SUPPLEMENTARY APPENDIX 3: Evidence Report

2022 American College of Rheumatology (ACR) Guideline for Exercise, Rehabilitation, Diet, and Additional Integrative Interventions for Rheumatoid Arthritis

Prepared for: American College of Rheumatology

Literature review team:

Jonathan R Treadwell, Ph.D.

Rawan AlHeresh, MSOT, Ph.D., OTR/L

Kamil Barbour, PhD, MPH

Thomas Bye PT, DPT, MS

Dana Guglielmo, MPH

Rebecca Haberman, MD

Tate Johnson, MD

Anatole Kleiner, MD, FAAAAI, FACR

Chris Lane, PT, DPT

Linda Li, PT, PhD

Hiral Master, PT, PhD, MPH

Daniel Pinto, PT, PhD

Janet Poole, PhD, OTR/L

Kimberly Steinbarger, PT, MHS, DHSc

Daniel Sztubinski

Louise Thoma, PT, PhD

Vlad Tsaltskan, MD

Marat Turgunbaev, MD

Courtney Wells, PhD, MPH, MSW

Introduction

Interventions: The following interventions were within the scope (in MS Word, ctrl-click a PICO to skip to its evidence summary; when done, ctrl-home to come back here):

PICO 16. Should patients with RA use environmental adaptations?

Psychosocial and vocational

PICO 17: Should patients with RA participate in comprehensive occupational therapy?

PICO 18: Should patients with RA participate in a comprehensive physical therapy program?

PICO 19: Should patients with RA use a standardized, evidence-based self-management program?

PICO 20: Should patients with RA use mind-body approaches?

PICO 21. Should patients with RA, who are currently employed or want to become employed, use vocational rehabilitation?

PICO 22: Should patients with RA, who are currently employed or want to become employed, receive work site evaluations and modifications?

Adjunctive therapies

PICO 23: Should patients with RA use acupuncture?

PICO 24: Should patients with RA receive massage therapy?

PICO 25: Should patients with RA receive thermal modalities?

PICO 26: Should patients with RA receive electrotherapy?

PICO 27. Should patients with RA receive chiropractic therapy?

Other

PICO 28. Should patients with RA who are current smokers engage in a smoking cessation program?

Study designs

- While randomized controlled trials (RCTs) were the preferred source of evidence, we also included non-randomized comparative studies that addressed a PICO. Like RCTs, these studies must have had two or more separate groups of RA patients who received different management strategies (e.g., comprehensive physical therapy vs. waiting list).
- We required that studies reported data exclusive to RA patients, or if studies did include other patients, at least 80% in each group had RA.
- Studies must have reported data on at least 10 patients per group at follow-up.

- Studies must have reported data at least two weeks after the start of the intervention. We defined short-term data as <12 weeks, and long-term data as 12 weeks or more.

Comparators

- For all 28 PICOs, we included comparisons to an inactive treatment (e.g., splint vs no splint for PICO 9).
- For seven of the 28 PICOs, we also included "active" comparisons within the same category (e.g., one diet versus a different diet for PICO 1). These were PICOs 1, 2, 4, 5, 6, 7, and 20.

Outcomes

- We considered 14 categories of outcomes: Disease activity, Radiographic progression, Functional status, Pain, Fatigue, QOL, Treatment related harms, Mortality, CVD, Joint replacement, Self efficacy, Work-related outcomes, Sleep status, and Mental health status.
- Not all outcome categories were included for all PICOs. For specifics, see individual PICO evidence summaries.
- For disease activity, we considered only the following as direct measures: DAS28, CDAI, SDAI, RAPID3, PAS, PAS2, DAS (44 joints), ACR20, ACR50, and ACR70. Other instruments, radiographic progression, and fatigue were considered surrogate measures of disease activity.
- For all PICOs, pain and function were deemed "critical" outcomes for the purpose of applying the GRADE system. For PICOs 1/2/3/28, disease activity was also a critical outcome, because the mechanism of action of those interventions (dietary interventions and smoking cessation) can target disease activity. By contrast, the interventions for other PICOs pain and function, not disease activity. Therefore, disease activity was only critical for PICOs 1/2/3/28.
- Not every included study reported critical outcomes. Each outcome was analyzed separately.
- Many exercise interventions (PICOs 4, 5 and 6) involved multiple forms of exercise (e.g., both aerobic and resistance exercise). To address this complexity, we created a new PICO, called PICO 4-5-6, which included all studies of multicomponent exercise.
- For the four exercise PICO's (4 through 7), we also combined the 12 weeks+ data from RCTs with inactive control groups that reported critical outcomes (pain, function), in order to estimate the effects of any exercise.

Quality Assessment

- Quality assessment was performed separately for each outcome using the GRADE system, which results in one of four possible evidence grades that reflect level of confidence in the effect estimate: high, moderate, low, and very low.
- Study design is the starting point for quality assessment: randomized controlled trials (RCTs) start at high quality and non-randomized studies start at low quality.
- Five factors can lower the quality of evidence grade: risk of bias, inconsistency, indirectness, imprecision, and publication bias.
- Risk of bias refers to limitations in study design or execution (e.g., lack of allocation concealment or blinding).

- Inconsistency refers to unexplained heterogeneity in results of studies evaluating the same outcome.
- Indirectness refers to lack of direct comparisons of interventions of interest (e.g. studies comparing drug A vs. placebo and drug B vs. placebo when the comparison of interest is drug A vs. drug B), lack of applicability in the interventions or populations being evaluated, or use of indirect (surrogate) outcome measures.
- Imprecision refers to uncertainty in the estimate of effect due to very low numbers of patients or events and/or wide 95% confidence intervals that cross a clinical decision threshold (i.e. between recommending and not recommending treatment).
- Publication bias refers to selective publication of studies that show greater treatment effects (i.e. negative studies are suppressed).
- Quality of evidence can vary from outcome to outcome. The final quality assessment for the PICO question is based on the critical outcome with the lowest quality assessment.
- The level of evidence listed in this report for either an individual paper or a group of papers is not meant to be an absolute statement about the quality of the study (or studies) under consideration. Rather, the intention is to rate the paper(s) *in relation to the question being asked in this guideline*. Because of this, a very well-conducted study might actually be rated down in this evidence report, possibly because the population or intervention being studied does not match the population or intervention being examined by the PICO question in this guideline (in other words, downgrading for indirectness). The level of evidence may also be downgraded due to imprecision in the effect estimate (wide confidence intervals that cross the line of no effect, or a low number of patients or events). A combination of these factors may result in quality of evidence from a well-conducted study being rated as low.

