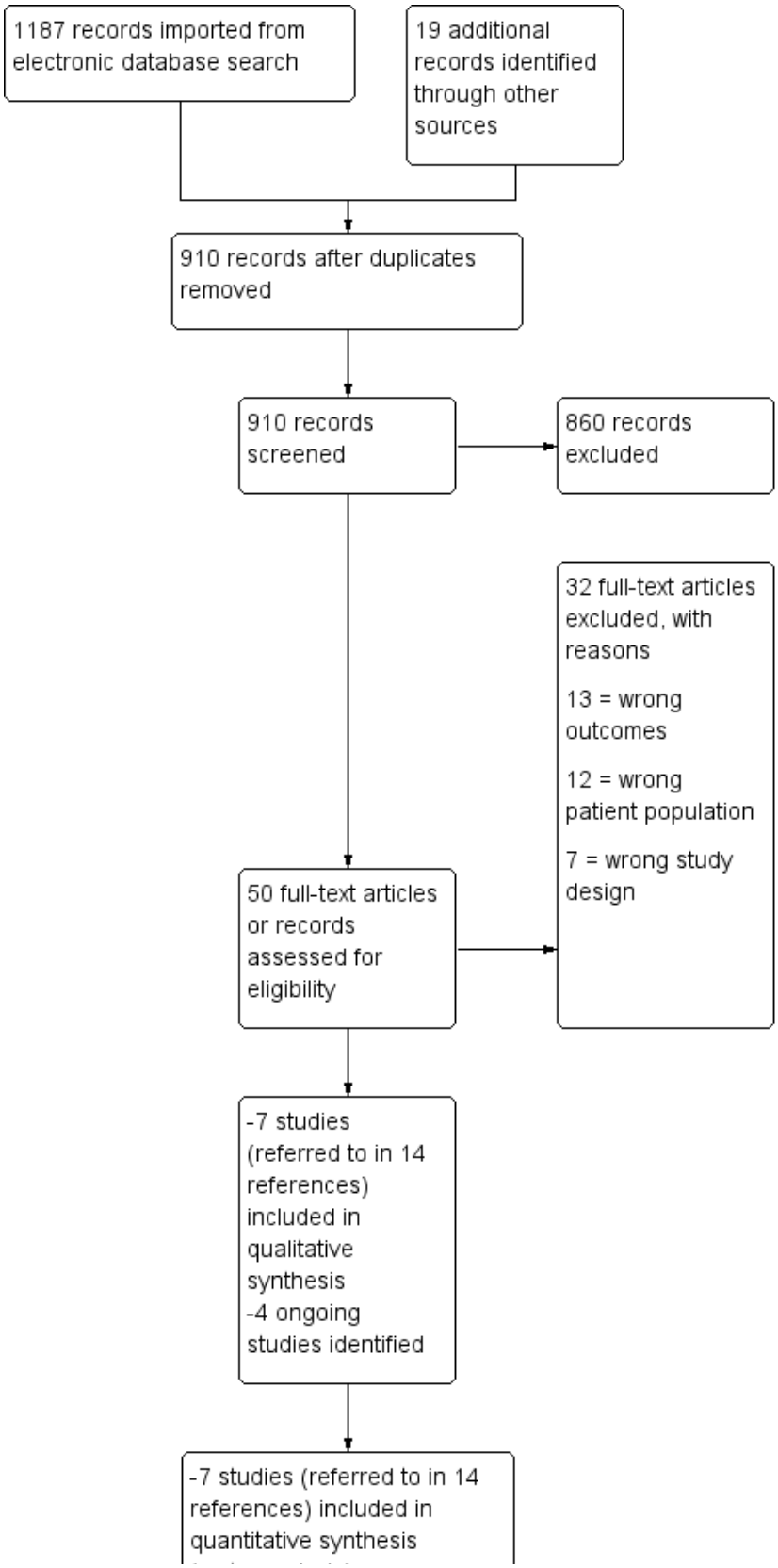
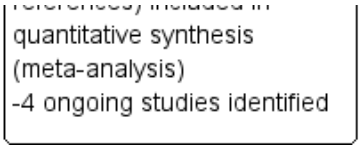


Figure 1. (Continued)



Data extraction and management

We performed data extraction using a standard data extraction form that included the following:

Characteristics of the study

- Study sponsors and author affiliations
- Study funding
- Methods, including study design, method of sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcomes, participant attrition and exclusions, and selective outcome reporting
- Full-text availability versus abstract only

Characteristics of the study population

- Country of enrolment
- Inclusion/exclusion criteria
- Study definition of AIMSS
- Number of participants in each treatment arm
- Mean and range of participant ages
- Type of AI prescribed to the participant

Characteristics of the intervention

- Description of the intervention
  - Aerobic/resistance/combination/other
  - Exercise intensity: mild/moderate/vigorous
  - Frequency and duration of sessions
  - Duration of intervention period
  - Supervised versus home-based; group versus individual
- Details of control or waiting list group
- Compliance with intervention
- Safety

Characteristics of the outcomes

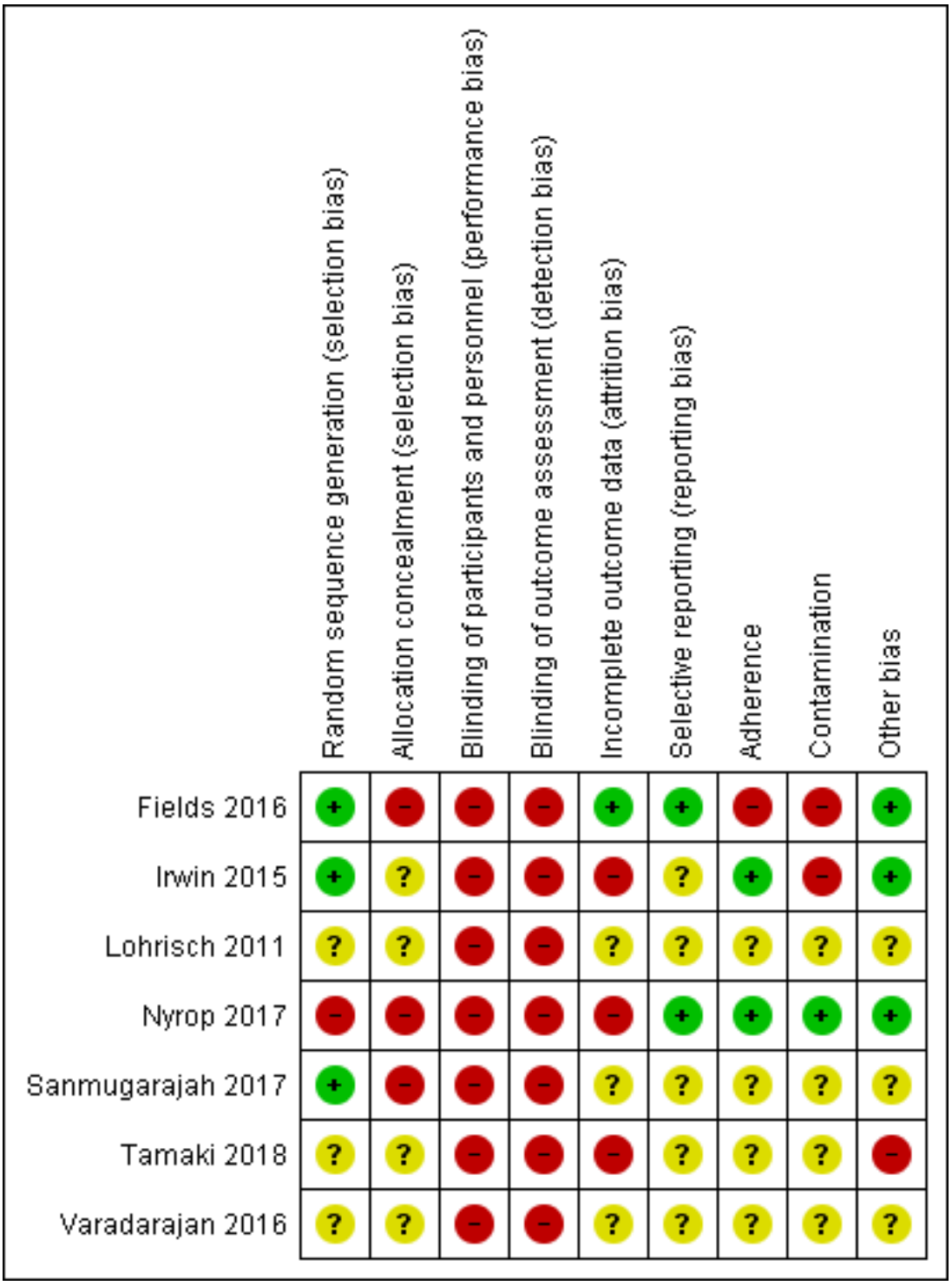
- Scoring systems used (and documentation of participant-reported outcomes versus investigator-reported outcomes)
- Outcomes of pain, stiffness, grip strength and health-related quality of life
- Change in incidence of AIMSS
- Timing of outcome data collection, including length of time between intervention and last collected outcome measurement
- Follow-up period

Two review authors (KER and SF) performed data extraction and a third review author (NW) resolved any disagreements. KER and SF entered data into Review Manager 5 (Review Manager 2014). Where there was more than one publication for a study, we extracted the data from all publications as required, but we considered the most recent version of the study to be the primary reference. We combined records relating to the same study under an overall study name or ID.

Assessment of risk of bias in included studies

We performed assessment of risk of bias for all RCTs using Cochrane's 'Risk of bias' assessment tool (Higgins 2017). This included the assessment of seven specific domains: random sequence generation; allocation concealment; blinding of participants and personnel; blinding of outcome assessment; incomplete outcome data; selective reporting; and other sources of bias. We assessed each study domain as high risk, low risk or unclear risk. Two review authors (KER, SF, NW or KR) independently assessed each study for risk of bias and a third review author (KER or NW) resolved any disagreements. Where there was incomplete reporting of the conduct of a study, we attempted to contact the authors of the study to clarify any uncertainties. 'Risk of bias' tables for each study are presented in the Characteristics of included studies table and a summary table, listing the 'Risk of bias' judgement for all studies is presented Figure 2.

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



Measures of treatment effect

We expected that studies would use a variety of different tools to measure the outcomes of interest (pain, stiffness, grip strength and health-related quality of life) and would mainly be reporting continuous outcomes. Therefore, we calculated the treatment effect by undertaking a standardised mean difference (SMD) analysis (SMD = difference in mean outcomes/standard deviation of outcomes among participants; Deeks 2017), to combine data from different instruments measuring the same domain. When studies used the same participant-reported outcome tool for a single outcome, we combined the data for meta-analysis using mean difference (MD). If there was minimal heterogeneity between studies, we used a fixed-effect model. This is different from our protocol, as we originally had proposed to use only the random-effects model, due to expected heterogeneity amongst the varying

interventions and assessment tools. In our revised method, we still used a random-effects model, but have also reported the results of the fixed-effect model. Due to the small number of studies, and small number of participants in some studies, we also performed a random-effects meta-analysis using the Hartung, Knapp, Sidik and Jonkman (HKSJ) approach (IntHout 2014).

For studies which we could not obtain standard deviations (SD), we imputed the SD as per Higgins 2011a. Where the SD was not available in the published study, or from study authors, we used the following formula to determine the SD: SD = √n × (upper limit 95% CI – lower limit 95% CI)/(2 x T value calculated by the T distribution), where n is the sample size and CI is the confidence interval. We estimated appropriate T values for smaller sample sizes using the TINV function (TINV(1-0.95,n-1)) in Excel. We could then use the calculated SD to calculate the SMD. These calculations were guided by the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b), and results were confirmed with the calculator available on Review Manager 5 (Review Manager 2014).

We calculated the published confidence intervals in the HOPE study (Irwin 2015), on the difference in the means for each group, so we calculated a SD for the change in means, rather than the final value for each arm of the study. By using these calculations, our review ran the risk of giving greater weight to the studies that reported change-from-baseline SD, as the SD in these studies may have been more precise than studies only reporting final value SD, due to the smaller SD (Deeks 2017). Therefore, where possible, we performed a separate analysis on final values and change-from-baseline values, and compared the results. Where we used a combination of final value confidence intervals and change-from-baseline values in a meta-analysis, we highlighted it in the text for the result.

Fields 2016 reported interquartile range (IQR) and median values, rather than mean and standard deviations. This is often an indicator that the data are skewed, so should be incorporated into a meta-analysis with caution (Higgins 2011a). To calculate SD from IQR, we used the following formula: SD = (q3 - q1)/1.35, where q3 is quartile 3 and q1 is quartile 1 (Higgins 2011b).

For dichotomous outcomes, we measured the treatment outcome by the odds ratio (OR), in combination with a 95% confidence interval (CI).

Unit of analysis issues

There were no studies that may have created unit of analysis issues, such as cross-over trials or trials with multiple treatment arms.

Dealing with missing data

In the case of missing data, wherever possible, we sourced additional information through clinical trials registries or data repositories. Where the required data were still not available, we contacted original corresponding authors via email and gave them three weeks to reply to the request. If the corresponding authors did not reply, we attempted further contact with the original investigators, and either the first or last author of each paper (if not the primary corresponding author). Where we were unable to obtain missing data, we have included an explanation for this in the Discussion section of our review.

Assessment of heterogeneity

We assessed the percentage of total variation across studies that is due to heterogeneity rather than chance using I^2 statistic (Higgins 2003). We also used the Chi^2 test and visual inspection of forest plots, as recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2017). Based on this, an I^2 statistic value of:

- 30% to 60% may represent moderate heterogeneity;
- 50% to 90% may represent substantial heterogeneity; and
- 75% to 100% represents considerable heterogeneity (Deeks 2017).

The importance of the I^2 statistic result depends on the magnitude and direction of effects, and the strength of evidence for heterogeneity. Based on Deeks 2017, a Chi^2 test greater than the degrees of freedom (df) and a small P value (e.g. $P < 0.05$) indicates significant heterogeneity, and we applied this guideline in the current analysis.

Assessment of reporting biases

We included one funnel plot in the assessment of reporting biases for the outcome with the largest number of studies. We could not undertake any further assessments due to the small number of studies contributing data to each outcome.

Data synthesis

We performed statistical analysis using Review Manager 5 software (Review Manager 2014). Where there was only low statistical heterogeneity, we performed a fixed-effect meta-analysis. Where there was at least moderate statistical heterogeneity present, we used a random-effects meta-analysis, using the inverse variance method to combine the data.

We reported the meta-analysis mainly by forest plots and the 'Summary of findings' table (see Summary of findings for the main comparison). For outcomes where there was an insufficient number of studies for us to pool for meta-analysis (i.e. fewer than two studies), or we could not combine the data, we presented our findings in a narrative manner.

Summary of findings and assessment of the certainty of the evidence

We developed a 'Summary of findings' table to assess the certainty of evidence using the GRADE approach, as detailed in the *Cochrane Handbook for Systematic Reviews of Interventions*, Chapter 11 (Schünemann 2017). The GRADE approach assesses the evidence using five considerations: study limitations, consistency of effect, imprecision, indirectness and publication bias. The key outcomes we included in the Summary of findings for the main comparison were:

- overall change in worst pain scores;
- overall change in stiffness scores;
- overall change in grip strength;
- overall change in health-related quality-of-life scores;
- overall change in cancer-specific quality of life;
- adverse effects, secondary to the intervention; and

- persistence and compliance of participants continuing to take their AI medication due to the intervention.

The 'Summary of findings' table in our review was different to the 'Summary of findings' table that we proposed in our protocol. We had initially intended to assess the overall change in the incidence of AIMSS. There were no studies that addressed the incidence of AIMSS as a result of exercise, and therefore we substituted the overall change in the incidence of AIMSS for one of our secondary outcomes, the overall change in health-related quality of life. Quality of life was further defined by incorporating both 'health-related quality of life' and 'cancer-specific quality of life'. We developed the 'Summary of findings' table using GRADEpro GDT software (GRADEpro GDT). Two review authors (KER, NW) independently assessed the evidence using the GRADE approach and a third review author (KR) resolved any disputes.

Subgroup analysis and investigation of heterogeneity

We did not undertake any subgroup analyses, as there were insufficient studies and participants to undertake any meaningful subgroup analysis within this review.

Sensitivity analysis

There were not enough studies to in our review to undertake meaningful sensitivity analyses.

RESULTS

Description of studies

Results of the search

The searches of the identified databases retrieved 1187 results. Our searches of other resources, such as bibliography and citation searching, and searching of the grey literature identified 19 additional records that appeared to meet the inclusion criteria. Once duplicates had been removed, there were 910 records for title and abstract screening, where we excluded 860 records. Where possible, we obtained the full text of the remaining 50 papers or register records. We excluded 32 articles (see Characteristics of excluded studies). We included seven studies (relating to 14 references) and identified four ongoing studies relevant to our inclusion criteria (see Characteristics of ongoing studies). The process is detailed further in the study flow diagram (see Figure 1; Moher 2009).

Included studies

The final selection of studies, based on review author consensus, resulted in seven studies for inclusion. Study characteristics and outcomes can be viewed in the Characteristics of included studies table. Three of these studies had been published as full texts (Fields 2016; Irwin 2015; Nyrop 2017), and four studies (Lohrisch 2011; Sanmugarajah 2017; Tamaki 2018; Varadarajan 2016), were published in abstract or poster form only. We were able to retrieve additional data from the Sanmugarajah 2017 trial via study author correspondence.

Three studies enrolled participants in the USA (Irwin 2015; Nyrop 2017; Varadarajan 2016), one in the UK (Fields 2016), one in Canada (Lohrisch 2011), one in Japan (Tamaki 2018) and one in Australia (Sanmugarajah 2017). All studies were RCTs, but four of these studies were designed as feasibility studies (Fields 2016; Nyrop 2017; Sanmugarajah 2017; Varadarajan 2016). The majority of the studies were investigating the treatment of AIMSS, with only one study investigating the prevention of AIMSS, which enrolled participants at the time of AI initiation (Sanmugarajah 2017).

Population

There were 400 participants enrolled across the seven studies. The sample sizes of the included studies ranged from 22 to 108. Four studies reported participant mean ages (Fields 2016; Irwin 2015; Lohrisch 2011; Nyrop 2017), and ranged from 61 to 63.8 years. Two studies gave age ranges (Fields 2016; Tamaki 2018), and ranged from 50 to 73 years. The majority of participants were on an AI at the time of enrolment, which included either anastrozole, letrozole or exemestane, with three studies reporting the average length of time on an AI (Irwin 2015; Nyrop 2017; Tamaki 2018), and ranging from 1.7 to 2.1 years. In Sanmugarajah 2017, participants commenced the exercise intervention within 12 weeks of being initiated on an AI. For detailed information on inclusion and exclusion criteria for each study, see the Characteristics of included studies table.

For the studies that only included participants experiencing AIMSS at baseline, the definition of AIMSS varied widely. Some studies reported their inclusion criteria as incorporating women experiencing any joint symptoms whilst taking an AI (Fields 2016; Irwin 2015; Nyrop 2017; Tamaki 2018; Varadarajan 2016), and only a few of these had stipulated a minimum pain score to qualify for inclusion (Irwin 2015; Nyrop 2017). Only one study specified arthralgia/myalgias, which were related to the AI as an inclusion criteria, although they did not report their criteria for this (Lohrisch 2011). Only one study reported the exclusion of women with pre-morbid musculoskeletal conditions such as rheumatoid arthritis (Varadarajan 2016). All of the studies had excluded metastatic disease as per their inclusion and exclusion criteria, but one study reported 16% of their participants (n = 10) as having stage IV disease in the baseline demographics (Nyrop 2017).

Interventions

The included studies investigated a variety of different exercise programmes. Two studies investigated walking programmes, with one of these being Nordic walking, which utilises walking plus the addition of hand-held poles (Fields 2016). The other walking study was a home-based exercise programme of 150 minutes' walking per week (Nyrop 2017). Three studies used a combination of resistance training plus aerobic exercise (Irwin 2015; Lohrisch 2011; Sanmugarajah 2017). One study only described their intervention as an "exercise program", without further details available (Varadarajan 2016). Tamaki 2018 enabled participants who were randomised to the exercise arm to choose between three types of exercises, which included either walking/running, gentle callisthenics, or weak exercise such as going up the stairs.