Presentation of effects

- Most outcomes were reported on continuous scales as means and standard deviations.
- We presented effect sizes as between-group differences, along with its 95% confidence intervals.
- When multiple studies measured the same construct using different instruments (e.g., DAS28 or CDAI for measuring disease activity), we used the standardized mean difference (SMD) in order to perform meta-analyses. We then converted this SMD into a more easily understood metric by multiplying it by the typical standard deviation of the more commonly used scale. For DAS28, we used an SD of 1.1. For other outcomes, the SD for presentation was based on the studies included in the meta-analysis.
- Relative effects for dichotomous outcomes in the tables are expressed as relative risk (RR) or odds ratio (OR). RR is the default effect size because it is more easily interpretable.

Interpreting the evidence

- It is important to take into account the information presented specifically as it relates to the question of interest. For example, when the only evidence for a given PICO question is indirect due to the comparison or patient population, it appropriately gets downgraded for indirectness as shown under the column labeled “indirectness.” Also, if the 95% confidence interval around an effect size is wide and

crosses the line of no difference between treatments, the evidence for that outcome is downgraded due to imprecision. Study design and risk of bias also may result in downgrades in the quality of evidence. The overall quality of evidence takes all these factors into account, and is appropriately rated as high, moderate, low, or very low. This quality of evidence is key to your decisions.

Moving from evidence to recommendations

- In GRADE, recommendations can be either strong or conditional. Generally, strong recommendations are restricted to high or moderate quality evidence. Low quality evidence almost invariably mandates a weak recommendation.
- There are, however, situations in which low quality evidence can lead to strong recommendations. For instance, if there is low quality evidence favoring an intervention but high quality evidence of important harm then a strong recommendation against the intervention may be appropriate.

Bibliography of included studies

- Separate reference lists of studies included for each PICO question with an evidence base appear at the end of the summaries for each question.

Diet

PICO 1: Should patients with RA use a formally defined diet?

PICO 1 included 21 articles, 19 RCTs and 2 non-randomized studies (Fraser, 2000 ; Siddique, 2020). They investigated 10 different diets:

- Vegan diet (5 studies) (Tables 1,2)
- Mediterranean diet (4 studies) (Tables 3, 4, 5)
- Anti-Inflammatory diet (4 studies) (Table 6)
- Exclusion/elimination diet (3 studies) (Tables 7, 8)
- Fasting (2 studies) (Tables 9, 10)
- Elemental Peptide diet (1 study) (Tables 11, 12)
- Graastener Diet (1 study) (Table 13)
- High Saturated Fat/Low Unsaturated Fat/Hypoallergenic Diet (1 study) (Table 14)
- Arthritis Diet (1 study) (Tables 15, 16)
- Low dose of food sensitivities (1 study) (Table 17)

All were compared to no change in diet, and in addition, one study compared fasting to ketogenic diet, and one study compared High SatFat/Low UnsatFat/Hypoallergenic to a well-controlled diet. One study (Gianfranceschi 1996) was a crossover study that contained elements in both the exclusion/elimination and the food sensitivity diets.

Five studies were included in the comparison for the use of a vegan diet (Hanninen 2000, Elkan 2008, Helve 1998, Peltonen 1997, Hafstrom 2001). Several studies reported positive outcomes for a vegan diet, however the level of certainty was very low. Results favored the vegan diet in the study by Peltonen for reduction in disease activity. In the study by Hafstrom, a greater percentage of patients on the vegan diet achieved ACR20 than those with no dietary change. Certainty of evidence for these outcomes was low to very low. Helve and colleagues (1998) also found improvements in pain, joint swelling, morning stiffness, and ability to move for those consuming an uncooked vegan diet prepared in a specialized kitchen vs. no change in diet.

Evidence reported for the Mediterranean diet (MD) demonstrated positive effects on disease activity, function, and pain when compared to no change in diet. There were 4 studies that explored the effects of the Mediterranean diet on patients with RA. Skoldstam (2003) and Hagfors (2005) found significant reductions in the HAQ, swollen joint count, global VAS (patient's global assessment of disease activity by means of a 0 –

100 mm VAS), Pain VAS, and the duration of morning stiffness, as well as improvement in the SF-36 subsets of physical and social functioning when compared to no change in diet. The certainty of evidence for these measures was Low. Garcia-Morales (2020) found the combination of MD + dynamic exercise program (DEP) showed more significant improvements in health-related quality of life (SF-36) than either intervention alone. Certainty of evidence for this study was low. Pineda-Juarez (2020) also compared groups utilizing a dynamic exercise program (DEP)+ a Mediterranean diet (MD), just DEP, and just MD, and found that the DEP + MD group had greater improvements in HAQ score (disability) than the MD-alone and DEP-alone groups

One of four anti-inflammatory diet studies reported a significant improvement in swollen joint and tender joint scores favoring those who consumed an anti-inflammatory diet vs. no diet change (Adam, 2003). The level of certainty for this study was very low. The other three studies, Vadell (2020), Turesson Wadell (2021), and Ghaseminasab-Parizi (2022) also compared an anti-inflammatory diet to no diet change, but found no statistically significant results. Certainty of evidence overall was very low for this type of diet.

Two of three studies of an Exclusion/Elimination diet (Darlington, 1986; Gianfranceschi, 1996) found significant improvement in pain and morning stiffness measures, compared to no change in diet, with very low certainty of evidence. The third study (Guagnano, 2021) reported only medians showing less pain and increased SF-36 scores after 3 months on this type of diet. Pfeiffer and colleagues (1998) found significant improvement in the ACR20 for those on an Elemental Peptide Diet vs. no change in diet. The level of certainty was very low.

Fraser and colleagues (2000) compared a 1 week fasting + 2 weeks of a lacto-vegetarian diet with a 1 week Ketogenic diet + 2 weeks of a lacto-vegetarian diet. The fasting group demonstrated significant improvement in disease activity when compared to the ketogenic group. Siddique (2020) found no statistically significant between-group difference in the amount of improvement with fasting vs non-fasting, and determined that fasting did not increase disease activity, so was safe for those who wished to fast during the holidays.

Hansen and colleagues (1996) found significantly fewer swollen joints in those who consumed the Graastener Diet compared to those who had no dietary changes. The Graastener diet had been composed of an energy intake adjusted so as to obtain near-standard BMI (body mass index), with lower fat and increased protein (vegetable and fish) consumption, plus supplements of vitamins A, C, E and selenium. Certainty of evidence was low for this single study.

Other diets that were reported had only single studies, low to very low certainty, and/or no statistically significant results. These included High Saturated Fat/Low Unsaturated Fat/Hypoallergenic Diet (Sarzi-Puttini 2000), the Arthritis Diet (Panush 1983), and Low dose of food sensitivities (Gianfranceschi, 1996).

Below, we discuss the details regarding each of the comparisons made by the included studies.

Vegan diet versus no change in diet

Five studies included this comparison (Hanninen 2000, Elkan 2008, Helve 1998, Peltonen 1997, Hafstrom 2001) (Table 1). Results favored the vegan diet in the study by Peltonen for high disease activity improvement, and results favored the vegan diet in the study by Hafstrom (Table 2) for those achieving ACR 20. Certainty of evidence for these outcomes was very low, due to an indirect outcome and wide confidence intervals. Helve and colleagues (1998) also found improvements in pain, joint swelling, morning stiffness, and ability to move for those consuming an uncooked vegan diet prepared in a specialized kitchen. This evidence is reported in a word table, as no means or medians were reported for these measures.

Mediterranean diet versus no change in diet

There were 4 studies that explored the effects of the Mediterranean diet on patients with RA. Skoldstam (2003) and Hagfors (2005) (Table 3) found significant reductions in the HAQ, swollen joint count, global VAS, Pain VAS, and morning stiffness, as well as improvement in the SF-36 subsets of physical and social functioning. The certainty of evidence for these measures was Low. Garcia-Morales (2020) (Tables 4, 5) found the combination of MD + DEP showed more significant improvements in health-related quality of life on a global score in the SF-36 than either intervention alone. Certainty of evidence for this study was low. For some outcomes, effect sizes could not be computed and so the data for those appear in a separate table. Pineda-Juarez (2020) (Table 4) also compared groups utilizing a dynamic exercise program (DEP)+ a Mediterranean diet (MD), just DEP, and just MD, and found that the DEP+ MD group had a significant decrease in their HAQ score, indicating improvement in disability, when compared to the MD and DEP groups alone. Both study time periods were for 24 months.

Anti-Inflammatory diet (ADIRA) versus no change in diet

There were four studies in this comparison group (Vadell, 2020; Turesson Wadell, 2021; Adam, 2003; Ghaseminasab-Parizi, 2022). Adam and colleagues (Table 7) found that the anti-inflammatory diet group had significant improvement in pain in comparison to the control group. The level of certainty was low due to a high risk of selection, performance, detection, and reporting bias. Vadell, Turesson Wadell, and Ghaseminasab-Parizi (Table 6) found no statistically significant differences between the anti-inflammatory diet in rheumatoid arthritis (ADIRA) and control group for disease activity, HAQ, or quality of life. Certainty of evidence was very low due to wide confidence intervals, small sample size, and single-blind study.

Exclusion/Elimination diet versus no change in diet

There were three studies in this comparison (Darlington, 1986; Gianfranceschi, 1996; Guagnano, 2021). Darlington and Gianfraneschi (Tables 8, 9) found those on an exclusion/elimination diet demonstrated significant improvements in painful joints, pain during the day, pain during 24 hours, and morning stiffness in comparison to the control groups. Level of certainty was very low due to no participant blinding, wide confidence intervals, and small sample sizes. The third study (Guagnano, 2021) (Table 9) reported only medians, showing less pain and increased SF-36 scores after 3 months on this type of diet, compared to no change in diet.

Fasting versus no change in diet and Fasting vs Ketogenic diet

Fraser and colleagues (2000) (Table 11)), in a non-blinded, non-randomized study, compared a 1 week Fasting + 2 weeks of a lacto-vegetarian diet with a 1 week Ketogenic diet + 2 weeks of a lacto-vegetarian diet. After the full 3 weeks, the fasting group demonstrated significant improvement in disease activity. Siddique (2020) (Table 10) found improvements in both the fasting and non-fasting groups in their study, but the magnitude of the difference was larger for the fasting group. The groups were determined by religious belief in fasting, therefore were not randomized, and the certainty of evidence was low.

Elemental peptide diet versus no change in diet

This comparison only had one study (Pfeiffer, 1998) (Tables 12, 13). The elemental peptide diet group had significant improvement in the ACR20 response in comparison to the control group. The level of certainty was very low due to wide confidence intervals and small sample size.

Graastener diet versus no change in diet

There was only one study (Hansen 1996) (Table 14) in this comparison group. The only significant results were for the critical outcome of fewer swollen joints, which favored those who consumed the Graastener Diet. The Graastener diet was composed of an energy intake adjusted so as to obtain near-standard BMI (body mass index), with lower fat and increased protein (vegetable and fish) consumption, plus supplements of vitamins A, C, E and selenium. Certainty of evidence was low for this single study.