The length of the intervention varied between studies, ranging from 6 weeks to 12 months. The intensity of the exercise intervention was variably reported, with only two of the studies reporting the desired level of exercise intensity. One study aimed for 60% to 70% of maximum heart rate, with no further details given (Sanmugarajah 2017). The other study aimed for 60% to 80% of maximum heart rate, based on VO_2 max testing (Irwin 2015).

The majority of studies included a mix of both supervised and home-based exercise; two studies had supervised components initially (Fields 2016; Lohrisch 2011), but the remainder of the

these studies was unsupervised; two studies had supervised strength training plus home-based aerobic exercise (Irwin 2015; Sanmugarajah 2017); one study was completely home-based (Nyrop 2017); one study used completely supervised exercise in the intervention arm (Varadarajan 2016); and one study was unclear (Tamaki 2018). The majority of studies included at least 150 to 200 minutes of exercise weekly.

Three studies reported adherence to the exercise intervention (Fields 2016; Irwin 2015; Nyrop 2017). Fields 2016 aimed for four independent Nordic walking sessions each week in weeks 7 to 12 of their intervention, but only 8% of participants were compliant with this. In contrast, 68% to 85% completed one to two independent sessions weekly. In Irwin 2015, there was 70% mean attendance at the twice-per-week strength-training sessions, and a mean of 119 minutes of aerobic exercise weekly. The recommended amount of aerobic exercise in this study was 150 minutes a week. Control arms also varied widely, including a waiting list control group (Nyrop 2017), unsupervised moderate physical activity (Varadarajan 2016), written information about exercise in cancer (Fields 2016; Sanmugarajah 2017), or no exercise instruction until the end of the study (Irwin 2015). One study only described the control arm as "usual care" (Tamaki 2018), and another did not describe the details of their control arm (Lohrisch 2011).

Excluded studies

The reasons for exclusions are summarised in Characteristics of excluded studies. We excluded the majority of studies because they either were not RCTs; they had the incorrect participant population (e.g. participants on tamoxifen, rather than aromatase inhibitors, or by including women who had metastatic disease); or they were looking at different outcomes, such as other health-related quality-of-life symptoms, including fatigue or hot flushes, rather than AIMSS.

Risk of bias in included studies

We have documented details for the risk of bias of the included studies in the 'Risk of bias' tables, listed in the Characteristics of included studies. We requested additional information from study authors where the risk of bias rating was unclear, and was provided by the following studies: Fields 2016; Irwin 2015; Sanmugarajah 2017. The 'Risk of bias' summary can be viewed in Figure 2.

Allocation

There were a number of studies that were at high risk of selection bias. We judged one study as high risk of selection bias, because during recruitment, three participants who were randomised to the home-based walking intervention were inadvertently assigned to the waiting list control, and three participants who were randomised to the waiting list control were inadvertently assigned to the exercise intervention (Nyrop 2017). It was unclear whether this was due to inadequate random sequence generation or inadequate allocation concealment, and we judged the study at high risk of both components of selection bias. We judged three studies to be at unclear risk of selection bias since these studies failed to report sufficient information to adequately assess their means of random sequence generation (Lohrisch 2011; Tamaki 2018; Varadarajan 2016). We judged the remaining three studies to have adequate random sequence generation (Fields 2016; Irwin 2015; Sanmugarajah 2017), and were therefore judged to be at low risk of selection bias caused by random sequence generation.

Three studies were at high risk of selection bias due to concerns with the allocation concealment in their studies, because allocation of the intervention was not concealed such that investigators and participants could not foresee assignment to the study groups. One study, as described above, reported randomisation errors, and although they did not report the actual cause of the error nor when this became apparent, it may have been because investigators were aware of the allocation (Nyrop 2017). The study was therefore judged to be at high risk of selection bias (allocation concealment). Another study was at high risk of selection bias (allocation concealment) as the study did not fully implement allocation concealment due to resource constraints (Fields 2016). One study did not implement allocation concealment (Sanmugarajah 2017), and was also judged as high risk. We rated the remainder of the studies as having an unclear risk of selection bias (allocation concealment) as they did not describe the method of allocation concealment in enough detail to adequately allow definitive judgement (Irwin 2015; Lohrisch 2011; Tamaki 2018; Varadarajan 2016).

Blinding

None of the included studies reported blinding of participants and personnel. It is not feasible to blind participants to an exercise intervention because of the nature of the intervention. We therefore assessed all included studies as being at high risk of performance bias. None of the studies had blinding of outcome assessment. The majority of outcomes were participant-reported outcomes, and it was not practical to blind participants to these outcomes in an exercise intervention. Two of our outcomes, overall survival and breast cancer-specific survival, would not be affected by blinding, but none of the studies in our review measured these outcomes. We assessed all the studies as being at high risk of detection bias.

Incomplete outcome data

Three of seven studies reported that they had analysed data according to the intention-to-treat (ITT) principle (Fields 2016; Irwin 2015; Nyrop 2017), but only one of these studies had completion of outcome assessments for all randomised participants to enable a judgement of low risk of attrition bias (Fields 2016). We assessed three studies to be at high risk of attrition bias (Irwin 2015; Nyrop 2017; Tamaki 2018), basing this judgement on disparities in dropout rates between intervention and control group (Irwin 2015; Nyrop 2017), or high dropout rates of greater than 20% (Tamaki 2018). We assessed three studies to be at unclear risk of attrition bias due to insufficient information available to make a judgement (Lohrisch 2011; Sanmugarajah 2017; Varadarajan 2016).

Selective reporting

We judged two studies as low risk of reporting bias, because either they reported all of their proposed outcomes (Nyrop 2017), or only minor outcomes included in the initial trial registration were not reported in the study and these outcomes were not of interest to our review (Fields 2016). We judged five studies as unclear risk of reporting bias. The reasons for judgement of unclear risk were: in one study, at least one relevant missing unreported outcome amongst a very high number of planned outcomes in the protocol (Irwin 2015); not enough information being provided on

outcomes and study protocol or registration not being available (Tamaki 2018; Varadarajan 2016); the protocol being available but not enough information on outcomes provided (Sanmugarajah 2017); or the abstract publication reporting different outcomes to those mentioned in the trial registration, and the trial only being reported in abstract publication (Lohrisch 2011).

Other potential sources of bias

Four studies were only reported as abstracts, and therefore were difficult to assess for other sources of bias due to inadequate information, and we rated three of these studies as unclear risk (Lohrisch 2011; Sanmugarajah 2017; Varadarajan 2016). We rated one study, which was also reported in abstract/poster form, as being at high risk of other potential sources of bias, for allowing participants in the intervention arm to choose between three different exercise interventions with a wide range of exercise intensities (Tamaki 2018). Three studies were at low risk of other sources of bias (Fields 2016; Irwin 2015; Nyrop 2017).

We have added two additional domains to be assessed across all studies: adherence and contamination. Studies reported different approaches for measuring adherence. Some studies did not provide this information. Adherence was the level of exercise achieved once the participant had agreed to undertake the intervention. In two studies, participants adhered to the exercise intervention adequately, and both studies were assessed as low risk of bias due to adherence (Irwin 2015; Nyrop 2017). In four studies, risk of bias due to adherence was unclear (Lohrisch 2011; Sanmugarajah 2017; Tamaki 2018; Varadarajan 2016). In the remaining study, adherence to the intervention was so low in the independent sessions during weeks 7 to 12 that we classified it as high risk of bias (Fields 2016). Two studies reported exercise in the non-exercising control groups (contamination; Fields 2016; Irwin 2015), and we assessed them as high risk of bias for contamination. Four studies did not report contamination and therefore we judged these studies as unclear risk of bias (Lohrisch 2011; Sanmugarajah 2017; Tamaki 2018; Varadarajan 2016). One study had a minimal increase in baseline activity levels in the control group, and therefore we judged it as low risk of bias for contamination (Nyrop 2017).

Effects of interventions

See: **Summary of findings for the main comparison** Exercise therapies compared to standard care for the management of aromatase inhibitor-induced musculoskeletal symptoms

Prevention of symptoms

Only one study investigated the use of exercise in the prevention of AIMSS (Sanmugarajah 2017). This study was stopped early due to lack of funding, after accruing only 20 of the 120 participants intended for the study. We obtained further results from the study and also the study protocol via author correspondence. Tamaki 2018 allowed the inclusion of women who were only just commencing their AI medication at the time of enrolment, but baseline characteristics showed that the majority of participants were already taking an AI prior to the study: AI administration 25.6 ± 13.8 months in the intervention arm, and 25.3 months ± 14.2 months in the control arm. Therefore we have included Tamaki 2018 in the analysis of treatment of symptoms section (outlined below).

Pain

Sanmugarajah 2017 used Brief Pain Index (BPI) scores to assess symptoms of pain. The study reported an increase of one BPI unit between baseline and 12-month follow-up, compared to an increase of mean BPI scores of five units in the control group. They did not provide any values. Correspondence with the study authors confirmed that differences in pain scores between groups were not statistically significant.

Stiffness

Sanmugarajah 2017 did not report stiffness as an outcome in the prevention of AIMSS.

Grip strength

Sanmugarajah 2017 reported a trend towards improved grip strength between baseline and six months in the exercise group. Study author correspondence confirmed that the change in grip strength was not statistically significant between groups and they did not provide values.

Safety of the intervention

The study did not report this outcome.

Incidence of AIMSS

The study did not report this outcome.

Persistence and compliance of women continuing to take their AI medication due to the intervention

Sanmugarajah 2017 collected data on AI adherence relating to preventing AIMSS, but this study has not been published in full, and adherence data were not available. We made contact with Sanmugarajah 2017, who did not provide data for AI adherence, but confirmed that the difference between groups was not statistically significant.

Health-related quality of life

Overall change in health-related quality of life

The study did not report this outcome.

Overall change in cancer-specific quality of life

The study did not report this outcome.

Breast cancer-specific survival

The study did not report this outcome.

Overall survival

The study did not report this outcome.

Treatment of symptoms

Six studies investigated the treatment of AIMSS (Fields 2016; Irwin 2015; Lohrisch 2011; Nyrop 2017; Tamaki 2018; Varadarajan 2016). Two of these studies had specified a minimum pain criteria for eligibility into the study, including either a pain score of at least 3 on a 5-point scale of joint pain, stiffness or achiness in the past seven days (Nyrop 2017), or arthralgia for at least two months that was at least mild in severity (consisting of a score of at least 3 for worst pain on a BPI; Irwin 2015). One study included women who had described any joint symptoms in the previous 12 months via

an amended C-PET (Checklist for Patients on Hormone Therapy) in clinic (Fields 2016). One study (Lohrisch 2011), listed arthralgias/myalgias as part of their inclusion criteria, and another included women who had been experiencing joint discomfort/stiffness when attempting activities of daily living (Varadarajan 2016). One study reported "any arthralgia level" as one of their inclusion criteria (Tamaki 2018).

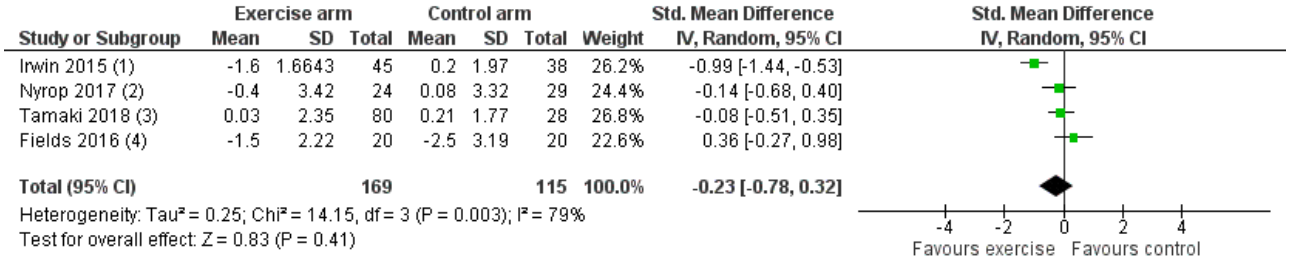
Pain

All of the six studies used participant-reported outcomes to assess pain symptoms. These included the Visual Analogue Scale (VAS; Nyrop 2017), Western Ontario and McMaster Universities Index (WOMAC; Irwin 2015; Nyrop 2017), Arthritis self-efficacy scale (Nyrop 2017), BPI-Short Form (BPI-SF; Fields 2016), BPI (Irwin 2015; Sanmugarajah 2017; Tamaki 2018), Pain Disability Index (Tamaki 2018), Pain Self-Efficacy Questionnaire (PSEQ; Fields 2016) and a Pain Scale (PS; Tamaki 2018).

We performed a meta-analysis on the effect of exercise on worst pain. Due to the availability of data, four studies were eligible for inclusion in the meta-analysis. Three of these studies reported BPI worst pain scores (Fields 2016; Irwin 2015; Tamaki 2018), and the remaining study reported WOMAC pain scores and VAS pain scores (Nyrop 2017). It should be noted that there was a discrepancy in the reporting of results between the two posters published for Tamaki 2018, with the initial poster (Tajaesu 2017), reporting a change of 0.09 points for worst pain at 12 months in the exercise group, and the final results poster reporting a change of 0.03 ± 2.35 points for worst pain in the exercise group. We used the result from the most recent poster/abstract in our analysis. The same study (Tamaki 2018), included three types of exercise in the intervention arm, and the participants randomised to the intervention arm were able to choose their exercise group. No details were reported on the number of participants in each exercise group, which ranged from weak exercise (going up the stairs) to strong exercise (120 to 150 minutes per week of walking or running).

Due to the different scoring systems used for measuring pain, we performed the analysis using SMD. In the meta-analysis, Nyrop 2017 used the WOMAC pain subscale and Irwin 2015 (the Hormones and Physical Exercise (HOPE) trial), used the BPI worst pain scores. The effect of exercise therapies on overall change in worst pain scores using the random-effects model resulted in an SMD of −0.23 (95% confidence interval (CI) −0.78 to 0.32; I^2 = 79%; 4 studies, 284 participants; very low-certainty evidence; Figure 3; Analysis 1.1). There was considerable statistical heterogeneity amongst the studies involved in the meta-analysis, which is likely to be explained by the wide range of outcome assessment tools used and also the range of exercise interventions utilised between the studies. The results using other models remained the same (fixed-effect model: SMD −0.29, 95% CI −0.54 to −0.04; HKSJ random-effects model: SMD −0.23, 95% CI −1.13 to 0.67). We performed a separate analysis using the results from the VAS scale in Nyrop 2017, which showed similar results (SMD −0.25, CI −0.80 to 0.30; I^2 = 79%; Analysis 1.2; fixed-effect model: SMD −0.31, 95% CI −0.56 to −0.06).

Figure 3. Forest plot of comparison 1. Primary outcomes, outcome 1.1: overall change in worst pain, using WOMAC pain subscale for Nyrop 2017



Footnotes
(1) Bried Pain Index (BPI) for worst pain: baseline compared with completion of 12 month intervention period. Standard deviation based on...
(2) Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale: baseline compared with completion of 6 week...
(3) Brief Pain Indec (BPI) for worst pain: baseline compared with completion of 12 month intervention period
(4) Brief Pain Index Short Form (BPI-SF) for worst pain: baseline compared with completion of 12 week intervention period

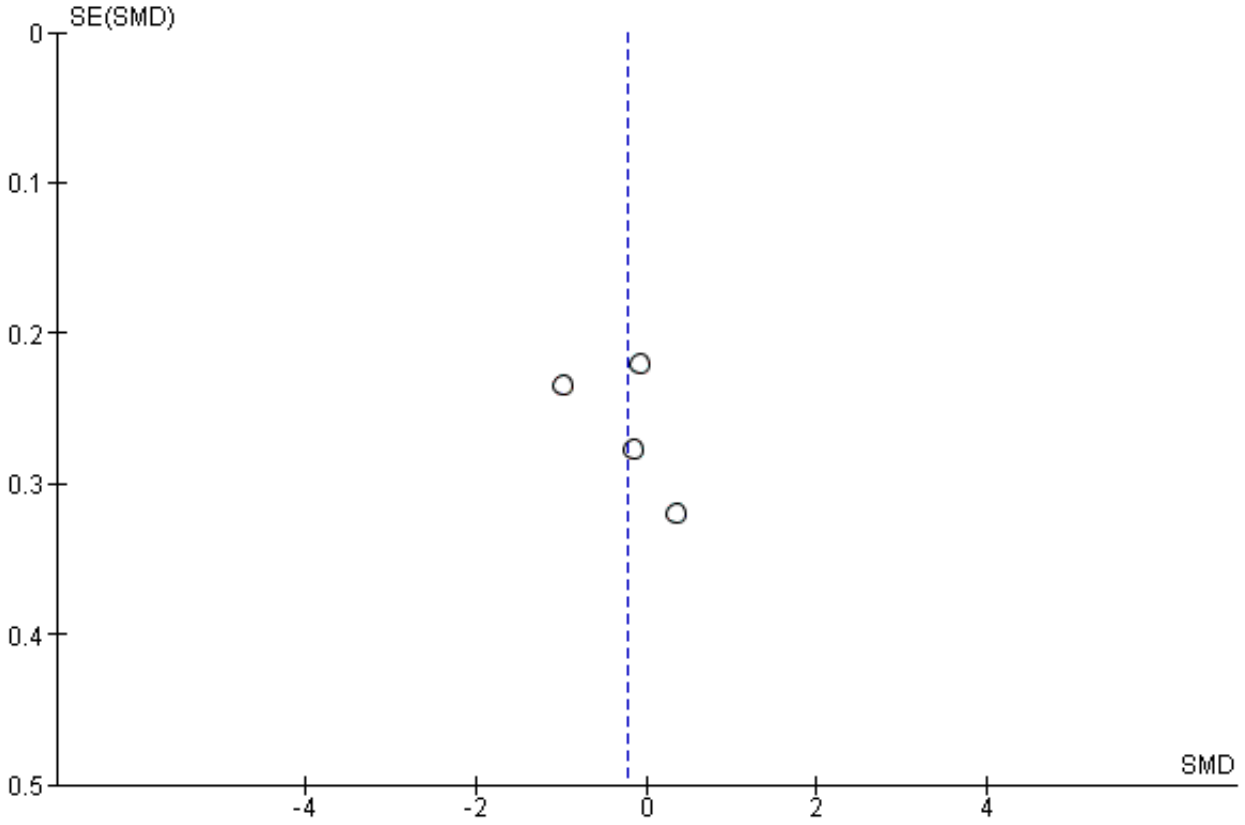
Irwin 2015 only reported BPI worst pain using change-from-baseline SD, rather than the SD of final values. Due to the potential risk of a change-from-baseline SD giving greater weight to the study, as discussed in our Measures of treatment effect, we also performed an analysis using SD from final values, obtained via study author correspondence. Of note, these data did not use a mixed-effect model with covariate adjustment as used in the published study results. The results of this analysis were also similar (SMD −0.18, CI −0.63 to 0.26; I^2 = 68%; Analysis 1.3).

Only limited results were available for the studies that we did not include in the meta-analysis. Lohrisch 2011 reported that the exercise intervention did not have a measurable improvement in AIMSS using the post-intervention, 12-week SF-36 pain scores, but did not report actual pain scores. Varadarajan 2016 reported that the intervention group showed a slight improvement in the pain scale, but did not report numerical values.

We were unable to determine the effect of exercise on worst pain scores, because we rated the evidence as very uncertain due to serious concerns with the risk of bias, such as lack of blinding, lack of allocation concealment, one study having inadequate random sequence allocation, and concerns about participation adherence to the exercise programme and contamination of the control group. The sample size in the meta-analysis was small, and therefore there were serious concerns regarding imprecision. There was also statistical heterogeneity between studies, and multiple studies that

did not publish their results in full. See the funnel plot in Figure 4 and Summary of findings for the main comparison.