High sat fat/low unsat fat/hypoallergenic diet versus well-controlled diet

There was one study included in this comparison (Sarzi-Puttini 2000) (Table 15). There was no significant difference between high saturated fats/low unsaturated fat/hypoallergenic diet versus a well-controlled diet for the critical outcomes of disease activity and pain. Certainty of evidence was moderate.

"Arthritis diet" versus no change in diet

There was only one study included in this comparison (Panush 1983) Tables 16, 17). There was no significant difference in the measure of "improvement" between the arthritis diet group versus control. Certainty of the evidence was low.

Low dose of food sensitivities versus no change in diet

This comparison had one study (Gianfranceschi, 1996) (Table 18). No significant differences were found between the low dose food sensitivities and control group in painful joints, morning stiffness, or swollen joints. Level of certainty was very low due to small sample size and because the diet may be difficult to administer in practice.

Table 1: Vegan diet vs no change in diet

Authors: Elkan 2008 (one year), Helve 1998 (12 weeks), Peltonen 1997 (4 weeks), Hafstrom 2001 (one year)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Vegan diet	Control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from CRP 4 weeks – 1 year												
2	randomised trials	not serious	not serious	Serious[e]	Serious[b]	none	49	48	-	SMD 0.19 higher (0.23 lower to 0.61 higher) On the scale of DAS-28, this corresponds to MD= 0.21, 95% CI 0.25 lower to 0.67 higher	⊕⊕◯◯ Low	CRITICAL Not statistically significant
Disease activity inferred from ESR 4 weeks – 1 year												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Vegan diet	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a,e]	serious[b]	none	19	20	-	MD 0.3 lower (13.76 lower to 13.16 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Disease activity inferred from Number of tender joints > 12 weeks

1	randomised trials	not serious	not serious	serious[a]	Serious[b]	none	19	20	-	MD 2.57 lower (5.37 lower to 0.23 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Disease activity inferred from Number of swollen joints 4 weeks – 1 year

1	randomised trials	not serious	not serious	serious[a]	Serious[b]	none	19	20	-	MD 0.61 lower (2.61 lower to 1.39 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Vegan diet	Control	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	not serious	not serious	not serious	Serious[b]	none	49	48	-	SMD 0.04 lower (0.44 lower to 0.36 higher) On the 0-3 scale of HAQ, this corresponds to MD= 0.02, 95% CI 0.24 lower to 0.19 higher	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Function: HAQ 4 weeks – 1 year

Disease activity inferred from Duration of morning stiffness 4 weeks – 1 year

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Vegan diet	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	19	20	-	MD 7.68 higher (20.41 lower to 35.77 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Pain at rest 4 weeks – 1 year

1	randomised trials	not serious	not serious	serious[a]	serious	none	19	20	-	MD 0.36 higher (11.08 lower to 11.8 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Pain in movement 4 weeks – 1 year

1	randomised trials	not serious	not serious	serious[a]	serious	none	19	20	-	MD 0.44 higher (11.14 lower to 12.02 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Vegan diet	Control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from High disease activity improvement 4 weeks – 1 year												
1	randomised trials	serious[c]	not serious	Serious[e]	serious[b]	none	5/22 (22.7%)	0/21 (0.0%)	RR 10.52 (0.62 to 179.27)	0 fewer per 1,000 (from 0 fewer to 0 fewer)	⊕⊕◯◯ Low	CRITICAL Statistically Significant Favors vegan diet
Achieved ACR20 4 weeks – 1 year												
1	randomised trials	serious[d]	not serious	serious[e]	serious[f]	none	12/35 (34.3%)	1/26 (3.8%)	RR 8.91 (1.24 to 64.30)	304 more per 1,000 (from 9 more to 1,000 more)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors vegan diet

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Vegan diet	Control	Relative (95% CI)	Absolute (95% CI)		
Disease activity: DAS28 4 weeks – 1 year												
1	randomised trials	not serious	not serious	not serious		none	30	28	-	MD 0.3 lower (4.91 lower to 4.31 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

CI: confidence interval; MD: mean difference; RR: risk ratio; SMD: standardised mean difference

Explanations

a. Indirect intervention in the Helve article - Experimental diet was prepared in a specialized kitchen - not a reproducible diet for most people.

b. Wide CI, on both sides of effect line

c. High performance and reporting bias, unknown selection bias

d. High performance, detection, and attrition bias

e. Indirect outcome

f. Wide CI

Table 2: Additional data on Vegan diet vs no change in diet, Hanninen, 2000; Helve 1998

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
276 Hanninen 2000	Controlled intervention	3 months	42 RA patients	The intervention in this study was living food (extreme uncooked vegan diet). The food was premade in a specialized kitchen.	There was no statistically significant difference between the intervention and control groups

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
6473, Helve, 1998	Randomized controlled trial	3 months	43 patients with RA were randomized into intervention (n=22) and control (n=21) groups. 39 patients	Intervention group: "living food" diet (uncooked vegan, rich in lactobacilli) prepared in a specialized kitchen. Participants were supervised and tutored daily to follow the diet. Control group: continued eating	**Data were collected immediately and 3 months after completion of the intervention on a 0-10 scale. No means or medians were reported, only total scores and p-values.** **During the intervention period** Rheumatic pains (n=42) were significantly different between the intervention and control group Swelling of joints (n=42) were significantly different between the intervention and control group

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
			completed the study	omnivorous diet, with no tutoring	Morning stiffness (n=42) were significantly different between the intervention and control group Ability to move (n=42) was not significantly different between the intervention and control group **After the intervention period** Rheumatic pains (n=42) were significantly different between the intervention and control group Swelling of joints (n=42) were significantly different between the intervention and control group Morning stiffness (n=42) were significantly different between the intervention and control group Ability to move (n=42) was not significantly different between the intervention and control group 6473 Summary of findings: Objective disease activity measures were not statistically different between groups. Subjective disease measures showed significant improvements in disease activity.