Figure 4. Funnel plot of comparison 1. Primary outcomes, outcome 1.1: overall change in worst pain, using Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale for Nyrop 2017



Stiffness

Two studies investigated stiffness as an outcome (Irwin 2015; Nyrop 2017). Irwin 2015 used the WOMAC index, which incorporates three domains of pain, stiffness and physical function. However, the study authors did not report the stiffness subscale of the WOMAC index, and we were unable to obtain this result via study author correspondence. Nyrop 2017 reported stiffness using a VAS, and also the stiffness subscale of the WOMAC. For both scoring tools, higher scores indicate greater symptom severity. In the study by Nyrop 2017, involving 53 people, the WOMAC stiffness subscale reported a decrease in mean scores in the intervention arm (unsupervised walking programme) of −0.94 (95% CI −1.78 to −0.11) versus a decrease in mean scores of −0.18 (95% CI −0.94 to 0.57) in the waiting list control arm at the end of the six-week programme. Our own calculations showed the effect of exercise on stiffness as a mean difference (MD −0.76, 95% CI −1.67 to 0.15; 1 study, 53 people; low-certainty evidence; Analysis 1.4), using the WOMAC stiffness scale. The VAS stiffness scale reported a change in mean scores of −0.24 (95% CI −1.53 to 1.05) from baseline until the end of the six-week programme in the intervention arm, versus a change in mean scores of 0.18 (95% CI −1.02 to 1.38) in the control arm. Our own calculations showed the overall change in stiffness scores using the VAS tool as MD −0.42 (95% CI −2.10 to 1.26; 1 study, 53 people; low-certainty evidence; Analysis 1.4).

We rated the evidence for this outcome as low certainty due to concerns with the risk of bias, including inadequate random sequence generation and allocation concealment, plus lack of blinding and attrition bias. The small sample size, with only one study publishing results on this outcome has raised concerns about imprecision. See Characteristics of included studies and Summary of findings for the main comparison.

Grip strength

Two studies investigated grip strength (Irwin 2015; Varadarajan 2016). Irwin 2015 reported no statistical difference in grip strength between the intervention and control groups at the end of the 12-month intervention period, with mean change from baseline 0.1 (95% CI −0.5 to 0.7) and 0.4 (95% CI −0.2 to 0.9) respectively, P = 0.47. Our calculations showed a MD of 0.30 (95% CI −0.55 to 1.15; 1 study, 83 people; Analysis 1.5). Varadarajan 2016 reported an improvement in both left and right grip strength in the intervention group as compared to the control group, but did not report any numerical values.

We rated the certainty of the evidence for this outcome as very low, due to concerns with the risk of bias, including risk of performance bias, detection bias, and contamination in the control arm. One study had not been published in full, and results were not available, so we were unable to incorporate the results for this study in this

analysis. The sample size for this analysis was small, and therefore we downgraded the evidence further for imprecision. See Summary of findings for the main comparison.

Safety of the intervention

Four studies involving 331 women addressed safety (Fields 2016; Irwin 2015; Nyrop 2017; Tamaki 2018). All four studies examined exercise for treating AIMSS. Three studies (Irwin 2015; Nyrop 2017; Tamaki 2018), reported no adverse events related to the intervention. Fields 2016 reported two participants dropping out after the first six weeks of supervised exercise in the intervention arm due to longstanding musculoskeletal issues, which were felt to be unrelated to the study intervention. Fields 2016 also reported new pain in two participants, but one of these had newly identified metastases. There was no new lymphoedema in any of the participants in the same study (Fields 2016). The other three studies (Lohrisch 2011; Sanmugarajah 2017; Varadarajan 2016), did not report safety of the intervention. We rated the certainty of evidence for this outcome as low, due to concerns with the small sample size of this analysis, and the risk of bias for each study. In particular, none of the studies were blinded to participants or personnel, which we believe may have led to risk of bias for this outcome. See Summary of findings for the main comparison.

Incidence of AIMSS

None of the studies reported this outcome.

Persistence and compliance of women continuing to take their AI medication due to the intervention

Two studies assessing exercise for treating AIMSS reported on AI adherence (Irwin 2015; Tamaki 2018). For the two studies that assessed AI adherence secondary to exercise, the random-effects analysis of the two studies resulted in an odds ratio (OR) of 2.43 (95% CI 0.41 to 14.63; I^2 = 55%; 2 studies, 224 women; very low-certainty evidence; Analysis 2.1). The event rate of discontinuation was 10 participants in the exercise arm and 12 participants in the control arm. The OR in the fixed-effect model was 1.78 (95% CI 0.71 to 4.45); and the OR in the HKSJ random-effects model was 2.43 (95% CI 0 to 271558). Using the HKSJ random-effects model, the upper confidence interval changed dramatically from 14.63 to 271558. This is due to the fact that we have only two studies for this outcome. The adjustment is done as a function of the exponentiated T value, which in this case is 12.7 and thus it led to a huge change. We note as well that the effects estimate from the two studies is very different (1.24 and 8.42). The interpretation in both cases remains the same: the upper confidence interval is high and there is a big difference in the effect estimated by the two studies. We graded the certainty of evidence for this outcome as very low, due to serious concerns with imprecision as a result of the small sample size and event rate. There were also concerns regarding the risk of bias in each study, and we downgraded the evidence further as multiple studies have not been published in full, leading to a suspicion of publication bias. See Summary of findings for the main comparison.

Health-related quality of life

Overall change in health-related quality of life

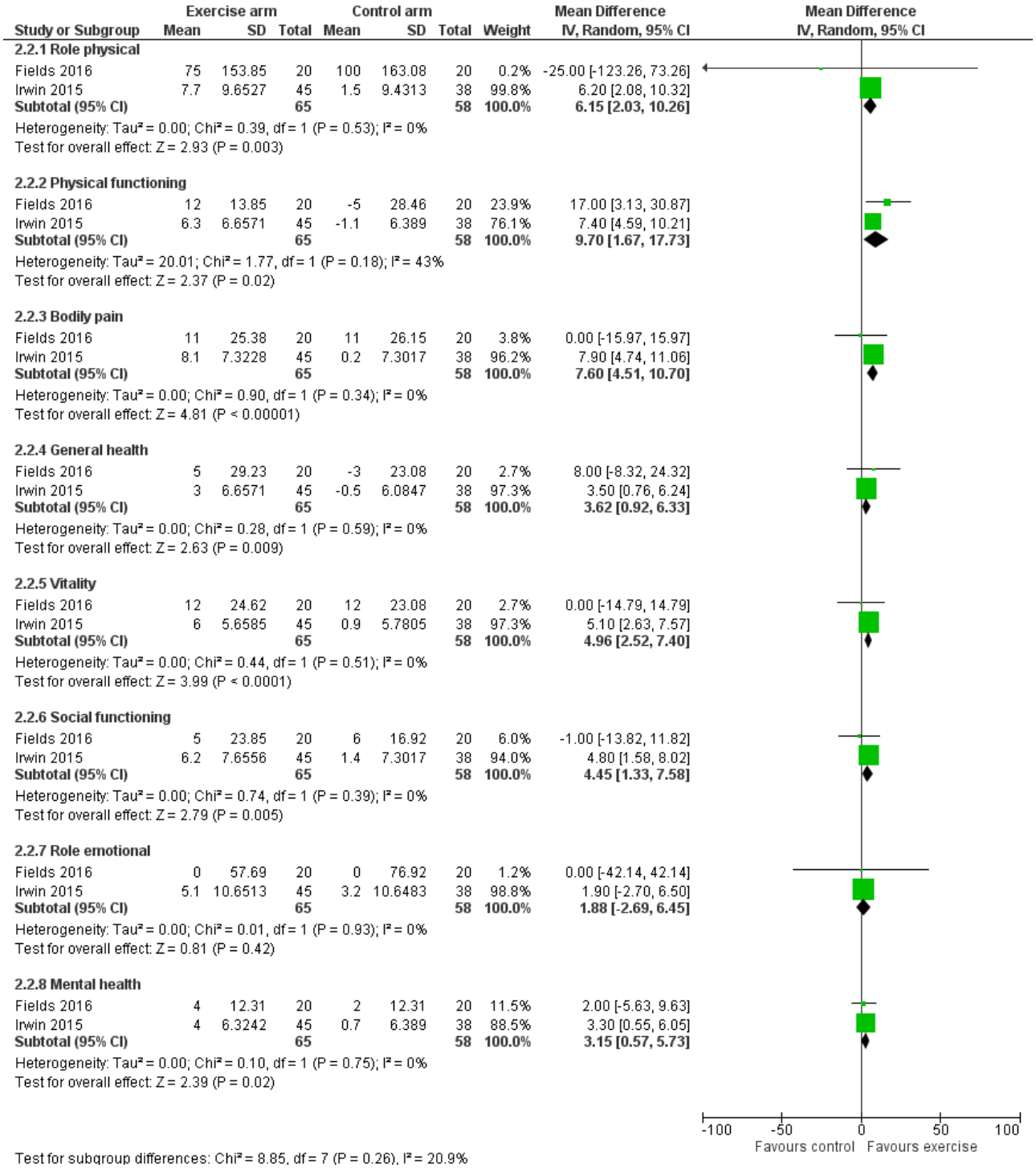
Two studies (Fields 2016; Irwin 2015), reported general health-related quality of life, in the form of the SF-36 (Rand Health Care). An additional study (Lohrisch 2011), collected data from SF-36, but did not report them. Irwin 2015 published quality-of-life data in a separate publication (Baglia 2019). There is not a total score for the SF-36 tool, instead, the subscales can be grouped into a Physical Component Score and a Mental Health Component Score. One study did not give the Physical Component score (Fields 2016). Therefore, we analysed the eight subscales within the SF-36 separately, as these data were available. Using the SF-36, a higher score indicated better health status.

The results using a random-effects model from the eight subscales included:

- role physical (MD 6.15, 95% CI 2.03, 10.26);
- physical functioning (MD 9.70, 95% CI 1.67 to 17.73; fixed-effect model MD 7.78, 95% CI 5.02 to 10.54; HKSJ random-effects model MD 9.7, 95% CI −42.32 to 61.72);
- bodily pain (MD 7.60, 95% CI 4.51 to 10.70);
- general health (MD 3.62, 95% CI 0.92 to 6.33)
- vitality (MD 4.96, 95% CI 2.52 to 7.40);
- social functioning (MD 4.45, 95% CI 1.33 to 7.58);
- role emotional (MD 1.88, 95% CI −2.69 to 6.45);
- and mental health (MD 3.15, 95% 0.57 to 5.73).

All subscale analyses included two studies involving 123 women and very low-certainty evidence (see Figure 5). Our analysis does not include the single-item Change in Health questionnaire, as this was not available from one study (Irwin 2015 (see quality-of-life data in Baglia 2019)). Fields 2016 used median values and interquartile ranges for reporting data, and stated that their reasoning for doing this was skewed data. Irwin 2015, as reported in Baglia 2019, provided mean values and change score confidence intervals, which we combined with the median scores and final value SDs calculated from Fields 2016. Ideally, the combination of change score and final value confidence intervals should not be done, as discussed in the Measures of treatment effect section. The greatest improvements were seen in the physical component scores, with the effect of exercise resulting in improvements in physical functioning and bodily pain, and which may correspond with clinically significant benefits (minimal clinically important improvement (MCII) of 7.1 and 4.9, respectively; Ward 2014). We graded the overall certainty of this evidence as very low, due to concerns with risk of bias for both studies (including concerns with performance bias, detection bias, poor adherence to the intervention in one study, and contamination of the control arm in both studies), a suspicion of publication bias as at least one study investigating this outcome has not been published in full, and also a considerable degree of imprecision with wide confidence intervals and a small sample size for the analysis. See Summary of findings for the main comparison.

Figure 5. Forest plot of comparison 2. Secondary outcomes, outcome 2.2: health-related quality of life



We did not assess other aspects of general quality of life, such as fatigue and depression, as part of this review. A number of studies investigated other aspects of quality of life using various participant-reported outcomes, such as the Centre for Epidemiological Studies Depression Scale (Irwin 2015), the Functional Assessment of Chronic Illness Therapy - Fatigue (FACIT-F; Irwin 2015), and the Patient Health Questionnaire-4 (PHQ-4; Varadarajan 2016).

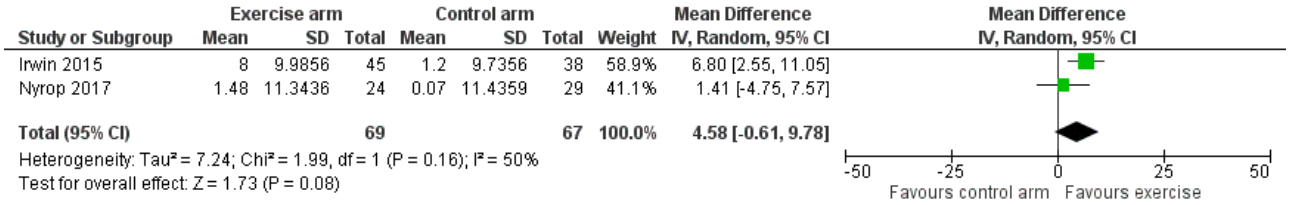
Overall change in cancer-specific quality of life

We added overall change in cancer-specific quality of life as an additional outcome, after our protocol was published. Our rationale was to try and report health-related quality of life in a more useful way for the reader, rather than one overall global health-related quality of life, which we felt would have been an incomplete assessment. Two studies assessed cancer-specific quality of life with the Functional Assessment of Cancer Therapy - General (FACT-

G) score (Irwin 2015; Nyrop 2017). Irwin 2015 published its quality-of-life data in a separate publication (Baglia 2019). In the FACT-G assessment tool, higher scores indicated better quality of life.

Our meta-analysis included quality-of-life assessments performed at the end of the exercise intervention in each study, which was after six weeks in Nyrop 2017, and 12 months in Irwin 2015, comprising a total of 136 participants. The effect of exercise resulted in a MD of 4.58 (95% CI −0.61 to 9.78; 2 studies; 136 participants; very low-certainty evidence; Figure 6; fixed-effect model: MD 5.06, 95% CI 1.56 to 8.56; HKSJ random-effects model: MD 4.58, 95% CI −29.12 to 38.28). The minimal clinically important change (MCID) score for FACT-G is 5 to 6 points (Eton 2004). We graded the certainty of this evidence as very low, due to the small sample size used in this analysis, leading to serious concerns with imprecision, and also concerns with risk of bias for both studies (including one study with high selection bias, and both studies having lack of blinding for the intervention and outcome assessments) and a suspicion of publication bias due to our knowledge of multiple studies not being published in full. See Summary of findings for the main comparison.

Figure 6. Forest plot of comparison 2. Secondary outcomes, outcome 2.3: cancer-specific quality of life



Breast cancer-specific survival

None of the studies reported on breast cancer-specific survival.

Overall survival

None of the studies reported on overall survival.

DISCUSSION

Summary of main results

The primary aim of this Cochrane Review was to investigate exercise therapies for the prevention or treatment of AIMSS in early breast cancer. We included data from seven RCTs with a total of 400 enrolled participants: one study assessed exercise for preventing AIMSS while six studies examined treating AIMSS. The comparator arm of studies was either usual care/information or walking, or the comparator was not reported. There was not enough evidence to determine the effect of exercise on the prevention of AIMSS based on a single study. Overall, the certainty of evidence was very low for multiple outcomes from the studies assessing prevention and treatment of AIMSS, and therefore it is unclear whether exercise has a positive or negative effect on pain, hand strength (grip strength), the number of women continuing to take AI medication, or the quality of life of women on AI medications. The evidence suggests that exercise results in little to no change in stiffness in women suffering from AIMSS, although the certainty of the evidence for this outcome was also low. Importantly, exercise is probably safe, with no harms reported, although this was not reported in four of the seven studies, and the follow-up interval was short. There were no data available to assess the effect of exercise on survival in this specific setting of postmenopausal women with breast cancer on adjuvant AI with AIMSS. Limited evidence from four studies suggests that exercise therapy resulted in little to no increase in adverse events compared to the comparator arm. No serious adverse events were reported. However safety data should be interpreted with caution given the low-certainty evidence in this review. There was insufficient evidence to determine the impact of exercise on the incidence of AIMSS due to scarcity of data.

Overall completeness and applicability of evidence

The review included studies with considerable clinical and methodological heterogeneity. The exercise interventions varied considerably. There was variability in the type of exercise intervention and in the frequency and intensity of the exercise therapy (e.g. Nordic walking; resistance and aerobic training; home-based, self-directed walking; three grades of exercise chosen by participant preference including a “weak” exercise arm); in the supervision and incorporation of behavioural support (home-based versus supervised); the duration of the intervention; and the nature of the control arm (usual care; written information; walking). Similar variability in exercise interventions has been noted in other Cochrane Reviews of exercise, including 'Exercise training for advanced lung cancer' (Peddle-McIntyre 2019), and 'Exercise for women receiving adjuvant therapy for breast cancer' (Furmaniak 2016). Limited information about the details of the interventions, and particularly the comparator arms, were available from some of the studies. Due the limitations of the available data, it was not possible to make any definitive conclusions about whether all relevant exercise interventions have been investigated. Due to the paucity of studies available, this review was unable to undertake further subgroup analyses to determine the effect of variations in exercise intensity, the effect of the setting or supervision of the exercise (supervised versus home-based) or the effect of differing types of exercise (such as aerobic/resistance/combination) on AIMSS. There are undoubtedly other exercise therapies, or variations of these, that could be trialled for AIMSS, but we are unable to determine whether these would affect the body of evidence.

Interventions were conducted in a variety of different settings, such as home-based, outpatient clinic etc. For applicability of the exercise interventions, consideration of the differences in the standard of care of women with early breast cancer would need to be considered in different populations and varying healthcare environments or systems, as these context factors may influence the effect of exercise interventions (Hawe 2004b; Schünemann 2017). Similar to the comments in the Cochrane Review, 'Exercise for women receiving adjuvant treatment for

breast cancer' (Furmaniak 2016), the exercise interventions being investigated for the treatment of AIMSS can be considered a complex intervention (Hawe 2004a). The Medical Research Council document 'A Framework for the Development and Evaluation of Randomised Controlled Trials for Complex Interventions' argues that "the greater the difficulty in defining precisely what exactly are the 'active ingredients' of an intervention and how they relate to each other, the greater the likelihood that you are dealing with a complex intervention" (Medical Research Council 2000). Hawe 2004b propose that for a complex intervention, such as exercise for AIMSS, which requires behavioural change, "the function and process of the intervention should be standardised, not the components themselves" allowing the intervention to be tailored to the local conditions.

Women with early breast cancer included in this review did have racial diversity (predominantly white/Asian) however certain racial groups were not adequately represented. Lack of racial diversity was noted in the Cochrane Review of 'Exercise for women receiving adjuvant treatment for breast cancer' (Furmaniak 2016), who noted the preponderance of white women included in the reviewed trials. The median age of participants diagnosed in the USA with breast cancer is 62 years (SEER database, Howlader 2019). The median age range of participants in the included studies is similar to this, and representative of post-menopausal breast cancer in high-income countries. However there were no reported participants included in the reviewed trials who were older than 75 years of age, although reporting of age ranges was poor. In addition, improvements in disease-free survival outcomes were observed in pre-menopausal women in the SOFT and TEXT studies, which were two landmark trials that showed the benefit of aromatase inhibitor use in pre-menopausal women with breast cancer in the adjuvant setting, in combination with ovarian function suppression (Francis 2015; Francis 2018). In these two trials, 88.7% of women who received ovarian suppression plus exemestane reported musculoskeletal symptoms compared to 69.0% of the women who received tamoxifen alone. Few young post-menopausal women with AIMSS were included in the studies of our Cochrane Review, likely due to time frames when results from the SOFT and TEXT studies became available (Francis 2015; Francis 2018). Women are being increasingly treated with ovarian suppression with AI and into the future, will represent an increasing group of women with AIMSS. These women may potentially have different baseline symptom levels, or responses to exercise interventions that have not yet been assessed. Mao 2009 reported AIMSS appeared to be inversely correlated with time since menopause, raising the possibility that young women with abrupt oestrogen withdrawal may be most at risk of symptoms.