Table 3:

Mediterranean diet vs no change in diet

Authors: Skoldstam 2003, Hagfors 2005

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity: DAS28 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	25	-	MD 0.4 lower (1.15 lower to 0.35 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant
Function: HAQ score 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	25	-	MD 0.2 lower (0.48 lower to 0.08 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Disease activity inferred from Swollen Joint 12 weeks

1	randomised trials	not serious	not serious	Serious[b]	serious[a]	none	26	25	-	MD 2.3 lower (5.27 lower to 0.67 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Disease activity inferred from Tender joint 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	Serious[b]	serious[a]	none	26	25	-	MD 1.6 lower (4.78 lower to 1.58 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Pain VAS 0-100 12 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	25	-	MD 14 lower (23.63 lower to 4.37 lower)	⊕⊕⊕◯ Moderate	CRITICAL Statistically Significant Favors Med diet

Disease activity inferred from Morning Stiffness (min) 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	Serious[b]	serious[a]	none	26	25	-	MD 26 lower (58.08 lower to 6.08 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Function inferred from Signals of functional impairment SOFI 12 weeks

1	randomised trials	not serious	not serious	Serious[b]	serious[a]	none	26	25	-	MD 0.7 higher (2.77 lower to 4.17 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

SF36 Physical Role change 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious[b]	Very serious[a]	none	26	25	-	MD 27.3 higher (4.83 higher to 49.77 higher)	⊕⊕◯◯ Low	IMPORTANT Statistically Significant Favors Med diet Inferred from functional measure

Pain: SF36 Bodily pain change 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	26	25	-	MD 0.5 higher (11.72 lower to 12.72 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
SF36 General health change 12 weeks												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	26	25	-	MD 5 higher (5.19 lower to 15.19 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant
SF36 Vitality change 12 weeks												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	26	25	-	MD 7.1 higher (3.1 lower to 17.3 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant
SF36 social functioning change 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	26	25	-	MD 10.2 higher (0.18 lower to 20.58 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 Emotional role change 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	26	25	-	MD 7.6 higher (11.11 lower to 26.31 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 mental health change 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	26	25	-	MD 2.8 higher (5.31 lower to 10.91 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

Function inferred from Grip ability test 12 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	26	25	-	MD 1 lower (7.6 lower to 5.6 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

SF36 physical function 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	27	-	MD 0.2 higher (0.25 lower to 0.65 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

SF36 role physical 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	35	27	-	MD 0.28 higher (0.2 lower to 0.76 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 role emotional 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	35	27	-	MD 0.16 higher (0.39 lower to 0.71 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 vitality 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	35	27	-	MD 0.26 higher (0.18 lower to 0.7 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 mental health 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	35	27	-	MD 0.1 lower (0.52 lower to 0.32 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 social function 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	35	27	-	MD 0.01 higher (0.48 lower to 0.5 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 bodily pain 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Mediterranian diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	27	-	MD 0.28 higher (1.26 lower to 1.82 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

SF36 global health 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	35	27	-	MD 0.29 higher (0.17 lower to 0.75 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

CI: confidence interval; MD: mean difference

Explanations

a. Small sample size

b. indirect measure

Table 4: Additional data on Mediterranean vs no change in diet, Garcia-Morales, 2020; Pineda-Juarez 2020

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1720, Garcia-Morales, 2020	Randomized controlled clinical trial	24 weeks	144 female RA patients enrolled; 130 competed study	Four groups: Dynamic exercise program (DEP): twice weekly training sessions lasting 80-90 minutes comprised of 5 stages (warm-up, aerobic exercise, anaerobic exercise, recreational games, cool down) Mediterranean diet (MD): individualized diet prescribed according to basal energy expenditure. MD and control group received general physical activity recommendations DEP and control group received general nutritional recommendations	(Reported as median changes between baseline and 24 weeks) MD+DEP (n=32) -38 (-0.62 to 0) DEP (n=36) -0.25 (-0.50 to 0) MD (n=35) 0 (-0.31 to 0.18) Control (n=27) 0 (-0.25 to 0.25) 1720 Summary of findings: The combination of MD + DEP showed more significant improvements in

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					health-related quality of life than either intervention alone. Outcome measure used was the SF-36