The included studies have recruited in highly resourced health economies, which may have resources allowing interventions such as supervised exercise programmes. Hence the studied exercise interventions may not be applicable in all health systems. Rates of adherence or acceptability may not be the same or even feasible in different settings or populations. Post-menopausal breast cancer interventions should preferably be broadly applicable to older women. Kilari 2016 recommend that when designing exercise clinical trials for older adults with cancer, the exercise interventions should ideally be cost effective and "not burdensome to the patient/payer/society".

Many debilitating symptoms that are characteristic of AIMSS have not been adequately investigated, such as joint stiffness, which only has findings reported from one study; and other outcomes have uncertain evidence. Problematically, there is inconsistent definition of AIMSS, lack of objective outcome measures, and multiple participant-reported outcomes in the trials to date (Hershman 2015; Niravath 2013), and this lack of consensus limits the interpretation of the degree of completeness of the outcome measures in our review. Multiple outcomes were assessed and it was difficult to combine several of the outcome measures in meta-analysis. Data reporting was of low quality for some of the included studies, and some important outcome measures were missing and we were unable to obtain them. Similarly, there was considerable heterogeneity in the timing of the outcome measures (6 weeks to 12 months) due to the variation in exercise protocols, and in the length of follow-up. Differences in the timing of outcome measures between studies may limit comparability and determination of effect size. Duration of follow-up considerations may be important in determining longer-term benefits or harms of the exercise intervention.

Very few studies investigated the effect of exercise on quality of life in women with AIMSS. Due to the wide range of symptomatology of AIMSS, and the potential severity of symptoms, AIMSS can affect multiple facets of health and well-being for women. Safety data were not available for many of the studies, although reported adverse event rates do appear minor and of low incidence in the remaining studies. As only limited safety data are available, a degree of caution needs to be observed. Adherence rates to exercise were reported in two studies, and this is important also in investigating the tolerability of a particular exercise intervention, and also potential differences between the interventions. Adherence rates to aromatase inhibitors were only reported in two studies. Continuing adherence to AI treatment would seem to be an important outcome for improvement due to an exercise intervention, so it does appear the evidence body is not complete. Many of the above factors do limit conclusions about the benefits or harms of exercise in women with AIMSS. In addition, as studies included were generally of small size and at risk of bias, caution should be advised in interpretation of this review.

Quality of the evidence

There were only a small number of studies available, and only a small number of participants in most of the studies. The number of participants in the intervention arm ranged from 11 to 80 across all studies. The overall number of participants in each of our outcome assessments was low, which led to downgrading the certainty of evidence for all outcomes due to serious concerns about imprecision. Furthermore, a number of outcomes produced results with wide confidence intervals, where the 95% confidence interval included both no effect and appreciable benefit or harm. One study (Fields 2016), included in the analysis, reported results with medians and interquartile ranges rather than means and confidence intervals/standard deviations, due to skewed data. The quality-of-life data, in particular, reported very imprecise results with wide confidence intervals. All outcome assessments using results from this study should be viewed with caution, due to the skewed data in their study.

Due to the nature of the intervention, blinding of participants and personnel was not possible. The extent to which absence of blinding may have affected the results of each outcome is

unclear, but this contributes a high risk of performance bias for each of the included studies. In addition, there was a high risk of detection bias for each of the studies, as the majority of the outcomes were self-reported, and the remaining outcomes were not blinded to outcome assessors. Multiple studies were at risk of selection bias, with one study reporting mistakes in their random sequence generation, and multiple studies had inadequate allocation concealment. A number of studies had inadequate outcome data and selective reporting of outcomes. We judged that particularly the problems with allocation concealment and random sequence generation could potentially impact the results of these studies to a considerable degree, and therefore we downgraded by at least one point the grading of evidence for all of our outcomes due to serious concerns with the risk of bias of all studies. In addition as studies included were generally of small size and at risk of bias, caution should be advised in interpretation of this review.

Because of the heterogeneous nature of the various exercise regimens amongst the studies, we downgraded the certainty of evidence for the worst pain outcome by one point due to moderate statistical heterogeneity. For the remaining outcomes where there was only mild to moderate statistical heterogeneity, we did not downgrade the certainty of the evidence.

Four of the seven studies had only been reported in abstract form, and not published in full. Three of these studies reported non-significant differences in pain between the intervention and control groups, and therefore it is likely that they were not published in full because they were deemed to be negative studies. It is therefore possible that some of our outcomes may have over-emphasised the benefit of exercise, due to publication bias amongst studies on this topic. A number of other outcomes of interest to this review were assessed by the studies that had not been published in full, and we were unable to obtain the results. The publication of these results may have improved the certainty of evidence in our review across multiple outcomes.

Potential biases in the review process

A strength of this review is the extensive search methods strategy employed, with no language limitation, which we expected would have identified the main studies. However, it should be acknowledged that the searches identified only English language studies, raising the possibility that we missed studies in other languages, with possible publication bias for studies included in our review. Inconsistency in the definitions of AIMSS or exercise therapy may have potentially led to bias in the search strategy for studies. We attempted to account for this by making the search strategy broad. We designed a search strategy with terms for all the generally accepted exercise therapies that have been investigated in the literature, and for all aromatase inhibitors in clinical practice. Two review authors independently reviewed each of the searched studies to assess the risk of bias.

We attempted to contact all study authors for further information. Three study authors provided additional information (Fields 2016; Irwin 2015; Sanmugarajah 2017). We were either unable to contact four study authors or they could not provide additional information at the time of our data collection (Lohrisch 2011; Nyrop 2017; Tamaki 2018; Varadarajan 2016). A language barrier may have played a role in our inability to make contact with the authors of the Japanese study (Tamaki 2018). Three authors providing additional data increased the available data for analysis. A limitation of our analysis of exercise for prevention or treatment of AIMSS are unavailable data, and this introduces selection bias into our review; we had to exclude certain outcomes from the meta-analysis due to our inability to obtain further requested data.

Due to resource constraints, we did not systematically evaluate the strengths and weaknesses of the measurement instruments. However as per comments in the above section, there are enormous discrepancies in subjective and objective outcome measures used to assess AIMSS (Hershman 2015; Niravath 2013), and this is hence a limitation of our review, as outcome results will need to be interpreted in the context of heterogeneity of measurement of outcomes.

Agreements and disagreements with other studies or reviews

This is the first Cochrane Review of exercise for prevention or management of AIMSS. Other review evidence (systematic reviews and meta-analyses) were systematically reviewed. Our search strategy was broad to catch all potentially relevant papers. In agreement, although at an earlier search date than ours, a systematic review of systematic reviews (Kim 2018), identified the same three systematic reviews that included exercise (Nahm 2018; Roberts 2017; Yang 2017). However, in this Cochrane Review, we excluded systematic reviews of acupuncture alone for AIMSS.

Several of the authors involved in this review were also previously involved in a review entitled 'Management of aromatase inhibitor induced musculoskeletal symptoms in postmenopausal early breast cancer: a systematic review and meta-analysis' (Roberts 2017). This review considered all clinical trials, both prospective and retrospective, including RCTs, cohort and case-control studies and preventative trials of interventions for AIMSS; included all pharmacological, non-pharmacological (including exercise) and complementary and alternative medicine (CAM) interventions. Roberts 2017 considered only English language publications, identified three RCTs (Fields 2015; Irwin 2015; Lohrisch 2011), and grouped physical therapies narratively. They performed meta-analysis on two RCTs of exercise (Irwin 2015; Fields 2015), and the overall mean difference in the worst BPI using a random-effects model was −0.29 (95% CI −3.32 to 2.75), with significant between-study heterogeneity. Roberts 2017 is not directly comparable to this Cochrane Review, as due to paucity of data, they included non-randomised data and the scope of interventions was broader. They determined evidence quality to be poor overall, but it was not graded. Roberts 2017 did not undertake any reporting or analysis of stiffness, health-related quality of life or adherence outcomes. No adverse events were reported in agreement with this Cochrane Review.

Yang and colleagues conducted a review entitled 'Interventions for the treatment of aromatase inhibitor associated arthralgia in breast cancer survivors. A systematic review and meta-analysis' (Yang 2017). They included all studies that were RCTs and “quasi-experimental design”. The primary outcome was pain, described as a mean score, and assessed by BPI at the end time point of the intervention. They included a subgroup analysis of three studies of exercise in the meta-analysis: two were RCTs (Fields 2015; Irwin 2015), and one was a pre-test post-test study (DeNysschen 2014). Physical exercise was reported to show “no significant effect on pain, although they had a trend to decreasing joint pain” (SMD −0.562, 95% CI −1.499 to 0.375). This subgroup meta-analysis was

consistent with our review. However they undertook no reporting or analysis of stiffness or health-related quality-of-life outcomes, and did not report any adverse events, in contrast to this Cochrane Review.

Nahm and colleagues performed a systematic review entitled 'Efficacy of management strategies for aromatase inhibitor-induced arthralgia in breast cancer patients: a systematic review' (Nahm 2018). They identified one RCT of exercise (Fields 2015), and did not perform a meta-analysis. As this systematic review included all management interventions for AIMSS and included non-randomised studies, it is not directly comparable.

The most recent review, Kim 2018, was entitled 'Therapeutic options for aromatase inhibitor-associated arthralgia in breast cancer survivors: a systematic review of systematic reviews, evidence mapping, and network meta-analysis'. This was a systematic review of eligible systematic reviews, which were subjected to evidence mapping, and the RCTs included in the reviews were handsearched for network meta-analysis. The search strategy therefore differs, as does the statistical methodology, and the included studies to our Cochrane Review. Kim 2018's effectiveness outcome was mean and standard deviation of BPI. Two of the six RCTs included in the network meta-analysis were exercise studies (Fields 2015; Irwin 2015). In contrast to our review, Kim 2018 did not combine these but compared them separately to waiting list control, with the review stating, “aerobic exercise… significantly improved pain severity scores” for the study by Irwin 2015 (MD −0.80, 95% CI −1.33 to −0.016) in the abstract of the article. There was no combined meta-analysis of the data from Irwin 2015 with Fields 2015 of Nordic walking. The MD of BPI for Nordic walking versus control was −1.58 (95% CI −3.21 to 0.05). In agreement with our review, Kim 2018 concluded adverse event reporting to be poor, and as a result of this, they did not undertake network meta-analysis of adverse events.

AUTHORS' CONCLUSIONS

Implications for practice

Overall, this review highlighted that across multiple outcomes, there was very low-certainty evidence. This means that we were uncertain of the overall effect of exercise on pain, grip strength, health-related quality of life, cancer-specific quality of life or adherence to aromatase inhibitors in the treatment of aromatase inhibitor-induced musculoskeletal symptoms (AIMSS). There was insufficient evidence to assess the use of exercise in the prevention of AIMSS.

Despite these inconclusive findings, exercise still needs to be part of routine care for women with breast cancer. Current guidelines recommend exercise be embedded as part of routine care for women with breast cancer and in fact in all people with cancer (COSA 2018; European Society for Medical Oncology 2014; Rock 2012). Exercise has been shown to be beneficial in relieving cancer fatigue (Cramp 2012), and depressive symptoms (Patsou 2017), and improving cancer-related outcomes (Cormie 2017), and quality of life (Gerritsen 2016). Given that a supervised exercise programme has a low risk of harm, it is reasonable to incorporate aerobic and resistance exercise into the treatment of all people affected by cancer.

The updated guidelines for physical activity in cancer survivors, led by the Amercian College of Sports Medicine, recommends aerobic and resistance training for approximately 30 minutes, for three sessions per week (Campbell 2019). Despite longstanding guidelines recommending exercise in people with cancer, a review by Irwin (Irwin 2009), estimated that approximately 70% of cancer survivors were not performing the recommended exercise targets. Lack of adherence constitutes an important barrier to implementation of recommended exercise targets. Similarly, resource constraints may limit the implementation of certain types of exercise interventions studied in AIMSS, and their adoption into practice in certain healthcare systems. In keeping with the previously mentioned guidelines, exercise should continue to be recommended for women with breast cancer undergoing adjuvant AI treatment, although patient preference, lifestyle and health factors, and resource constraints may need to be considered for individual women.

Implications for research

Research in AIMSS has been hampered by a poor understanding of the aetiology of the syndrome (Niravath 2013), making it challenging to design and implement trials that address the problem. To sufficiently answer the question of whether exercise can prevent or treat AIMSS, further high-quality, adequately powered, randomised, phase three controlled trials of targeted exercise interventions on key outcomes would be required. With current guidelines recommending exercise be embedded as part of routine care for women with breast cancer (COSA 2018; European Society for Medical Oncology 2014; Rock 2012), any further exercise trials versus a waiting list control would need to incorporate guideline recommendations appropriate for the local population into the control arm. The control arm should be adequately described in any future or ongoing trials.

To improve precision of results, all studies, even if deemed 'negative trials', should endeavour to publish their results, to enable inclusion in future meta-analyses.

Research has been limited also by heterogeneous patient populations (Hershman 2015). Better understanding of the risk factors for AIMSS may help stratify women who are most at risk of AIMSS (Hershman 2015). Prior chemotherapy, prior hormone replacement therapy and increased body weight have been identified as increasing the risk of AIMSS (Hershman 2015). These women may also have the most to benefit from exercise intervention strategies, and research into subgroups who have a higher baseline incidence of AI arthralgia may potentially be a focus for future research.

Research in AIMSS has also been hampered by lack of objective outcome measures, the use of multiple different patient-reported outcomes (Hershman 2015), and by lack of consensus on the constellation of clinical symptoms and signs that constitute or define the syndrome (Hershman 2015; Niravath 2013). This variability was reflected in the included exercise studies in our review. To reduce clinical and methodological heterogeneity, consensus on definitions and outcome measures in AIMSS needs to be developed as a research priority. Such consensus would improve the ability to answer the relevant research questions. Better definition of AIMSS may lead to more rigorous entry criteria, as it is possible that some of the included studies included participants with pre-existing musculoskeletal or neuropathic

pain syndromes. Patient-reported outcomes should certainly be the primary focus of AIMSS research, as the most reliable measures of quality of life include patient-reported outcomes (Deshpande 2011). Efforts are being made to develop more reliable tools. A recent study developed and validated a reliable arthralgia severity measurement instrument, the Patient-Reported Arthralgia Inventory (PRAI), in postmenopausal breast cancer patients with AIMSS (Castel 2015). Other newly developed tools including electronic/web-based patient-reported outcomes may also improve research capability. Better standardisation may improve the quality of the evidence generated, and allow for relevant outcomes to be compared in future meta-analysis. Patient-reported outcome monitoring may potentially improve care delivery and patient outcomes (Nipp 2017), and incorporation of these newer tools should be a priority for further research. It is also important for future AIMSS trials to incorporate quality-of-life patient-reported outcomes in their outcome assessments, as we become increasingly aware of the impact of AIMSS on aspects of quality of life other than pain. The optimal timing of outcome measures remains to be established, as AIMSS fluctuates with time (Hershman 2015). Future trials should endeavour to include longer-term follow-up assessments to establish if any positive effects of exercise continue beyond the intervention period.

The next phase of ongoing research may examine comparison between different types, frequencies and intensities of exercise in AIMSS. Goal setting, setting of graded tasks and behavioural instructions are the common features of exercise interventions in people with cancer that meet the recommendation targets with high adherence rates (Turner 2018). These differences between the studies in AIMSS are yet to be examined. Similarly, patient preference between different exercise interventions trialled in AIMSS also remains unknown, and this should be examined. If the benefit of exercise was firmly established in other AIMSS outcomes, such research questions may be of greater merit. However, this research would need to be conducted in the complex environment of current recommended exercise guidelines for women with breast cancer that reports low compliance rates, and be appropriate for different populations and healthcare settings.

ACKNOWLEDGEMENTS

We would like to acknowledge Ava Grace Tan-Koay, Cochrane Breast Cancer's Information Specialist, for peer reviewing the search strategy and for searching Cochrane Breast Cancer's Specialised Register. We would also like to acknowledge our reviewers: Dr Rachel Dear, Medical Oncologist, St Vincent's Hospital, Darlinghurst; Rebecca Seago-Coyle, Breast Cancer Patient Advocate; and Gian Luca Di Tanna, Head of Statistics, The George Institute for Global Health.

REFERENCES

References to studies included in this review

Fields 2016 *{published data only}*

Fields J, Richardson A, Hopkinson J, Fenlon D. Nordic walking as an exercise intervention to reduce pain in women with aromatase inhibitor- associated arthralgia: a feasibility study. *Journal of Pain and Symptom Management* 2016;**52**(4):548-59.

Irwin 2015 *{published and unpublished data}*

Arem H, Sorkin M, Cartmel B, Fiellin M, Capozza S, Harrigan M, et al. Exercise adherence in a randomized trial of exercise on aromatase inhibitor arthralgias in breast cancer survivors: the Hormones and Physical Exercise (HOPE) study. *Journal of Cancer Survivorship* 2016;**10**(4):654-62.

Baglia ML, Lin IH, Cartmel B, Sanft T, Ligibel J, Hershman DL, et al. Endocrine-related quality of life in a randomized trial of exercise on aromatase inhibitor-induced arthralgias in breast cancer survivors. *Cancer* 2019;**125**(13):2262-71.

Crespo-Bosque M, Brown C, Cartmel B, Harrigan M, Irwin M, Neogi T. Pain and sensitization in women with aromatase inhibitor-associated arthralgias. *Arthritis and Rheumatology* 2016;**68**:2524-6.

Irwin M. Exercise in the management of arthralgia. *Asia-Pacific Journal of Clinical Oncology* 2016;**12**:69.

* Irwin M, Cartmel B, Gross C, Ercolano E, Li F, Yao X, et al. Randomized exercise trial of aromatase inhibitor–induced arthralgia in breast cancer survivors. *Journal of Clinical Oncology* 2015;**33**(10):1104-11.

Irwin ML, Cartmel B, Gross C, Ercolano E, Fiellin M, Capozza S, et al. Randomized trial of exercise vs. usual care on aromatase inhibitor-associated arthralgias in women with breast cancer: the Hormones and Physical Exercise (HOPE) study. *Cancer Research* 2013;**73**(24 Suppl):S3-03.

Printz C. Exercise helps ease joint pain for women with breast cancer. *Cancer* 2014;**120**(8):1133.

Thomas GA, Cartmel BR, Harrigan M, Fiellin M, Capozza S, Zhou Y, et al. The effect of exercise on body composition and bone mineral density in breast cancer survivors taking aromatase inhibitors. *Obesity* 2017;**25**(2):346-51.

Lohrisch 2011 *{published data only}*

Lohrisch CA, McKenzie D, Truong P, Jesperson D, Gelmon KA, Premji S, et al. A randomized trial of exercise versus control for musculoskeletal symptoms from adjuvant anastrozole (A) for postmenopausal early breast cancer (PEBC). *Journal of Clinical Oncology* 2011;**29**(15 Suppl):636.

Nyrop 2017 *{published data only (unpublished sought but not used)}*

Nyrop K, Callahan L, Cleveland R, Arbeeva L, Hackney B, Muss H. Randomized controlled trial of a home-based walking program to reduce moderate to severe aromatase inhibitor-associated arthralgia in breast cancer survivors. *Oncologist* 2017;**22**(10):1238-48.

Sanmugarajah 2017 *{published data only}*

Sanmugarajah J, Allan S, Bagchi R, Laakso EL. Can a supervised exercise program compared to usual care prevent aromatase inhibitor-induced musculoskeletal pain in women with breast cancer?. *Cancer Research* 2017;**77**(4 Suppl):P5-12-02.

Tamaki 2018 *{published data only (unpublished sought but not used)}*

Tamaki K, Takaesu M, Nagamine S, Terukina S, Kamada Y, Uehara K, et al. Final results of the randomized trial of exercise intervention vs. usual care for breast cancer patients with aromatase inhibitor to prevent and improve the aromatase inhibitor induced arthralgia. *Cancer Research* 2018;**78**(4 Suppl):P6-11-01.