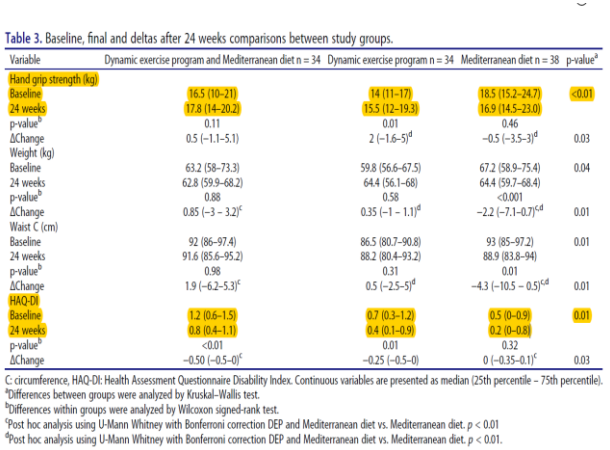
Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
803 Pineda-Juarez 2020	RCT, single blind	24 weeks	106 participants with active RA, all female and over 18 years old	DEP Group - Dynamic Exercise program 2x a week MD Group - prescribed a Mediterranean diet, received general physical activity recommendations DEP/MD Group - received both interventions	 **Table 3.** Baseline, final and deltas after 24 weeks comparisons between study groups. Variable | Dynamic exercise program and Mediterranean diet n = 34 | Dynamic exercise program n = 34 | Mediterranean diet n = 38 | p-value[a] Hand grip strength (kg) Baseline | 16.5 (10–21) | 14 (11–17) | 18.5 (15.2–24.7) | <0.01 24 weeks | 17.8 (14–20.2) | 15.5 (12–19.3) | 16.9 (14.5–23.0) | p-value[b] | 0.11 | 0.01 | 0.46 | ΔChange | 0.5 (−1.1–5.1) | 2 (−1.6–5)[d] | −0.5 (−3.5–3)[d] | 0.03 Weight (kg) Baseline | 63.2 (58–73.3) | 59.8 (56.6–67.5) | 67.2 (58.9–75.4) | 0.04 24 weeks | 62.8 (59.9–68.2) | 64.4 (56.1–68) | 64.4 (59.7–68.4) | p-value[b] | 0.88 | 0.58 | <0.001 | ΔChange | 0.85 (−3 – 3.2)[c] | 0.35 (−1 – 1.1)[d] | −2.2 (−7.1–0.7)[c,d] | 0.01 Waist C (cm) Baseline | 92 (86–97.4) | 86.5 (80.7–90.8) | 93 (85–97.2) | 0.01 24 weeks | 91.6 (85.6–95.2) | 88.2 (80.4–93.2) | 88.9 (83.8–94) | p-value[b] | 0.98 | 0.31 | 0.01 | ΔChange | 1.9 (−6.2–5.3)[c] | 0.5 (−2.5–5)[d] | −4.3 (−10.5 – 0.5)[c,d] | 0.01 HAQ-DI Baseline | 1.2 (0.6–1.5) | 0.7 (0.3–1.2) | 0.5 (0–0.9) | 0.01 24 weeks | 0.8 (0.4–1.1) | 0.4 (0.1–0.9) | 0.2 (0–0.8) | p-value[b] | <0.01 | 0.01 | 0.32 | ΔChange | −0.50 (−0.5–0)[c] | −0.25 (−0.5–0) | 0 (−0.35–0.1)[c] | 0.03 C: circumference, HAQ-DI: Health Assessment Questionnaire Disability Index. Continuous variables are presented as median (25th percentile – 75th percentile). [a]Differences between groups were analyzed by Kruskal–Wallis test. [b]Differences within groups were analyzed by Wilcoxon signed-rank test. [c]Post hoc analysis using U-Mann Whitney with Bonferroni correction DEP and Mediterranean diet vs. Mediterranean diet. $p < 0.01$ [d]Post hoc analysis using U-Mann Whitney with Bonferroni correction DEP and Mediterranean diet vs. Mediterranean diet. $p < 0.01$. There was a significant difference between before and after scores on the HAQ-DI for the DEP/MD group compared to the DEP group.

Table 5:

Dynamic exercise program + Mediterranean diet vs dynamic exercise program alone

Author: Garcia-Morales 2020

Certainty assessment	№ of patients	Effect	Certainty	Importance

№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Active dynamic exercise program + mediterranian diet	dynamic exercise programEP+MD	Relative (95% CI)	Absolute (95% CI)		
Pain: SF36 Bodily pain 24 weeks												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	32	36	-	MD 0.24 higher (0.25 lower to 0.73 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant
SF36 role physical 24 weeks												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	32	36	-	MD 0.21 lower (0.7 lower to 0.28 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant
SF36 role emotional 24 weeks												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	32	36	-	MD 0.06 lower	⊕⊕◯◯ Low	IMPORTANT

										(0.51 lower to 0.39 higher)		Not statistically significant

SF36 vitality 24 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	32	36	-	MD 0.33 lower (1.05 lower to 0.39 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 mental health 24 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	32	36	-	MD 0.28 lower (1.01 lower to 0.45 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 social function 24 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	32	36	-	MD 0.22 lower (0.71 lower to 0.27 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 global health 24 weeks

1	randomised trials	not serious	not serious	not serious	very serious[a]	none	32	36	-	MD 0.22 lower (0.71 lower to 0.27 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

CI: confidence interval; MD: mean difference

Explanations

a. indirect measure

b. small group size (n=36)

Table 6:

Anti-inflammatory diet vs no change in diet

Authors: Vadell 2020, Turesson Wadell 2021, Adam 2003

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
Swollen Joint Score (see note) (3 months)												
1	randomised trials	very serious[g]	not serious	serious[b]	serious[i]	none	30	30	-	MD **6.4 lower** (11.85 lower to 0.95 lower)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors Anti-inflam diet
Tender Joint Score (see note) (3 months)												
1	randomised trials	very serious[g]	not serious	serious[b]	serious[i]	none	30	30	-	MD **6 lower** (11.77 lower to 0.23 lower)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors Anti-inflam diet
Disease activity inferred from ACR20 Achieved (3 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[g]	not serious	serious[i]	serious[i]	none	13/34 (38.2%)	8/34 (23.5%)	**OR 2.01** (0.70 to 5.76)	**147 more per 1,000** (from 58 fewer to 404 more)	⨁◯◯◯ Very low	CRITICAL Not statistically significant

Pain VAS pain (0-100) short term (10 weeks)

2	Randomised trials	very serious[g]	very serious[h]	not serious	serious[i]	none	55	55	-	SMD **0.63 lower** (1.58 lower to 0.32 higher) This corresponds to MD=11.47, 95% CI 28.76 lower to 5.82 higher	⨁◯◯◯ Very low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
Function HAQ short term (10 weeks)												
1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 0.04 lower (0.17 lower to 0.09 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant
Pain :SF36 bodily pain short term (10 weeks)												
1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 1.5 higher (4.17 lower to 7.17 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		

SF36 physical functioning short term (10 weeks)

1	randomised trials	serious[e]	not serious	Not serious[b]	serious[f]	none	25	25	-	MD 5.39 higher (0.22 lower to 11.01 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

SF36 role-physical short term (10 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[e]	not serious	not serious	very serious[a]	none	25	25	-	MD 0.6 higher (7.4 lower to 8.6 higher)	⊕◯◯◯ Very low	IMPORTANT Not statistically significant

SF36 general health short term (10 weeks)

1	randomised trials	serious[e]	not serious	not serious	serious[a]	none	25	25	-	MD 3.16 lower (8.08 lower to 1.75 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		