Varadarajan 2016 *{published data only}*

Varadarajan R, Helm E, Arnold C, Huelsenbeck-Dill L, Ingraham-Lopresto B, Sonaad S, et al. Abstract P5-12-04: directed exercise intervention in breast cancer patients with arthralgias receiving aromatase inhibitors: a randomized pilot study. *Cancer Research* 2016;**76**(4 Supplement):P5-12-04-P5-12-04.

References to studies excluded from this review

Bower 2012 *{published data only}*

Bower JE, Garet D, Sternlieb B, Ganz PA, Irwin MR, Olmstead R, et al. Yoga for persistent fatigue in breast cancer survivors: a randomized controlled trial. *Cancer* 2012;**118**(15):3766-75.

Brown 2015 *{published data only}*

Brown JC, Schmitz KH. Weight lifting and physical function among survivors of breast cancer: a post hoc analysis of a randomized controlled trial. *Journal of Clinical Oncology* 2015;**33**(19):2184-9.

Cantarero-Villanueva 2013a *{published data only}*

Cantarero-Villanueva I, Fernandez-Lao C, Caro-Moran E, Morillas-Ruiz J, Galiano-Castillo N, Diaz-Rodriguez L, et al. Aquatic exercise in a chest-high pool for hormone therapy-induced arthralgia in breast cancer survivors: a pragmatic controlled trial. *Clinical Rehabilitation* 2013;**27**(2):123-32.

Cantarero-Villanueva 2013b *{published data only}*

Cantarero-Villanueva I, Fernandez-Lao C, Cuesta-Vargas A I, Del Moral-Avila R, Fernandez-de-Las-Penas C, Arroyo-Morales M. The effectiveness of a deep water aquatic exercise program in cancer-related fatigue in breast cancer survivors: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* 2013;**94**(2):221-30.

DeNysschen 2014 *{published data only}*

DeNysschen CA, Burton H, Ademuyiwa F, Levine E, Tetewsky S, O'Connor T. Exercise intervention in breast cancer patients with aromatase inhibitor-associated arthralgia: a pilot study. *European Journal of Cancer Care* 2014;**23**(4):493-501.

Desbiens 2017 *{published data only}*

Desbiens C, Filion M, Brien MC, Hogue JC, Laflamme C, Lemieux J. Impact of physical activity in group versus individual physical activity on fatigue in patients with breast cancer: a pilot study. *Breast* 2017;**35**:8-13.

Djuric 2011 *{published data only}*

Djuric Z, Ellsworth JS, Weldon AL, Ren J, Richardson CR, Resnicow K, et al. A diet and exercise intervention during chemotherapy for breast cancer. *Open Obesity Journal* 2011;**3**:87-97.

Galantino 2013 *{published data only}*

Galantino M, Cardena G, Piela N, Callens M, Mao JJ. Impact of tai chi on quality of life in breast cancer survivors with aromatase inhibitor associated arthralgia. *Rehabilitation Oncology* 2013;**31**(1):50-1.

Galiano-Castillo 2017 *{published data only}*

Galiano-Castillo N, Arroyo-Morales M, Lozano-Lozano M, Fernandez-Lao C, Martin-Martin L, Del-Moral-Avila R, et al. Effect of an internet-based telehealth system on functional capacity and cognition in breast cancer survivors: a secondary analysis of a randomized controlled trial. *Supportive Care in Cancer* 2017;**25**(11):3551-9.

Goodwin 2014 *{published data only}*

Goodwin PJ, Segal RJ, Vallis M, Ligibel JA, Pond GR, Robidoux A, et al. Randomized trial of a telephone-based weight loss intervention in postmenopausal women with breast cancer receiving letrozole: the LISA trial. *Journal of Clinical Oncology* 2014;**32**(21):2231-9.

Harrigan 2016 *{published data only}*

Harrigan M, Cartmel B, Loftfield E, Sanft T, Chagpar AB, Zhou Y, et al. Randomized trial comparing telephone versus in-person weight loss counseling on body composition and circulating biomarkers in women treated for breast cancer: the Lifestyle, Exercise, and Nutrition (LEAN) Study. *Journal of Clinical Oncology* 2016;**34**(7):669-76.

Kiecolt-Glaser 2014 *{published data only}*

Kiecolt-Glaser JK, Bennett JM, Andridge R, Peng J, Shapiro CL, Malarkey WB, et al. Yoga's impact on inflammation, mood, and fatigue in breast cancer survivors: a randomized controlled trial. *Journal of Clinical Oncology* 2014;**32**(10):1040-9.

Knobf 2017 *{published data only}*

Knobf MT, Jeon S, Smith B, Harris L, Thompson S, Stacy MR, et al. The Yale Fitness Intervention Trial in female cancer survivors: cardiovascular and physiological outcomes. *Heart and Lung: The Journal of Critical Care* 2017;**46**(5):375-81.

Lash 2011 *{published data only}*

Lash BW, Katz J, Gilman P. Feasibility of a defined exercise program for aromatase inhibitor–related arthralgia (AIRA) in breast cancer survivors. *Journal of Clinical Oncology* 2011;**29**(15 Suppl):e19725.

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.

Ligibel 2011 *{published data only}*

Ligibel JA, Segal R, Pond G, Dion MJ, Pritchard KI, Levine M, et al. Impact of the lifestyle intervention study in adjuvant treatment of early breast cancer (LISA) weight loss intervention upon physical activity. *Cancer Research* 2011;**71**(24 Suppl 3):P4-12-05.

Nikander 2012 *{published data only}*

Nikander R, Sievanen H, Ojala K, Kellokumpu-Lehtinen PL, Palva T, Blomqvist C, et al. Effect of exercise on bone structural traits, physical performance and body composition in breast cancer patients--a 12-month RCT. *Journal of Musculoskeletal & Neuronal Interactions* 2012;**12**(3):127-35.

Nyrop 2016 *{published data only}*

Nyrop KA, Callahan LF, Rini C, Altpeter M, Hackney B, DePue A, et al. Aromatase inhibitor associated arthralgia: the importance of oncology provider-patient communication about side effects and potential management through physical activity. *Supportive Care in Cancer* 2016;**24**(6):2643-50.

Pakiz 2016 *{published data only}*

Pakiz B, Ganz PA, Sedjo RL, Flatt SW, Demark-Wahnefried W, Liu J, et al. Correlates of quality of life in overweight or obese breast cancer survivors at enrollment into a weight loss trial. *Psycho-oncology* 2016;**25**(2):142-9.

Paulo 2019 *{published data only}*

Paulo TR, Rossi FE, Viezel J, Tosello GT, Seidinger SC, Simoes RR, et al. The impact of an exercise program on quality of life in older breast cancer survivors undergoing aromatase inhibitor therapy: a randomized controlled trial. *Health and Quality of Life Outcomes* 2019;**17**(1):17.

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 2009 *{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.

Peppone 2012 *{published data only}*

Peppone LJ, Mohile SG, Janelsins MC, Sprod L, Gewandter JS, Heckler CE, et al. YOCAS yoga for musculoskeletal symptoms in breast cancer patients receiving aromatase inhibitors: A URCC CCOP randomized, controlled clinical trial. *Journal of Clinical Oncology* 2012;**30**(15 Suppl):9028.

Peppone 2015 *{published data only}*

Peppone LJ, Janelsins MC, Kamen C, Mohile SG, Sprod LK, Gewandter JS, et al. The effect of YOCAS yoga for musculoskeletal symptoms among breast cancer survivors on hormonal therapy. *Breast Cancer Research and Treatment* 2015;**150**(3):597-604.

Pruthi 2012 *{published data only}*

Pruthi S, Stan DL, Jenkins SM, Huebner M, Borg BA, Thomley BS, et al. A randomized controlled pilot study assessing feasibility and impact of yoga practice on quality of life, mood, and perceived stress in women with newly diagnosed breast cancer. *Global Advances in Health and Medicine* 2012;**1**(5):30-5.

Reeves 2017 *{published data only}*

Reeves M, Winkler E, McCarthy N, Lawler S, Terranova C, Hayes S, et al. The Living Well after Breast Cancer pilot trial: a weight loss intervention for women following treatment for breast cancer. *Asia-Pacific Journal of Clinical Oncology* 2017;**13**(3):125-36.

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 & Prevention* 2009;**18**(5):1410-8.

Rogers 2009a *{published data only}*

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.

Segal 2011 *{published data only}*

Segal R, Pond GR, Vallis M, Dion M, Pritchard KI, Ligibel JA, et al. Randomized trial of a lifestyle intervention for women with early-stage breast cancer (BC) receiving adjuvant hormone therapy: initial results. *Journal of Clinical Oncology* 2011;**29**(15 Suppl 1):512.

Winkels 2017 *{published data only}*

Winkels RM, Sturgeon KM, Kallan MJ, Dean LT, Zhang Z, Evangelisti M, et al. The women in steady exercise research (WISER) survivor trial: the innovative transdisciplinary design of a randomized controlled trial of exercise and weight-loss interventions among breast cancer survivors with lymphedema. *Contempory Clinical Trials* 2017;**61**:63-72.

Winters-Stone 2012 *{published data only}*

Winters-Stone KM, Dobek J, Bennett JA, Nail LM, Leo MC, Schwartz A. The effect of resistance training on muscle strength and physical function in older, postmenopausal breast cancer survivors: a randomized controlled trial. *Journal of Cancer Survivorship* 2012;**6**(2):189-99.

References to ongoing studies

NCT03162133 *{published data only}*

NCT03162133. A randomized controlled trial evaluating the impact of Baduanjin exercise intervention on quality of life in breast cancer survivors receiving aromatase inhibitor therapy. clinicaltrials.gov/ct2/show/NCT03162133 (first received 22 May 2017).

NCT03786198 *{published data only}*

NCT03786198. Activity program during aromatase inhibitor therapy. clinicaltrials.gov/ct2/show/NCT03786198 (first received 24 December 2018).

NCT03953157 *{published data only}*

NCT03953157. Dietary and exercise interventions in reducing side effects in patients with stage I-IIIa breast cancer receiving aromatase inhibitors. clinicaltrials.gov/ct2/show/NCT03953157 (first received 16 May 2019).

NCT03956875 *{published data only}*

NCT03956875. Yoga for aromatase inhibitor-induced knee pain relief in breast cancer patients. clinicaltrials.gov/ct2/show/NCT03956875 (first received 21 May 2019).

Additional references

Aydiner 2013

Aydiner A. Meta-analysis of breast cancer outcome and toxicity in adjuvant trials of aromatase inhibitors in postmenopausal women. *Breast* 2013;**22**(2):121-9.

Baglia 2019

Baglia ML, Lin IH, Cartmel B, Sanft T, Ligibel J, Hershman DL, et al. Endocrine-related quality of life in a randomized trial of exercise on aromatase inhibitor-induced arthralgias in breast cancer survivors. *Cancer* 2019;**125**(13):2262-71.

Beckwee 2017

Beckwee D, Leysen L, Meuwis K, Adriaenssens N. Prevalence of aromatase inhibitor-induced arthralgia in breast cancer: a systematic review and meta-analysis. *Supportive Care in Cancer* 2017;**25**(5):1673-86.

Brier 2017

Brier MJ, Chambless DL, Gross R, Chen J, Mao JJ. Perceived barriers to treatment predict adherence to aromatase inhibitors among breast cancer survivors. *Cancer* 2017;**123**(1):169-76.

Burstein 2007

Burstein HJ. Aromatase inhibitor-associated arthralgia syndrome. *Breast* 2007;**16**(3):223-34.

Burstein 2019

Burstein HJ, Lacchetti C, Anderson H, Buchholz TA, Davidson NE, Gelmon KA, et al. Adjuvant endocrine therapy for women with hormone receptor-positive breast cancer: ASCO Clinical Practice Guideline Focused Update. *Journal of Clinical Oncology* 2019;**37**(5):423-38.

Campbell 2019

Campbell KL, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. *Medicine and Science in Sports and Exercise* 2019;**51**(11):2375-90.

Caspersen 1985

Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports* 1985;**100**(2):126-31.

Castel 2015

Castel LD, Wallston KA, Saville BR, Alvarez JR, Shields BD, Feurer ID, et al. Validity and reliability of the patient-reported arthralgia inventory: validation of a newly-developed survey instrument to measure arthralgia. *Patient Related Outcome Measures* 2015;**6**:205-14.

Chim 2013

Chim K, Xie SX, Stricker CT, Li QS, Gross R, Farrar JT, et al. Joint pain severity predicts premature discontinuation of aromatase inhibitors in breast cancer survivors. *BMC Cancer* 2013;**13**:401.

Cormie 2017

Cormie P, Zopf EM, Zhang X, Schmitz KH. The impact of exercise on cancer mortality, recurrence, and treatment-related adverse effects. *Epidemiologic Reviews* 2017;**39**(1):71-92.

COSA 2018

Clinical Oncology Society of Australia. COSA Position Statement on Exercise in Cancer Care. Sydney: Clinical Oncology Society of Australia, 2018.

Cramp 2012

Cramp F, Byron-Daniel J. Exercise for the management of cancer-related fatigue in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.CD006145.pub3]

Deeks 2017

Deeks JJ, Higgins JP, Altman DG (editors), on behalf of the Cochrane Statistical Methods Group. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Dennett 2016

Dennett AM, Peiris CL, Shields N, Prendergast LA, Taylor NF. Moderate-intensity exercise reduces fatigue and improves mobility in cancer survivors: a systematic review and meta-regression. *Journal of Physiotherapy* 2016;**62**(2):68-82.

Deshpande 2011

Deshpande PR, Rajan S, Sudeepthi BL, Abdul Nazir CP. Patient-reported outcomes: a new era in clinical research. *Perspectives in Clinical Research* 2011;**2**(4):137-44.

EBCTCG 2015

Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015;**386**(10001):1341-52.

Eton 2004

Eton DT, Cella D, Yost KJ, Yount SE, Peterman AH, Neuberg DS, et al. A combination of distribution-and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. *Journal of Clinical Epidemiology* 2004;**57**(9):898-910.

European Society for Medical Oncology 2014

European Society for Medical Oncology. ESMO Handbook on Rehabilitation Issues During Cancer Treatment and Follow-Up. Lugano, Switzerland: European Society for Medical Oncology, 2014.

Excel [Computer program]

Microsoft. Excel. Microsoft, 2016.

Ferlay 2012

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. International Agency for Research on Cancer. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. globocan.iarc.fr. Lyon, France: International Agency for Research on Cancer, (accessed prior to 14 March 2018). [http://globocan.iarc.fr,]

Fields 2015

Fields J. A feasibility study of a Nordic walking intervention for women experiencing aromatase inhibitor associated arthralgia [Doctoral Thesis]. Southampton, U.K.: University of Southampton, 2015.

Francis 2015

Francis P, Regan M, Fleming G, Láng I, Ciruelos E, Bellet M, et al. Adjuvant ovarian suppression in premenopausal breast cancer. *New England Journal of Medicine* 2015;**372**(5):436-46.

Francis 2018

Francis PA, Pagani O, Fleming GF, Walley BA, Colleoni M, Lang I, et al. Tailoring adjuvant endocrine therapy for premenopausal breast cancer. *New England Journal of Medicine* 2018;**379**(2):122-37.

Fransen 2014

Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: 10.1002/14651858.CD007912.pub2]

Furmaniak 2016

Furmaniak AC, Menig M, Markes MH. Exercise for women receiving adjuvant therapy for breast cancer. *Cochrane Database of Systematic Reviews* 2016, Issue 9. [DOI: 10.1002/14651858.CD005001.pub3]

Gerritsen 2016

Gerritsen JK, Vincent AJ. Exercise improves quality of life in patients with cancer: a systematic review and meta-analysis of randomised controlled trials. *British Journal of Sports Medicine* 2016;**50**(13):796-803.

GRADEpro GDT [Computer program]

GRADE Working Group, McMaster University. GRADEpro GDT. Version accessed October 2017. Hamilton (ON): GRADE Working Group, McMaster University.

Hadji 2014

Hadji P, Jackisch C, Bolten W, Blettner M, Hindenburg HJ, Klein P, et al. COMPliance and Arthralgia in Clinical Therapy: the COMPACT trial, assessing the incidence of arthralgia, and compliance within the first year of adjuvant anastrozole therapy. *Annals of Oncology* 2014;**25**(2):372-7.

Hawe 2004a

Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be?. *BMJ* 2004;**328**(7455):1561-3.

Hawe 2004b

Hawe P, Shiell A, Riley T, Gold L. Methods for exploring implementation variation and local context within a cluster randomised community intervention trial. *Journal of Epidemiology & Community Health* 2004;**58**(9):788-93.

Henry 2012

Henry NL, Azzouz F, Desta Z, Li L, Nguyen AT, Lemler S, et al. Predictors of aromatase inhibitor discontinuation as a result of treatment-emergent symptoms in early-stage breast cancer. *Journal of Clinical Oncology* 2012;**30**(9):936-42.

Hershman 2011

Hershman DL, Shao T, Kushi LH, Buono D, Tsai WY, Fehrenbacher L, et al. Early discontinuation and non-adherence to adjuvant hormonal therapy are associated with increased mortality in women with breast cancer. *Breast Cancer Research and Treatment* 2011;**126**(2):529-37.

Hershman 2015

Hershman DL, Loprinzi C, Schneider BP. Symptoms: aromatase inhibitor induced arthralgias. In: Ganz PA editor(s). Improving Outcomes for Breast Cancer Survivors: Perspectives on Research Challenges and Opportunities. Cham: Springer International Publishing, 2015:89-100.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60.

Higgins 2011a

Higgins JP, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. 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.handbook.cochrane.org..

Higgins 2011b

Higgins JP, Deeks JJ (editors). Chapter 7: Selecting studies and collecting data. 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.handbook.cochrane.org..

Higgins 2017

Higgins JP, Altman DG, Sterne JA (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Howlader 2019

Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et al. Table 1.11 Median age of cancer patients at diagnosis (2012-2016). SEER Cancer Statistics Review, (1975-2016). Bethesda, MD: National Cancer Institute., 2019. [http://seer.cancer.gov/csr/1975_2016/ Accessed on August 6, 2019]

IntHout 2014

IntHout J, Ioannidis J P, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology* 2014;**14**:25.

Irwin 2009

Irwin ML. Physical activity interventions for cancer survivors. *British Journal of Sports Medicine* 2009;**43**(1):32-8.

Kadakia 2016

Kadakia KC, Snyder CF, Kidwell KM, Seewald NJ, Flockhart DA, Skaar TC, et al. Patient-reported outcomes and early discontinuation in aromatase inhibitor-treated postmenopausal women with early stage breast cancer. *Oncologist* 2016;**21**(5):539-46.

Kemp-Casey 2017

Kemp-Casey A, Roughead EE, Saunders C, Boyle F, Bulsara M, Preen DB. Switching between endocrine therapies for primary breast cancer: frequency and timing in Australian clinical practice. *Asia-Pacific Journal of Clinical Oncology* 2017;**13**(2):e161-70.

Kilari 2016

Kilari D, Soto-Perez-de-Celis E, Mohile SG, Alibhai SM, Presley CJ, Wildes TM, et al. Designing exercise clinical trials for older adults with cancer: recommendations from 2015 Cancer and Aging Research Group NCI U13 meeting. *Journal of Geriatric Oncology* 2016;**7**(4):293-304.

Kim 2018

Kim TH, Kang JW, Lee TH. Therapeutic options for aromatase inhibitor-associated arthralgia in breast cancer survivors: a systematic review of systematic reviews, evidence mapping, and network meta-analysis. *Maturitas* 2018;**118**:29-37.

Lefebvre 2011

Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for 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.handbook.cochrane.org.

Lintermans 2013

Lintermans A, Laenen A, Van Calster B, Van Hoydonck M, Pans S, Verhaeghe J, et al. Prospective study to assess fluid accumulation and tenosynovial changes in the aromatase inhibitor-induced musculoskeletal syndrome: 2-year follow-up data. *Annals of Oncology* 2013;**24**(2):350-5.

Mao 2009

Mao JJ, Stricker C, Bruner D, Xie S, Bowman MA, Farrar J, et al. Patterns and risk factors associated with aromatase inhibitor-related arthralgia among breast cancer survivors. *Cancer* 2009;**115**(16):3631-9.

Medical Research Council 2000

Medical Research Council. A Framework for development for the development and evaluation of randomised controlled trials for complex interventions to improve health. London: Medical Research Council, 2000.

Merriam-Webster

Merriam-Webster Dictionary. Definition of therapy. www.merriam-webster.com (accessed 8 March 2018). [www.merriam-webster.com/]

Mishra 2012

Mishra SI, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, et al. Exercise interventions on health-related quality of life for cancer survivors. *Cochrane Database of Systematic Reviews* 2012, Issue 8. [DOI: 10.1002/14651858.CD007566.pub2]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Medicine* 6;**7**:e1000097. [DOI: 10.1371/journal.pmed1000097]

Nadji 2005

Nadji M, Gomez-Fernandez C, Ganjei-Azar P, Morales AR. Immunohistochemistry of estrogen and progesterone receptors reconsidered: experience with 5,993 breast cancers. *American Journal of Clinical Pathology* 2005;**123**(1):21-7.

Nahm 2018

Nahm N, Mee S, Marx G. Efficacy of management strategies for aromatase inhibitor-induced arthralgia in breast cancer patients: a systematic review. *Asia-Pacific Journal of Clinical Oncology* 2018;**14**(6):374-82.

Nipp 2017

Nipp R, Temel J. The patient knows best: incorporating patient-reported outcomes into routine clinical care. *Journal of the National Cancer Institute* 2017;**109**(9):djx044.

Niravath 2013

Niravath P. Aromatase inhibitor-induced arthralgia: a review. *Annals of Oncology* 2013;**24**(6):1443-9.

Osteras 2017

Osteras N, Kjeken I, Smedslund G, Moe RH, Slatkowsky-Christensen B, Uhlig T, et al. Exercise for hand osteoarthritis. *Cochrane Database of Systematic Reviews* 2017, Issue 1. [DOI: 10.1002/14651858.CD010388.pub2]

Partridge 2008

Partridge AH, LaFountain A, Mayer E, Taylor BS, Winer E, Asnis-Alibozek A. Adherence to initial adjuvant anastrozole therapy among women with early-stage breast cancer. *Journal of Clinical Oncology* 2008;**26**(4):556-62.

Patsou 2017

Patsou ED, Alexias GD, Anagnostopoulos FG, Karamouzis MV. Effects of physical activity on depressive symptoms during breast cancer survivorship: a meta-analysis of randomised control trials. *ESMO Open* 2017;**2**(5):e000271.

Peddle-McIntyre 2019

Peddle-McIntyre CJ, Singh F, Thomas R, Newton RU, Galvao DA, Cavalheri V. Exercise training for advanced lung cancer. *Cochrane Database of Systematic Reviews* 2019, Issue 2. [DOI: 10.1002/14651858.CD012685.pub2]

Presant 2007

Presant CA, Bosserman L, Young T, Vakil M, Horns R, Upadhyaya G, et al. Aromatase inhibitor-associated arthralgia and/or bone pain: frequency and characterization in non-clinical trial patients. *Clinical Breast Cancer* 2007;**7**(10):775-8.

Rand Health Care

Rand Health Care. 36-Item Short Form Survey (SF-36). www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html. [www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html]

Review Manager 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Roberts 2017

Roberts K, Rickett K, Greer R, Woodward N. Management of aromatase inhibitor induced musculoskeletal symptoms in postmenopausal early breast cancer: a systematic review and meta-analysis. *Critical Reviews in Oncology/Hematology* 2017;**111**:66-80.

Rock 2012

Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. *CA: a Cancer Journal for Clinicians* 2012;**62**(4):243-74.

Schünemann 2017

Schünemann HJ, Oxman AD, Higgins JP, Vist GE, Glasziou P, Akl E, et al. on behalf of the Cochrane GRADEing Methods

Group and the Cochrane Statistical Methods Group. Chapter 11: Completing 'Summary of findings' tables and grading the confidence in or quality of the evidence. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). Cochrane, 2017. Available from www.training.cochrane.org/handbook.

SIGN

Scottish Intercollegiate Guidelines Network (SIGN). Search filters. www.sign.ac.uk/methodology/filters.html#random2011.

Smaradottir 2017

Smaradottir A, Smith AL, Borgert AJ, Oettel KR. Are we on the same page? Patient and provider perceptions about exercise in cancer care: a focus group study. *Journal of the National Comprehensive Cancer Network* 2017;**15**(5):588-94.

Szlezak 2016

Szlezak AM, Szlezak SL, Keane J, Tajouri L, Minahan C. Establishing a dose-response relationship between acute resistance-exercise and the immune system: protocol for a systematic review. *Immunology Letters* 2016;**180**:54-65.

Tajaesu 2017

Tajaesu M, Tamaki K, Nagamine S, Kamada Y, Uehara K, Arakaki M, et al. Randomized trial of exercise intervention vs. usual care for breast cancer patients with aromatase inhibitor to prevent and improve the aromatase inhibitor induced arthralgia. *Cancer Research* 2017;**77**(4 Suppl):P5-12-01.

Turner 2018

Turner RR, Steed L, Quirk H, Greasley RU, Saxton JM, Taylor SJ, et al. Interventions for promoting habitual exercise in people living with and beyond cancer. *Cochrane Database of Systematic Reviews* 2018, Issue 9. [DOI: 10.1002/14651858.CD010192.pub3]

Ward 2014

Ward MM, Guthrie LC, Alba MI. Clinically important changes in Short Form 36 Health Survey Scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis Care & Research* 2014;**66**(12):1783-9.

Yang 2017

Yang GS, Kim HJ, Griffith KA, Zhu S, Dorsey SG, Renn CL. Interventions for the treatment of aromatase inhibitor-associated arthralgia in breast cancer survivors: a systematic review and meta-analysis. *Cancer Nursing* 2017;**40**(4):E26-e41.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Fields 2016

Methods	RCT; duration: 12 weeks
Participants	Inclusion • Women on AI reporting joint symptoms over preceding 12 months Exclusion • Metastatic disease • Already Nordic walking • Unable to exercise due to mobility. Safety to exercise determined by PARQ Recruited from clinic during routine follow-up, clinical team used checklist to identify potential participants. January-December 2011 n = 40 (20 in each group) Median age 63 years (55-71) More employed in intervention arm (13 vs 5). Intervention group had more chemotherapy (13 vs 7). 30% had pre-existing musculoskeletal disease
Interventions	Intervention • Weeks 1-6: supervised Nordic walking once a week for 30 min x 6 sessions; participants encouraged to have extra independent sessions of Nordic walking • Weeks 7-12: independent Nordic walking 4 x 30-min sessions a week for 6 weeks • Received written information about exercise after cancer

Fields 2016 *(Continued)*

	• Macmillan exercise diary • Phone call every 2 weeks Comparator • Enhanced usual care (phone call every 2 weeks) • Received written information about exercise after cancer • Macmillan exercise diary
Outcomes	Primary • Brief Pain Inventory- Short Form (BPI-SF): worst pain and pain interference, measured at baseline, week 6 and week 12 Secondary outcomes • PSEQ • Adherence • Depression (Center for Epidemiological Studies Depression Scale) • QoL (Medical Outcomes SF-36) Safety and exercise adherence data available
Notes	The study was funded by the Wessex Cancer Trust, Barbers' Institute RCN

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Women were randomised by an independent data manager using a random permuted blocks method, with a block size of 20 to ensure an even distribution of group size"
Allocation concealment (selection bias)	High risk	Quote: "Following randomisation, the data manager informed the researcher of the randomisation outcome, and then participants were contacted by phone by the researcher to inform them of their allocated study group. Allocation concealment was not fully implemented in view of the limited resources and staff in this feasibility study". From Fields' thesis (Fields 2015)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Unable to blind this intervention. Both participants and personnel were aware of intervention allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "It is recommended in a future study that collection of outcome measures and data analysis be carried by out those blind to group allocation to avoid any potential bias"
Incomplete outcome data (attrition bias) All outcomes	Low risk	All participants concisely accounted for. 10% attrition rate for walking intervention, however 40/40 participants completed both questionnaires therefore outcome is ITT population.
Selective reporting (reporting bias)	Low risk	Outcomes listed on p. 550-551 of the publication, and also Table 1 showing feasibility outcomes all accounted for in results. The original published protocol, registered with South Central Ethics had also proposed to include an analgesia diary, and 24-month outcomes for the waiting list control arm, both of which were not reported, but were not primary or secondary outcomes in our review. We have therefore classified this as low risk.

Fields 2016 *(Continued)*

Adherence	High risk	Patients attended an average of 5 (out of 6) supervised sessions. Only 8% of participants managed 4 x independent sessions in weeks 7-12. The majority managed 1-2 independent sessions weekly (68%-85%).
Contamination	High risk	The intervention group had a 39% increase in vigorous activity, and no change in walking activity. In the control arm, there was a 15% increase in vigorous activity, and a 45% increase in walking activity. Quote: "The exercise contamination observed in the enhanced usual care group could have led to a treatment effect".
Other bias	Low risk	No other sources of bias identified

Irwin 2015

Methods	RCT; duration: 12 months
Participants	Inclusion • Postmenopausal women, HR-positive, stage I-III BC diagnosed 0.5-4 years prior to enrolment • On AI for at least 6 months • Arthralgias for ≥ 2 months, with BPI-SF score ≥ 3/10 • Pre-existing joint pain allowed if worsened after AI • Physically inactive: baseline < 90 min exercise/week, no strength training Recruited from 5 hospitals in Connecticut, USA from June 2010-December 2012 n = 121 (61 in exercise and 60 in usual care group). Lost funding during the study, so not all participants able to complete entire 12-month programme. (45 participants completed programme in exercise arm vs 38 in control arm completed 12-month programme) Mean age: exercise group: 62 ± 7 years; usual care 60.5 ± 7 years No significant differences between groups: 85% vs 84% non-Hispanic white; 1.9 vs 1.8 years since starting AI; 52% vs 42% on pain medication; 32 vs 49% had pre-existing arthritis
Interventions	Intervention • Supervised resistance training twice a week at local health club, plus 150 min/week home-based aerobic exercise (brisk walking, cycling) for 12 months • Participants wore heart rate monitors during each workout • Physical activity log book Comparator • Usual care • Given written information regarding cancer topics • Monthly phone calls (attendance to monthly phone calls was 53%)
Outcomes	Primary outcome (Questionnaires done baseline, 3, 6, 9, 12 months) • Modified BPI assessed worst pain, pain severity and pain interference. Altered 'pain' phrase to 'joint pain/stiffness.' • WOMAC index

Irwin 2015 *(Continued)*

- DASH questionnaire
- Grip strength
- Adherence to exercise intervention (Arem 2016)

Secondary outcome

- Pain medications at 6 and 12 months
- AI adherence
- Change in weight and physical activity at 6 and 12 months

Safety data available

Collected QoL data using SF-36, FACT-G, FACT-B and FACIT-Fatigue

Preplanned subgroup analysis of those with pre-existing joint pain

Notes	Study funding from National Cancer Institute, Breast Cancer Research Foundation, Yale Cancer Centre, and National Center for Advancing Translational Science.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Permuted block randomization (at 1:1 ratio) with random block size was performed, stratified by joint pain before AI therapy and current bisphosphonate use"
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not described in article/protocol.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants and personnel unable to be blinded due to nature of the intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	PROs completed by unblinded participants. Grip strength may be influenced by motivational encouragement by assessors who are unblinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	Incomplete outcome data. Quote: "Given funding cuts, the last 25 of the 121 patients recruited were enrolled into a 6 month rather than a 12 month trial. Thus their study compliance was based on the 6 month data". Dropout imbalanced between arms: 95% completed 6 months in intervention arm, vs only 82% in usual care group. 94% vs 80% at 12 months. Dropout reason not stated.
Selective reporting (reporting bias)	Unclear risk	Protocol published prospectively with many outcomes. Most are published in the primary publication and a variety of secondary publications. Eg, QoL reported in Baglia 2019. Some primary psychological outcomes are not reported, and stiffness subscale not reported from WOMAC. Stiffness was one of the outcomes of our review.
Adherence	Low risk	Intervention arm averaged 119 min/week of aerobic exercise, with average 70% of strength training sessions completed.

Irwin 2015 *(Continued)*

Contamination	High risk	Increase in physical activity levels in the control arm. Women randomly assigned to exercise increased their physical activity an average of 159 min/week, compared with 49 min/week in the usual care group (P = 0.001)
Other bias	Low risk	No other sources of bias identified

Lohrisch 2011

Methods	Phase III RCT; duration: 48 weeks
Participants	Post-menopausal early BC, with arthralgias/myalgias related to anastrozole
Interventions	Intervention • Aerobic and resistance programme, 3 times/week for 48 weeks • Weeks 1-12 fully supervised • Weeks 13-24 supervised once weekly (2 x independent sessions/week) • Weeks 25-48 not supervised Comparator • Not described
Outcomes	Primary outcomes • Change in SF-36 bodily pain at week 12 Secondary outcomes • Change in BMD • Change in BMI • Change in hot flash index • Global QoL (SF-36 and VAS) • Change in musculoskeletal symptoms (VAS and NCIC-CTC toxicity)
Notes	Study closed early due to poor accrual, with only 22 of the proposed 72 participants enrolled. Only abstract available. Unable to get further information from the study authors.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomized". No further information given about the randomisation method.
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding of participants and personnel (performance bias) All outcomes	High risk	Unable to blind this type of intervention

Lohrisch 2011 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	PROs completed by unblinded participants. Strength testing may have been influenced by motivation from unblinded assessors. Weight measurements unlikely to be affected by unblinding
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not described. No details given for attrition
Selective reporting (reporting bias)	Unclear risk	No study protocol published. Registered trial (NCT00519883). Different secondary endpoints mentioned (e.g. difference in musculoskeletal symptoms - VAS, global QoL) and not reported, but only abstract available
Adherence	Unclear risk	Only abstract available. Not reported
Contamination	Unclear risk	Only abstract available. Not reported
Other bias	Unclear risk	Only abstract available

Nyrop 2017

Methods	RCT, duration: 6-week intervention, plus 6-month follow-up
Participants	Inclusion • Stage 0-III BC, age > 21 years (although 16%, n = 10, had stage IV disease at baseline) • Compliant on AI ≥ 4 weeks • ≥ 3/5 score on PROMIS pain intensity-short form • Exercising ≤ 150 min/week Exclusion • Concurrent chemotherapy or radiation therapy Identified through review of appointment schedule at BC clinic in tertiary care hospital. Recruited between February 2014-August 2015 n = 62 (31 in exercise, 31 in waiting list group) Median age: 63.8 ± 8.3 years Intervention group had more prior use of tamoxifen (11 vs 5) and vitamin D supplement (28 vs 19); more AI non-compliance in control group (forgets once a week = 2 vs 9). Baseline joint symptoms balanced
Interventions	Intervention • Home-based, self-directed walking, 6 weeks (Walk with Ease-Breast Cancer programme) • Given written information, workbook encouraging at least 150 min exercise/week, walking as comfortable pace • Physical activity log • No contact with investigators during the 6-week intervention period Comparator • Waiting list control- no intervention. Asked to await further contact from the research team at 6 weeks after study baseline • After 6 weeks given same materials and instructions as intervention group

Nyrop 2017 *(Continued)*

Outcomes	(Assessed at baseline, 6 weeks and 6 months after intervention. Waiting list control group had an extra questionnaire 6 weeks after completing intervention) • Physical activity levels (min/week) • VAS for pain, stiffness, fatigue • WOMAC index • FACT-G • RAI • Adherence to AI • ASE scale • Outcome expectations from exercise • Self-efficacy for physical activity • Feasibility, tolerability, adverse events
Notes	Study funding from National Cancer Institute

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "Randomized", however statements indicating serious problems - "misrandomised", "inadvertent randomisation errors". Unclear what the errors were however. Imbalances in baseline demographics, which may be a reflection of selection bias.
Allocation concealment (selection bias)	High risk	No information provided regarding allocation concealment. 5/62 participants were "mis-randomised". It is unclear whether this was a result of the randomisation process, or allocation process. To be identified as being "mis-randomised", it indicates that the allocations were not concealed from investigators.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Unable to blind this intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	All outcomes were PROs, completed by unblinded participants.
Incomplete outcome data (attrition bias) All outcomes	High risk	ITT analysis prospectively planned. However 23% of intervention arm did not complete 6-week data, vs 6% control arm. Only 77% completed 6-month data, but no mention about dropout between arms. No description of reason for not completing 6-week or 6-month questionnaires
Selective reporting (reporting bias)	Low risk	All outcomes reported
Adherence	Low risk	Mean walking time in min/week at end of 6 weeks intervention: 108.7 for intervention arm (aim was 150 min/week)
Contamination	Low risk	Mean increase in activity time in min/week between baseline and 6 weeks was 76.21 in intervention arm, and 10.52 in the waiting list control arm. The study authors did not collect information on the type or intensity of exercise that these women engaged in, only the number of min/week.
Other bias	Low risk	No other sources of bias identified

Sanmugarajah 2017

Methods	RCT; duration: 12-week intervention, plus 9-month follow-up
Participants	Included • ≥ 18 years old • Resected, stage I-III, HR-positive BC • Post-menopausal (or on LHRH agonist) • Completed adjuvant chemotherapy and radiation • ECOG 0-2 • Not on HRT • Started AI ≤ 12 weeks prior • Able to safely exercise as per PARQ Plus Exclusion • Previous AI or tamoxifen treatment • Locally advanced or metastatic BC • Significant medical conditions • Other study involvement not compatible with this one • Legal incapacity
Interventions	Intervention • Supervised programme twice weekly (consisting of moderate intensity aerobic exercise and graded resistance exercise programme) plus home walking, for total 150 min/week • 12-week programme Comparator • Information regarding importance of lifestyle management in cancer survivorship • Advice and written information about benefits of regular exercise
Outcomes	Primary outcomes • Change in BPI scores on serial assessments at 3, 6 and 12 months Secondary outcomes • Rate of AI discontinuation at 3, 6, 12 months • Compliance with exercise intervention at 3, 6, 12 months • Hand grip strength at 3, 6, 12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Randomization will be either computer generated, by a randomization programme, or by sealed envelopes"
Allocation concealment (selection bias)	High risk	Confirmed by study author correspondence that allocation concealment was not performed

Sanmugarajah 2017 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Unable to blind participants and personnel to this intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	PROs, completed by unblinded participants. Grip strength may be influenced by motivational encouragement by assessors who are unblinded.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	ITT analysis not described in the statistical plan in the protocol. The study has only been published in abstract form. Incomplete data available.
Selective reporting (reporting bias)	Unclear risk	Study only published in abstract form. Protocol made available by study author correspondence. Incomplete data available
Adherence	Unclear risk	Abstract only. Not reported
Contamination	Unclear risk	Abstract only. Not reported
Other bias	Unclear risk	Abstract only

Tamaki 2018

Methods	RCT; duration: 12 months
Participants	On an AI for 0-4 years, with no metastases. Participants could have any arthralgia level. < 75 years of age

Interventions	Intervention
	• 3 grades of exercise. Participants' choice:
	  ○ strong: 120/150 min/week of walking or running
	  ○ intermediate: daily gentle callisthenics
	  ○ weak: going up stairs
	• 12 months of exercise
	Comparator
	• No details provided

Outcomes	Primary outcomes • BPI after 12 months' exercise Secondary outcomes • BPI according to exercise strengths • Change in BPI • Adherence • Safety
Notes	Unclear how many participants in the study had AIMSS upon enrolment Only abstract and poster available. Unable to contact study authors

Tamaki 2018 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Only "randomization" mentioned. No description of method in abstract or poster
Allocation concealment (selection bias)	Unclear risk	Not described in abstract or poster
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants unable to be blinded for this intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	PROs, completed by unblinded participants. Uncertain risk of bias in adherence to AIs. Unsure if physicians in clinic were blinded to intervention
Incomplete outcome data (attrition bias) All outcomes	High risk	As per poster: 22/102 (22%) of participants dropped out of the exercise intervention group, and 9/37 (24%) dropped out of the usual care group.
Selective reporting (reporting bias)	Unclear risk	Unsure if protocol published prospectively. Only published as abstract and poster. Incomplete data available
Adherence	Unclear risk	Abstract/posters only. Not reported
Contamination	Unclear risk	Abstract/posters only. Not reported
Other bias	High risk	Study design: participants in exercise arm then able to choose between 1/3 exercise regimes, with wide variations in exercise intensity