SF36 physical component summary short term (10 weeks)

1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 0.02 higher (2.18 lower to 2.22 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

SF36 vitality short term (10 weeks)

1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 2.97 lower (10.05 lower to 4.1 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
SF36 social functioning short term (10 weeks)												
1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 0.58 lower (8.03 lower to 6.87 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant
SF36 role-emotional short term (10 weeks)												
1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 4.06 higher (3.19 lower to 11.3 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
SF36 mental health short term (10 weeks)												
1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 1.41 higher (4.03 lower to 6.84 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant
SF36 mental component summary short term (10 weeks)												
1	randomised trials	serious[e]	not serious	not serious	serious[f]	none	25	25	-	MD 0.34 higher (2.66 lower to 3.35 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from VAS fatigue (0-100) short term (10 weeks)												
1	randomised trials	serious[e]	not serious	serious[b]	serious[f]	none	25	25	-	MD 2.55 lower (11.67 lower to 6.56 higher)	⊕◯◯◯ Very low	IMPORTANT Not statistically significant
Disease activity inferred from VAS morning stiffness short term (10 weeks)												
1	randomised trials	serious[e]	not serious	serious[b]	serious[f]	none	25	25	-	MD 1.72 higher (5.77 lower to 9.21 higher)	⊕◯◯◯ Very low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from Morning stiffness (min) short term (10 weeks)												
1	randomised trials	serious[e]	not serious	serious[b]	serious[f]	none	25	25	-	MD 3.75 higher (9.52 lower to 17.02 higher)	⊕◯◯◯ Very low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect	Certainty		Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
HAQ-DI 12 weeks												
1	randomised trials	not serious	not serious	not serious	not serious	serious[b]	none	32	35	-	MD 0.17 higher (0.1 lower to 0.44 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not Statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
HAQ pain score change (mm VAS - 10 cm) 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	32	35	-	MD 8.8 higher (5.14 lower to 22.74 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not Statistically significant
Present Pain VAS change score (0 - 10) 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	32	35	-	MD 1.01 higher (0.04 lower to 2.06 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not Statistically significant
Disease activity inferred from Morning stiffness change (min) 12 weeks												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	32	35	-	MD 15.15 higher (8.98 lower to 39.28 higher)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant
Disease activity inferred from Disease feeling change? 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[c]	none	32	35	-	MD 0.07 higher (0.31 lower to 0.45 higher)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant

SF 36 general health change score 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 0.6 higher (10.44 lower to 11.64 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant

SF 36 physical functioning change score 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 10.8 lower (23.95 lower to 2.35 higher)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant

SF 36 physical rule limitation change score 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 14.6 higher (9.69 lower to 38.89 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant

SF 36 pain change score 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 14.4 lower (27.12 lower to 1.68 lower)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant

SF 36 physical health change score 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 1.5 lower (12.03 lower to 9.03 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant

SF 36 emotional well-being change score 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 1.9 lower (9.13 lower to 5.33 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant

SF 36 emotional role limitation change score 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 4.3 higher (20.49 lower to 29.09 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant

SF 36 vitality change score 12 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 3.8 lower (12.71 lower to 5.11 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant

SF 36 social functioning change score 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 7.3 lower (22.02 lower to 7.42 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant
SF 36 mental health change score 12 weeks												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	32	35	-	MD 2.2 lower (12.05 lower to 7.65 higher)	⊕⊕◯◯ Low	IMPORTANT Not Statistically significant
Disease activity inferred from ESR change score (mm/h) 12 weeks												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	32	35	-	MD 4.58 higher (3.01 lower to 12.17 higher)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant
Disease activity inferred from Rheumatoid factor change score (IU/ml) 12 weeks												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	32	35	-	MD 8.6 lower (16.9 lower to 0.3 lower)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Anti-inflammatory diet (ADIRA)	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from Anti- CCP change score (U/ml) 12 weeks												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	32	35	-	MD 12 higher (187.25 lower to 211.25 higher)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant
Disease activity inferred from C-reactive protein Change score (mg/L) 12 weeks												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	32	35	-	MD 2.39 lower (11.65 lower to 6.87 higher)	⊕⊕◯◯ Low	CRITICAL Not Statistically significant

CI: confidence interval; MD: mean difference; OR: odds ratio; SMD: standardised mean difference

Explanations

a. Single study with unclear ROB for 4/6 categories and 1/6 high ROB

b. Indirect measure

c. wide CI and small sample size (n=46)

d. small sample size (n=46)

e. single-blind study; unclear if researchers were blinded

f. small sample size (n=25)

g. High risk of selection, performance, detection, and reporting bias

h. high heterogeneity

i. small sample size

Note

Tender and swollen joint values are not joint counts, but rather composite scores where each join was weighted 0-3 for degree of tenderness or swelling

Table 7:

Exclusion/elimination Diet vs no change in diet

Authors: Darlington 1986 (6 weeks), Gianfranceschi 1996 (12 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Exclusion Diet	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity as inferred from Painful joints 6 - 12 weeks												
1	randomised trials	serious[a]	not serious	Serious[h]	serious[b]	none	45	42	-	MD 7.43 lower (12.53 lower to 2.33 lower)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors exclusion diet

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Exclusion Diet	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from Painful joints 6 - 12 weeks												
1	randomised trials	serious[c]	not serious	serious[d]	serious[e]	none	12	12	-	MD 3.05 lower (4.64 lower to 1.46 lower)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors exclusion diet
Pain during day 6 – 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	4/45 (8.9%)	12/42 (28.6 %)	OR 0.25 (0.07 to 0.90)	195 fewer per 1,000 (from 258 fewer to 21 fewer)	⊕⊕◯◯ Low	CRITICAL Statistically Significant Favors exclusion diet