Varadarajan 2016

Methods	Pilot RCT; duration: 8 weeks
Participants	Inclusion • Post-menopausal women with HR-positive BC • Currently on AI • Experiencing joint discomfort/stiffness when attempting ADLs Exclusion • Pre-existing rheumatoid arthritis or fibromyalgia • Systemic metastatic disease • ≥ ECOG 2
Interventions	Intervention • Exercise programme, supervised by physical therapist • 8-week intervention • No further information provided Comparator

Varadarajan 2016 *(Continued)*

	• Walking. No further details provided
Outcomes	Outcomes not provided
Notes	Only abstract available

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomly assigned" No other description of method for randomisation
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Unable to blind this intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	Unblinded participants completed PROs
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	ITT analysis not described. Attrition not mentioned
Selective reporting (reporting bias)	Unclear risk	Protocol not published. Abstract only. Not enough information provided on the outcomes
Adherence	Unclear risk	Abstract only. Not reported
Contamination	Unclear risk	Abstract only. Not reported
Other bias	Unclear risk	Abstract only

ADLs: activities of daily living; **AI:** aromatase inhibitor(s); **AIMSS:** aromatase inhibitor-induced musculoskeletal symptoms; **ASE:** Arthritis self-efficacy; **BC:** breast cancer; **BMD:** bone mineral density; **BMI:** body mass index; **BPI:** Brief Pain Inventory; **BPI-SF:** Brief Pain Inventory - Short Form; **DASH:** Disabilities of the Arm, Shoulder and Hand; **ECOG:** Eastern Cooperative Oncology Group; **FACIT:** Functional Assessment of Chronic Illness Therapy; **FACT-B:** Functional Assessment of Cancer Therapies for breast cancer; **FACT-G:** Functional Assessment of Cancer Therapies - General; **HR:** hormone receptor; **HRT:** hormone replacement therapy; **ITT:** intention-to-treat; **LHRH:** luteinising hormone-releasing hormone; **NCIC-CTC:** National Cancer Institute of Canada common toxicity criteria; **PARQ:** Physical Activity Readiness Questionnaire; **PRO:** patient/participant reported outcome; **PROMIS:** Patient Reported Outcomes Measurement Information System; **PSEQ:** Pain self-efficacy questionnaire; **QoL:** quality of life; **RAI:** Rheumatology attitudes index; **RCN:** Royal College of Nursing; **RCT:** randomised controlled trial; **SF-36:** Short-Form 36; **VAS:** visual analogue scale; **WOMAC:** Western Ontario and McMaster Universities Osteoarthritis

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

Study	Reason for exclusion
Bower 2012	Wrong outcomes. Did not investigate AIMSS

Study	Reason for exclusion
Brown 2015	Wrong patient population. No AI subgroup
Cantarero-Villanueva 2013a	Wrong study design. Not an RCT
Cantarero-Villanueva 2013b	Wrong outcomes. Did not investigate AIMSS. No AI subgroup
DeNysschen 2014	Wrong study design. Not an RCT
Desbiens 2017	Wrong outcome. Did not investigate AIMSS
Djuric 2011	Wrong patient population. No AI subgroup
Galantino 2013	Wrong study design. Not an RCT
Galiano-Castillo 2017	Wrong patient population. No AI subgroup
Goodwin 2014	Wrong outcomes. Did not investigate AIMSS. Unable to determine if population had AIMSS at baseline
Harrigan 2016	Wrong outcomes. Did not investigate AIMSS
Kiecolt-Glaser 2014	Wrong outcomes. Did not investigate AIMSS
Knobf 2017	Wrong patient population. Not specific to BC population
Lash 2011	Wrong study design. Not an RCT
Ligibel 2008	Wrong patient population. Majority received tamoxifen
Ligibel 2011	Wrong outcomes. Did not investigate AIMSS. Substudy of Goodwin 2014
Nikander 2012	Wrong patient population. No AI subgroup
Nyrop 2016	Wrong study design. No intervention
Pakiz 2016	Wrong study design. No intervention
Paulo 2019	Wrong outcomes. Did not investigate AIMSS.
Payne 2008	Wrong outcomes. Did not investigate AIMSS.
Penttinen 2009	Wrong outcomes. Did not investigate AIMSS.
Peppone 2012	Wrong patient population. Included women on tamoxifen
Peppone 2015	Wrong patient population. Included women on tamoxifen
Pruthi 2012	Wrong patient population. No AI subgroup
Reeves 2017	Wrong outcomes. Did not investigate AIMSS
Rogers 2009	Wrong patient population. No AI subgroup data available
Rogers 2009a	Wrong outcomes. Baseline AIMSS not investigated

Study	Reason for exclusion
Segal 2011	Wrong outcomes. Did not investigate AIMSS
Winkels 2017	Wrong patient population. No AI subgroup
Winters-Stone 2012	Wrong patient population. No pre-defined AIMSS subgroup

AI: aromatase inhibitor; **AIMSS:** aromatase inhibitor-induced musculoskeletal symptoms; **BC:** breast cancer; **RCT:** randomised controlled trial

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

NCT03162133

Trial name or title	Evaluating the impact of Baduanjin exercise intervention on quality of life in breast cancer survivors receiving aromatase inhibitor
Methods	RCT
Participants	86 participants, early BC, on an AI for at least 6 months
Interventions	Baduanjin exercise classes vs waiting list control
Outcomes	Primary outcome • QoL via EORTC QLQ-C30 Secondary outcomes • Change in fatigue via Piper fatigue questionnaire • Change in sleep quality via Pittsburgh sleep quality index • Change in AIMSS via BPI-SF • Change in climacteric syndrome via Kupperman Index • Change in BMD • Change in BMI • Change in 6-min walk test • Change in grip strength • Change in flexibility via test of scratching back and sit in the chair and reach • Change in balance via stand on one foot • Change in interleukin-6, tumour necrosis factor-α, interleukin-1 beta and C-reactive protein
Starting date	10 November 2016
Contact information	Kun Wang: 00862083827812 ext 50910; gzwangkun@126.com
Notes	China

NCT03786198

Trial name or title	Activity progam during aromatase inhibitor therapy
Methods	RCT

NCT03786198 *(Continued)*

Participants	350 participants, early breast cancer, on adjuvant AI
Interventions	Home-based walking intervention vs physical activity according to standard recommendations
Outcomes	Primary outcome • Incidence of muscle/joint symptoms by BPI-SF Secondary outcomes • Fatigue, EORTC QLQ-C30 • Nausea/vomiting, EORTC QLQ - Breast Cancer Module (BR-23) • Severity of muscle/joint symptoms via BPI-SF • Walking activity • Treatment adherence
Starting date	28 March 2019
Contact information	Daniele Oberti: +41313899191; trials@sakk.ch
Notes	Switzerland

NCT03953157

Trial name or title	Dietary and exercise interventions in reducing side effects in patients with stage I-IIIa breast cancer receiving aromatase inhibitors
Methods	RCT
Participants	20 participants, on an AI for at least 6 months, with mild/moderate arthralgia for at least 2 months as determined by BPI
Interventions	Anti-inflammatory diet vs exercise programme
Outcomes	Primary outcomes • BMD • Joint and muscle pain via BPI and VAS • Grip strength • Inflammatory markers
Starting date	1 August 2019
Contact information	Catherine L Carpenter: 3108258499; ccarpenter@sonnet.ucla.edu
Notes	USA

NCT03956875

Trial name or title	Yoga for aromatase inhibitor-related knee pain relief in breast cancer patients
Methods	RCT (cross-over)

NCT03956875 *(Continued)*

Participants	60
Interventions	Yoga vs massage
Outcomes	Primary outcome • WOMAC questionnaire Secondary outcome • Meridian Energy Analysis Device (M.E.A.D)
Starting date	10 March 2019
Contact information	Ching-Liang Hsieh: 0975682012; clhsieh@mail.cmuh.org.tw
Notes	Taiwan

AI: aromatase inhibitor; **AIMSS:** aromatase inhibitor-induced musculoskeletal symptoms; **BC:** breast cancer; **BMD:** bone mineral density; **BMI:** body mass index; **BPI-SF:** Brief Pain Inventory-Short Form; **EORTC QLQ:** European Organisation for Research and Treatment of Cancer Quality of Life questionnaire; **QoL:** quality of life; **RCT:** randomised controlled trial; **VAS:** visual analogue scale; **WOMAC:** Western Ontario and McMasters Universities Osteoarthritis

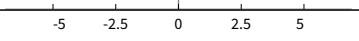
DATA AND ANALYSES

Comparison 1. Primary outcomes

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Overall change in worst pain, using WOMAC pain subscale for Nyrop 2017	4	284	Std. Mean Difference (IV, Random, 95% CI)	-0.23 [-0.78, 0.32]
2 Overall change in worst pain with VAS scale for Nyrop 2017	4	284	Std. Mean Difference (IV, Random, 95% CI)	-0.25 [-0.80, 0.30]
3 Overall change in worst pain using final values standard deviations for Irwin 2015	4	284	Std. Mean Difference (IV, Random, 95% CI)	-0.18 [-0.63, 0.26]
4 Overall change in stiffness scores	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.1 Stiffness as per WOMAC	1	53	Mean Difference (IV, Fixed, 95% CI)	-0.76 [-1.67, 0.15]
4.2 Stiffness as per VAS	1	53	Mean Difference (IV, Fixed, 95% CI)	-0.42 [-2.10, 1.26]
5 Overall change in grip strength	1	83	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.55, 1.15]

Analysis 1.1. Comparison 1 Primary outcomes, Outcome 1 Overall change in worst pain, using WOMAC pain subscale for Nyrop 2017.

Study or subgroup	Exercise arm N	Exercise arm Mean(SD)	Control arm N	Control arm Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Irwin 2015	45	-1.6 (1.7)	38	0.2 (2)		26.18%	-0.99[-1.44,-0.53]
Nyrop 2017	24	-0.4 (3.4)	29	0.1 (3.3)		24.42%	-0.14[-0.68,0.4]
Tamaki 2018	80	0 (2.4)	28	0.2 (1.8)		26.75%	-0.08[-0.51,0.35]
Fields 2016	20	-1.5 (2.2)	20	-2.5 (3.2)		22.64%	0.36[-0.27,0.98]
Total ***	**169**		**115**			**100%**	**-0.23[-0.78,0.32]**
Heterogeneity: Tau^2=0.25; Chi^2=14.15, df=3(P=0); I^2=78.79%							
Test for overall effect: Z=0.83(P=0.41)							



Favours exercise — Favours control

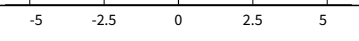
Analysis 1.2. Comparison 1 Primary outcomes, Outcome 2 Overall change in worst pain with VAS scale for Nyrop 2017.

Study or subgroup	Exercise arm N	Exercise arm Mean(SD)	Control arm N	Control arm Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Fields 2016	20	-1.5 (2.2)	20	-2.5 (3.2)		22.65%	0.36[-0.27,0.98]
Irwin 2015	45	-1.6 (1.7)	38	0.2 (2)		26.19%	-0.99[-1.44,-0.53]
Nyrop 2017	24	-0.7 (2.4)	29	-0.1 (2.4)		24.4%	-0.25[-0.8,0.29]
Tamaki 2018	80	0.1 (2.4)	28	0.2 (1.8)		26.77%	-0.05[-0.48,0.38]
Total ***	**169**		**115**			**100%**	**-0.25[-0.8,0.3]**
Heterogeneity: Tau^2=0.25; Chi^2=14.13, df=3(P=0); I^2=78.77%							
Test for overall effect: Z=0.91(P=0.37)							

Favours exercise — Favours control

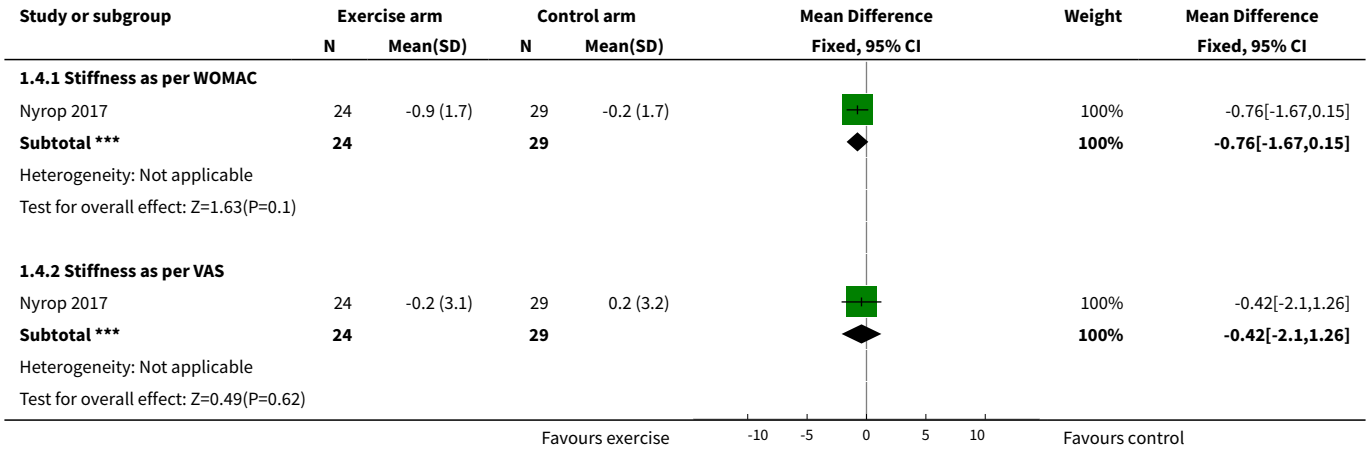
Analysis 1.3. Comparison 1 Primary outcomes, Outcome 3 Overall change in worst pain using final values standard deviations for Irwin 2015.

Study or subgroup	Exercise arm N	Exercise arm Mean(SD)	Control arm N	Control arm Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Fields 2016	20	-1.5 (2.2)	20	-2.5 (3.2)		21.43%	0.36[-0.27,0.98]
Irwin 2015	45	-2 (2.3)	38	-0.2 (2.4)		27.01%	-0.76[-1.21,-0.31]
Nyrop 2017	24	-0.4 (3.4)	29	0.1 (3.3)		23.97%	-0.14[-0.68,0.4]
Tamaki 2018	80	0 (2.4)	28	0.2 (1.8)		27.59%	-0.08[-0.51,0.35]
Total ***	**169**		**115**			**100%**	**-0.18[-0.63,0.26]**
Heterogeneity: Tau^2=0.14; Chi^2=9.33, df=3(P=0.03); I^2=67.85%							
Test for overall effect: Z=0.82(P=0.41)							

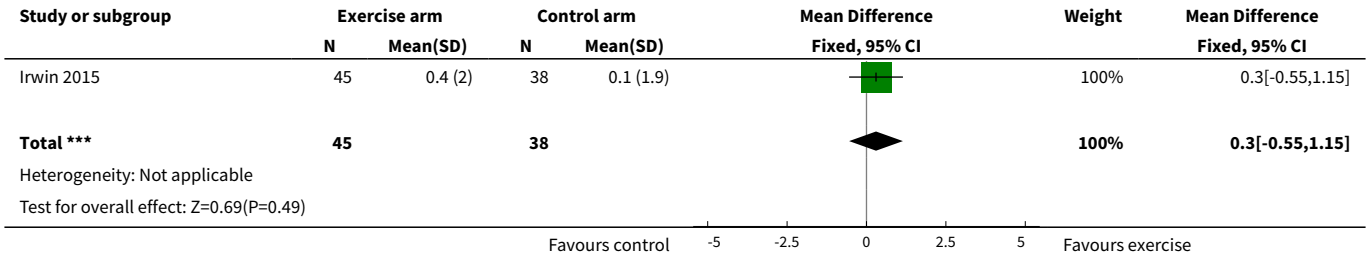


Favours exercise — Favours control

Analysis 1.4. Comparison 1 Primary outcomes, Outcome 4 Overall change in stiffness scores.



Analysis 1.5. Comparison 1 Primary outcomes, Outcome 5 Overall change in grip strength.

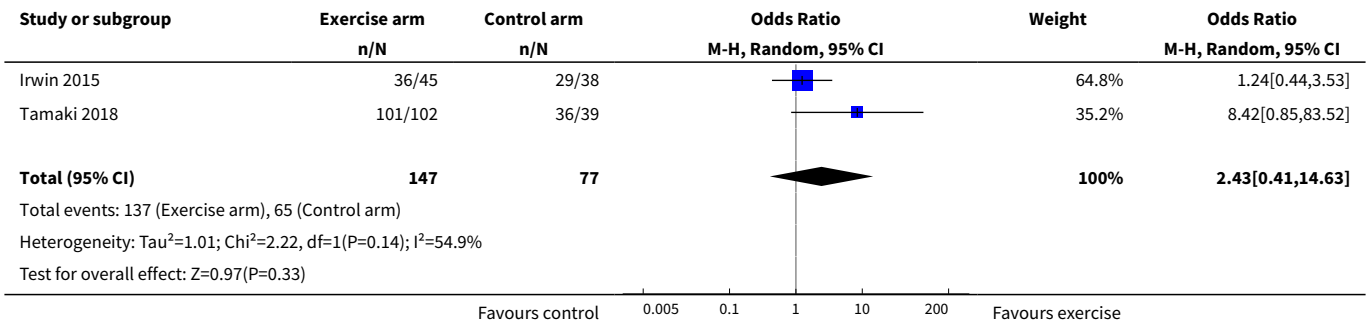


Comparison 2. Secondary outcomes

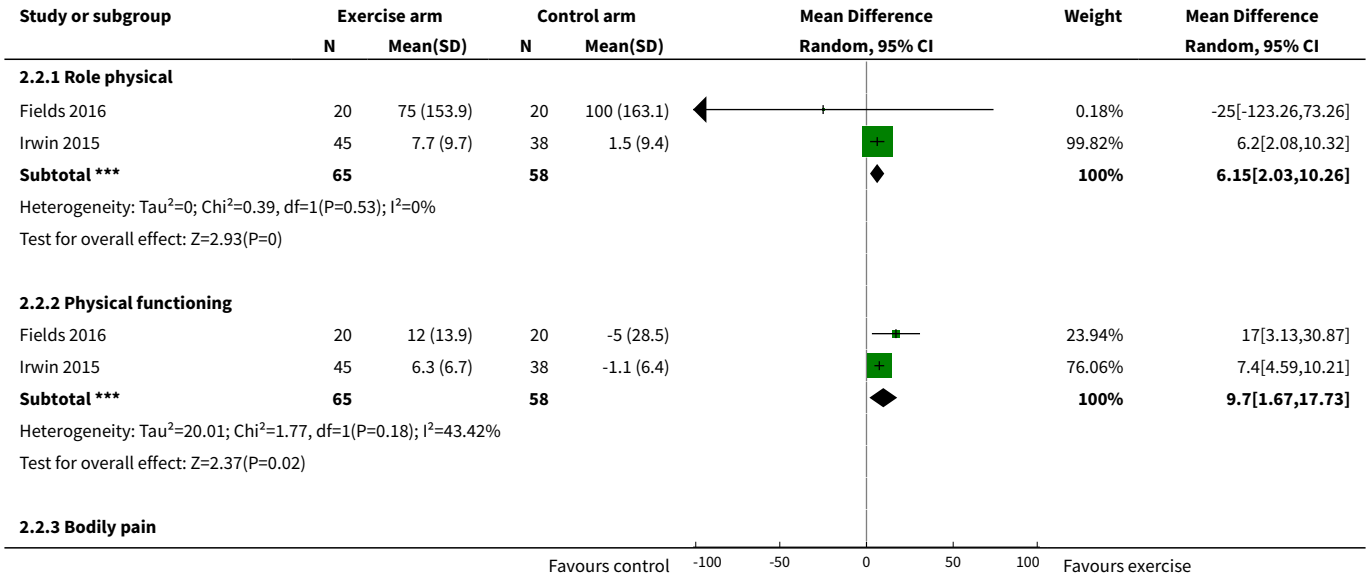
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Persistence and compliance with aromatase inhibitors	2	224	Odds Ratio (M-H, Random, 95% CI)	2.43 [0.41, 14.63]
2 Health-related quality of life	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Role physical	2	123	Mean Difference (IV, Random, 95% CI)	6.15 [2.03, 10.26]
2.2 Physical functioning	2	123	Mean Difference (IV, Random, 95% CI)	9.70 [1.67, 17.73]
2.3 Bodily pain	2	123	Mean Difference (IV, Random, 95% CI)	7.60 [4.51, 10.70]
2.4 General health	2	123	Mean Difference (IV, Random, 95% CI)	3.62 [0.92, 6.33]

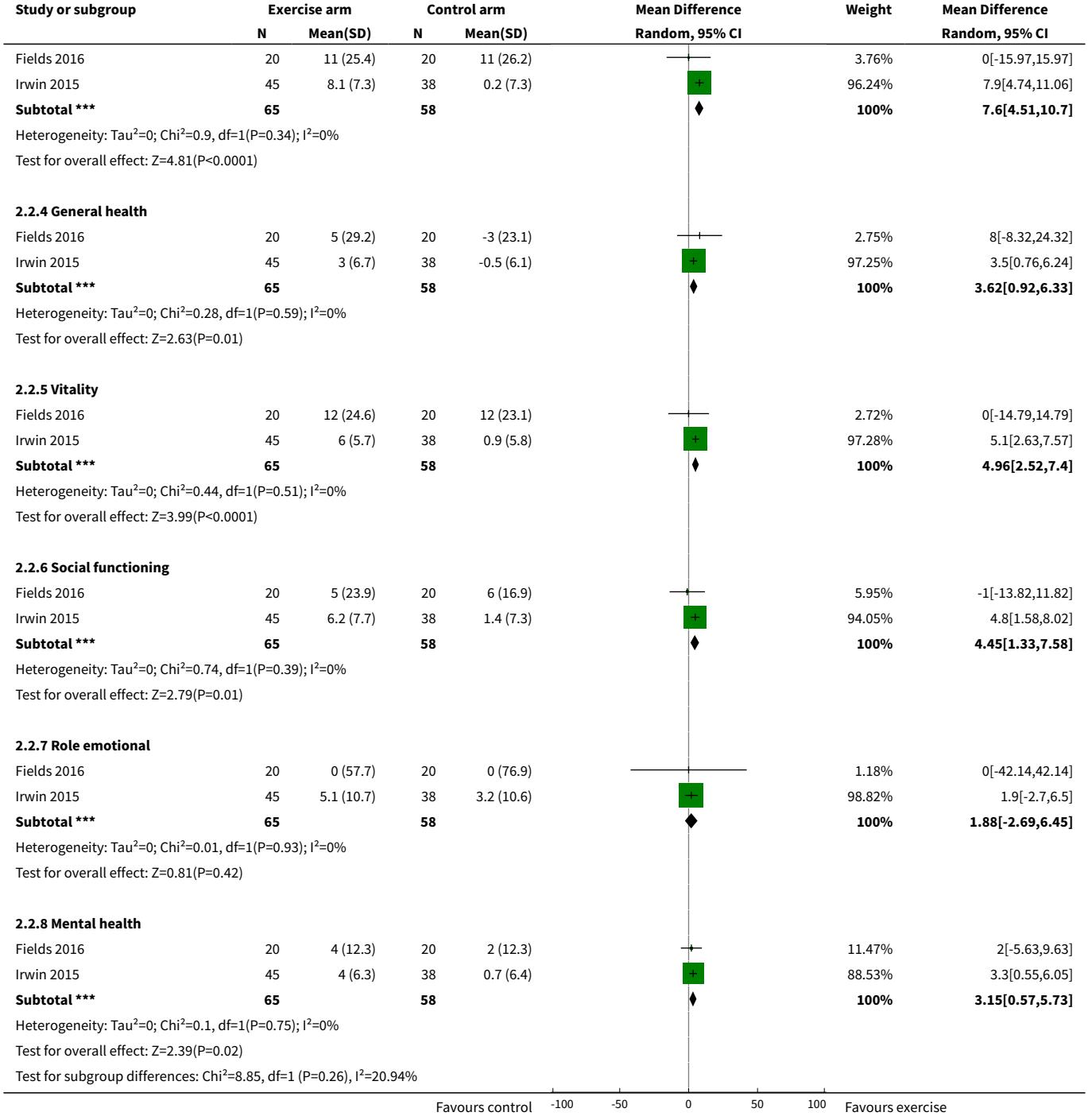
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.5 Vitality	2	123	Mean Difference (IV, Random, 95% CI)	4.96 [2.52, 7.40]
2.6 Social functioning	2	123	Mean Difference (IV, Random, 95% CI)	4.45 [1.33, 7.58]
2.7 Role emotional	2	123	Mean Difference (IV, Random, 95% CI)	1.88 [-2.69, 6.45]
2.8 Mental health	2	123	Mean Difference (IV, Random, 95% CI)	3.15 [0.57, 5.73]
3 Cancer-specific quality of life	2	136	Mean Difference (IV, Random, 95% CI)	4.58 [-0.61, 9.78]

Analysis 2.1. Comparison 2 Secondary outcomes, Outcome 1 Persistence and compliance with aromatase inhibitors.

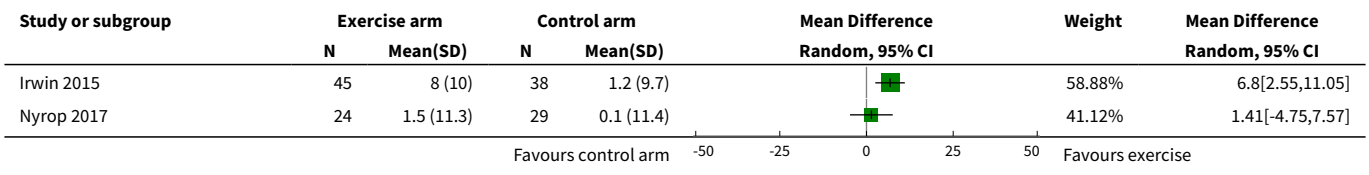


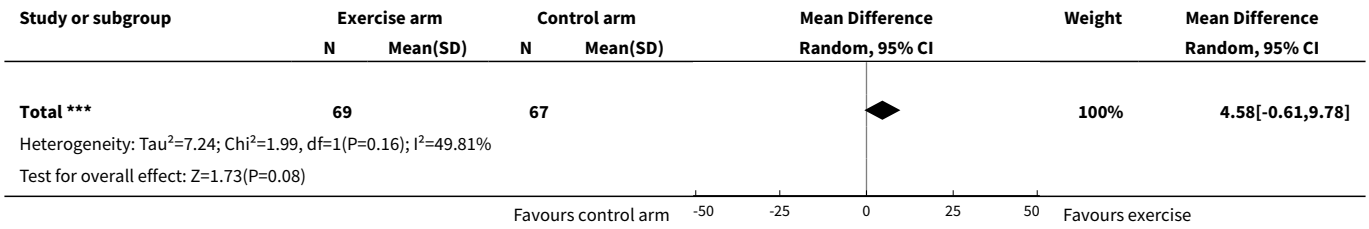
Analysis 2.2. Comparison 2 Secondary outcomes, Outcome 2 Health-related quality of life.





Analysis 2.3. Comparison 2 Secondary outcomes, Outcome 3 Cancer-specific quality of life.





APPENDICES

Appendix 1. CENTRAL

#1 MeSH descriptor: [Aromatase Inhibitors] explode all trees

#2 aromatase inhibit* (Word variations have been searched)

#3 anastrozole or exemestane or letrozole or aminoglutethimide* or atamestane or fadrozole or formestane or vorozole or arimidex or aromasin or femara or fadrozole or lentaron or rivizor or cytadren (Word variations have been searched)

#4 #1 or #2 or #3

#5 MeSH descriptor: [Breast Neoplasms] explode all trees

#6 breast near cancer*

#7 breast near (tumour* or tumor*)

#8 breast near malignan*

#9 breast near (carcinoma* or adenocarcinoma*)

#10 #5 or #6 or #7 or #8 or #9

#11 (physical or strength* or resistance or isometric*)

#12 (exercise* or activit* or therap* or program* or training)

#13 #11 near #12

#14 exercise near (therap* or program* or training*)

#15 MeSH descriptor: [Exercise Therapy] explode all trees

#16 MeSH descriptor: [Exercise] explode all trees

#17 sport or sports* or walk* or swim* or aquatic or danc* or running or jogging or aerobic* or pilates or qigong or "qi gong" or "chi kung" or "chi gung" or exercis* or gym* or isometric*

#18 tai chi or t'ai chi or taijiquan or yoga or yogi* or dhyan or pranayam or asana or bikram or vinyasa or hatha or ashtanga or iyengar or kundalini

#19 #13 or #14 or #15 or #16 or #17 or #18

#20 #4 and #10 and #19 [in trials]

Appendix 2. MEDLINE

1. "exemestane"[Supplementary Concept])

2. "Aromatase Inhibitors"[Mesh])

3. “Aromatase Inhibitors"[Pharmacological Action]

4. "letrozole"[Supplementary Concept]

5. "Aminoglutethimide"[Mesh]

6. "anastrozole"[Supplementary Concept]

7. "atamestane"[Supplementary Concept]

8. "Fadrozole"[Mesh]

9. "formestane"[Supplementary Concept]

10."vorozole"[Supplementary Concept]

11.aromatase inhibitor*

12.anastrozole

13.arimidex

14.exemestane

15.letrozole

16.aromasin

17.femara

18.fadrozole

19.formestane

20.rivizor

21.cytadren

22.aminoglutethimide

23. 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

24. (breast* OR mammary) AND (cancer OR cancers OR cancerous OR carcinoma OR malignant* OR tumor OR tumors OR tumour OR tumours OR adenocarcinoma*)

25. ("Breast"[Mesh] OR Breast Diseases"[Mesh]) AND "Neoplasms"[Mesh]

26."Breast Neoplasms"[Mesh]

27. 24 OR 25 OR 26

28."Exercise Therapy"[Mesh]

29. "Exercise Movement Techniques"[Mesh]

30. "Sports"[Mesh]

31. "Dancing"[Mesh]

32. "Exercise"[Mesh]

33. "Resistance training"[MeSH Terms]

34. dhyan*[Text Word] OR pranayam*[Text Word] OR asana [Text Word] OR bikram [Text Word] OR vinyasa [Text Word] OR hatha [Text Word] OR ashtanga [Text Word] OR iyengar [Text Word] OR kundalini [Text Word] OR yoga OR yogi*

35. (sport OR sports* OR walk* OR swim* OR aquatic OR danc* OR running OR jogging OR aerobic* OR pilates OR qigong OR "qi gong" OR "chi kung" OR "chi gung" OR exercis* OR gym* OR isometric* OR "tai chi" OR "t'ai chi" OR taijiquan)

36. (exercise* AND (therap* OR program* OR training*))

37. (physical OR strength* OR resistance OR isometric) AND (exercis* OR activit* OR therapy OR therapies OR therapeutic OR program* OR training)

38. 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 OR 35 OR 36 OR 37

39. randomized controlled trial[Publication Type]

40. controlled clinical trial[Publication Type]

41. randomized[Title/Abstract]

42.randomised[Title/Abstract]

43.randomly[Title/Abstract]

44. placebo[Title/Abstract]

45.trial[Title/Abstract])

46.groups[Title/Abstract]

47. 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45 OR 46

48. 23 AND 27 AND 38 AND 47

49. "Animals"[Mesh] NOT "Humans"[Mesh]

50. 48 NOT 49

Appendix 3. Embase

(((aromatase NEAR/2 inhibit* OR 'aromatase inhibitor'/exp OR anastrozole OR exemestane OR 'letrozole' OR aminoglutethimide* OR atamestane OR fadrozole OR formestane OR vorozole OR arimidex OR aromasin OR femara OR fadrozole OR lentaron OR rivizor OR cytadren)

AND

('breast cancer'/exp OR (breast OR mammary) NEAR/3 (cancer* OR carcinoma* OR malignan* OR tumo*r* OR adenocarcinoma*) OR ('neoplasm'/exp AND ('breast'/exp OR 'breast disease'/exp)))

AND

('sport'/exp OR 'dancing'/exp OR 'exercise'/exp OR 'walking'/exp OR 'physical activity'/exp ‘resistance training’/exp) OR (sport* OR walk* OR swim* OR aquatic OR danc* OR running OR jogging OR aerobic* OR pilates OR exercis* OR gym* OR isometric*) OR (sport*:ti,ab OR walk*:ti,ab OR swim*:ti,ab OR aquatic:ti,ab OR danc*:ti,ab OR running:ti,ab OR jogging:ti,ab OR aerobic*:ti,ab OR pilates:ti,ab OR exercis*:ti,ab OR gym*:ti,ab OR isometric*:ti,ab) OR (exercise* NEAR/3 (therap* OR program* OR training*)):ti,ab OR ((physical OR strength OR resistance OR isometric) NEAR/3 (exercis* OR activity* OR therap* OR program* OR training)):ti,ab OR (qigong:ti,ab OR 'qi gong':ti,ab OR 'chi kung':ti,ab OR 'chi gung':ti,ab) OR ('tai chi' OR 't?ai chi' OR taijiquan) OR (yoga:ti,ab OR yogi*:ti,ab) OR (dhyan:ti,ab OR pranayam:ti,ab OR asana:ti,ab OR bikram:ti,ab OR vinyasa:ti,ab OR hatha:ti,ab OR ashtanga:ti,ab OR iyengar:ti,ab OR kundalini:ti,ab)))

AND

random* OR factorial* OR crossover* OR cross NEXT/1 over* OR placebo* OR (doubl* AND blind*) OR (singl* AND blind*) OR assign* OR allocat* OR volunteer* OR 'crossover procedure'/exp OR 'double blind procedure'/exp OR 'randomized controlled trial'/exp OR 'single blind procedure'/exp

Appendix 4. CINAHL

S1 (MH "Aromatase Inhibitors+")

S2 TX (aromatase N3 inhibit*)

S3 TX exemestane

S4 TX letrozole

S5 TX Aminoglutethimide*

S6 TX atamestane

S7 TX fadrozole

S8 TX formestane

S9 TX vorozole

S10 TX arimidex

S11 TX aromasin

S12 TX femara

S13 TX fadrozole or TX anastrozole or TX rivizor or TX cytadren or TX lentaron

S14 TX hormon* W1 therapy*

S15 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

S16 (MH "Breast Neoplasms+")

S17 (MH "Breast+")

S18 (MH "Neoplasms+")

S19 S18 AND S19

S20 TX ((breast* OR mammary) N3 (cancer* OR carcinoma* OR adenocarcinoma* OR malignan* OR tumo#r*))

S21 S16 or S19 or S20

S22 (MH "Therapeutic Exercise+")

S23 (MH "Exercise+")

S24 (MH "Resistance Training")

S25 (MH "Physical Activity")

S26 (MH "Physical Fitness") OR (MH "Physical Performance") OR (MH "Sports+")

S27 (MH "Walking+") or (MH "Swimming")

S28 (MH "Dance Therapy") OR (MH "Dancing+")

S29 (MH "Yoga+") OR (MH "Tai Chi")

S30 (MH "Qigong")

S31 TX sport OR sports* OR walk* OR swim* OR aquatic OR danc* OR running OR jogging OR aerobic* OR pilates OR qigong OR "qi gong" OR "chi kung" OR "chi gung" OR exercis* OR gym* OR isometric*

S32 TX ((physical OR strength* OR resistance or isometric*) N3 (exercis* OR activit* OR therap* OR program* OR training)) or TX (exercise W6 therap*)

S33 TX dhyan* OR pranayam* OR asana OR bikram OR vinyasa OR hatha OR ashtanga OR iyengar OR kundalini OR yoga OR yogi*

S34 TX "tai chi"

S35 TX "t'ai chi" or or TX (tai ji) or TX (taijiquan)

S36 S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30 or S31 or S32 or S33 or S34 or S35

S37 (MH "Clinical Trials+")

S38 PT Clinical trial

S39 TX clinic* n1 trial*

S40 TX ((singl* n1 blind*) or (singl* n1 mask*)) or TX ((doubl* n1 blind*) or (doubl* n1 mask*)) or TX ((tripl* n1 blind*) or (tripl* n1 mask*)) or TX ((trebl* n1 blind*) or (trebl* n1 mask*))

S41 TX randomi* control* trial*

S42 (MH "Random Assignment")

S43 TX random* allocat*

S44 TX placebo*

S45 (MH "Placebos")

S46 (MH "Quantitative Studies")

S47 TX allocat* random*

S48 S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45 or S46 or S47

S49 S15 and S21 and S36 and S48

Appendix 5. WHO ICTRP

breast cancer AND aromatase AND exercise

breast cancer AND aromatase AND yoga

breast cancer AND aromatase AND training

breast cancer AND aromatase AND physical activity

Appendix 6. ClinicalTrials.gov

breast cancer AND aromatase | exercise OR physical OR yoga OR activity OR training OR walking | Studies with Female Participants

CONTRIBUTIONS OF AUTHORS

- Draft the protocol: KER, KR, NW, DV
- Study selection: KER, KR, NW, SF
- Extract data from studies: KER, SF
- Enter data into Review Manager 2014: KER, KR
- Carry out the analysis: KER, DV
- Interpret the analysis: KER, NW, DV
- Draft the final review: KER, KR, NW, DV, SF
- Disagreement resolution: NW, KER
- Update the review: KER, KR, NW, DV, SF

DECLARATIONS OF INTEREST

KER: received sponsorship from Roche, Amgen & BMS for accommodation, travel and registration to educational meetings. This financial support is unrelated to the topic under review.
KR: none to declare
DV: none to declare
SF: travel/accommodation by Eisai. This financial support is unrelated to the topic under review.
NW: consultancy fees from Roche, Novartis; grants from Medivation; expert panel review for Roche and Pfizer; travel/accommodation/meeting expenses for Roche, Novartis; stock in CSL. While both Pfizer and Novartis do market aromatase inhibitors (and these medicines are now off patent), this review focuses on managing side effects rather than assessing their efficacy. Relationship with Roche is not relevant to a review of managing aromatase inhibitor side effects.

SOURCES OF SUPPORT

Internal sources

- Not applicable, Other.

External sources

- Not applicable, Other.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol one of our secondary outcomes was 'overall change in quality of life'. Due to the diversity of quality-of-life scoring tools used amongst the studies, we felt that it was not appropriate to merge the findings of the quality-of-life tools into one overall meta-analysis. Instead, we divided this outcome into two subgroups, which included overall change in health-related quality of life, and overall change in cancer-specific quality of life. This enabled us to analyse different quality-of-life scoring tools and present the findings in a more appropriate way.

In our protocol, we had planned to only do random-effects meta-analysis as we had expected clinical heterogeneity between studies. In our revised method, we have also reported the results of the fixed-effect model for each assessment. Due to the small number of studies, and small number of participants in some studies, we also performed a random-effects meta-analysis using the Hartung, Knapp, Sidik and Jonkman (HKSJ) approach (IntHout 2014).

Due to the limited studies available, and limited reporting of results, we were unable to undertake further subgroup analysis. In our protocol, we had planned to undertake further subgroup analysis based on:

- type of exercise (i.e. aerobic/resistance/combination/other);
- supervised versus home-based; and
- intensity of treatment (i.e. mild/moderate/vigorous)

INDEX TERMS

Medical Subject Headings (MeSH)

*Exercise Therapy [methods]; Aromatase Inhibitors [*adverse effects] [therapeutic use]; Breast Neoplasms [drug therapy]; Musculoskeletal Pain [*chemically induced] [*prevention & control]; Neoplasm Staging; Randomized Controlled Trials as Topic

MeSH check words

Female; Humans