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Exclusion Diet	control	Relative (95% CI)	Absolute (95% CI)		
Pain at night 6 - 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	2/45 (4.4%)	6/42 (14.3%)	OR 0.38 (0.08 to 1.91)	83 fewer per 1,000 (from 130 fewer to 99 more)	⊕⊕◯◯ Low	CRITICAL Not statistically significant
Pain during 24 hours VAS 6 - 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[f]	none	45	42	-	MD 2.06 lower (2.99 lower to 1.13 lower)	⊕⊕◯◯ Low	CRITICAL Statistically Significant Favors exclusion diet
Disease activity inferred from Morning stiffness (min) 6 - 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Exclusion Diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[c]	not serious	serious[d]	serious[g]	none	12	12	-	MD 40.8 lower (63.66 lower to 17.94 lower)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors exclusion diet

Function inferred from 20 yd walk time sec 6 - 12 weeks

1	randomised trials	serious[a]	not serious	serious[h]	serious[f]	none	21	21	-	MD 1.1 lower (2.96 lower to 0.76 higher)	⊕◯◯◯ Very low	CRITICAL Not statistically significant

Disease activity inferred from Swollen joints 6 - 12 weeks

1	randomised trials	serious[c]	not serious	serious[e]	serious[e]	none	12	12	-	MD 0.4 lower (1.7 lower to 0.9 higher)	⊕◯◯◯ Very low	CRITICAL Not statistically significant

CI: confidence interval; MD: mean difference; OR: odds ratio

Explanations

a. no blinding of participants and incomplete data reporting

b. wide CI and small sample size (n=49)

c. unclear ROB for 3/6 categories

d. intervention may be difficult to administer in practice

e. small sample size (n=12)

f. small sample size (n=49)

g. wide CI and small sample size (n=12)

h. indirect measure

Table 8:

Additional data on Exclusion/elimination vs no change in diet, Guagnano, 2021; Darlington, 1986

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
3118 Guagnano	RCT	3 months	40 RA patients	Exclusion diet (meat, gluten, lactose)	Pain VAS 0-100 median [IQR] Exclusion 40.4 [20.2 57.5] Balanced 48.7 [28.9 48.7]

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
2021					DAS28 median [IQR] Exclusion 2.5 [2 3] Balanced 2.5 [2 3] SF36 median [IQR] Exclusion 55 [33 60] Balanced 45.1 [42 49] HAQ median [IQR] Exclusion 1 [0.52 2] Balanced 1 [0.13 2]
8512 Darlinton 1986	RCT	6 weeks	49 RA patients	Elimination diet for 6 weeks with reintroduction of sensitive foods (e.g. gluten/dairy)	All numeric values reported were significant change from baseline, groups B and C underwent the same treatment 6weeks apart Duration morning stiffness Control 45 min Group B 10 min Group C 10 min

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					Grip strength R/L mmHg Control no change Group B 174 +- 84 / 168 +- 84 Group C no change 20yd walk time Group B no change wk 6

Table 9:

Fasting vs no change in diet

Author: Siddique 2020

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fasting	not fasting	Relative (95% CI)	Absolute (95% CI)		

Disease activity DAS28 4 weeks

1	observational studies	serious[a]	not serious	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect	120	120	-	MD 0.22 lower (0.38 lower to 0.06 lower)	⊕⊕◯◯ Low	CRITICAL Statistically Significant Both groups (fasting and no change in diet)) showed improvement

CI: confidence interval; MD: mean difference

Explanations

a. High risk of selection bias (was not RCT) and performance bias. Participants self selected their study group based on fasting religious beliefs.

Table 10:

Additional data on Fasting vs ketogenic diet, Fraser 2000

Notes: Data presented here as only *medians* (with 95% CI) were provided. Of note, are presenting timepoints at 1W after the diet and at 3W. I am presenting 1W data (despite the 2 week f/u rule) as this is the post-intervention timepoint and 3W is after 1 week of the intervention and 2 weeks of re-feeding.

Outcome-- Median, (95%CI)	1 W		3W	
	Fasting	Ketogenic	Fasting	Ketogenic
ESR	21 (10-48)	28 (16-40)	29 (15-52)	30 (18-62)

CRP	13 (7-33)	19 (9-56)	21 (10-53)	12 (5-44)
Tender joint count	10 (2-17)	8 (5-14)	15 (5-8)	10 (6-16)

Table 11:

Inactive Elemental Peptide Diet vs no change in diet

Author: Pfeiffer 1998

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Elemental Peptide Diet	control	Relative (95% CI)	Absolute (95% CI)		
Disease activity inferred from ACR20 response 4 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	1/15 (6.7%)	0/15 (0.0 %)	RR 3.00 (0.13 to 68.26)	0 fewer per 1,000 (from 0 fewer to 0 fewer)	⊕◯◯◯ Very low	CRITICAL Statistically Significant Favors peptide diet

CI: confidence interval; RR: risk ratio

Explanations

a. 1 high risk ROB and 2 unclear ROB

b. indirect measure

c. wide CI and small sample size (n=30)

Table 12:

Additional data on Elemental peptide diet vs no change in diet, Pfieffer, 1988

Notes: I have provided 2 timepoints- 4W which is right after the interventions, and 6 months which reflects 1 month of intervention and 5 months of washout period or normal eating in both groups. Data presented here because the data is presented as median (10/90 percentile).

Outcomes	4 Weeks (right after intervention)		6 months (4 W intervention, 5M normal eating)	
	Peptide Diet Median (10/90 percentile)	Control Median (10/90%)	Peptide Diet Median (10/90%)	Control Median (10/90%)
Ritchie articular index Peptide n=12 Control n=12	9.5 (3.9/27.9)	11.5 (4.6/32.2)	10.0 (5.3/16.4)	10.0 (3.6/23.0)
Number of swollen joints Peptide n=11 Control n=13	8.0 (3.6/11.6)	10.0 (5.2/19.0)	7.0 (5.0/12.0)	9.0 (3.4/23.6)

ESR Peptide n=12 Control n=14	22 (12/80)	53 (13/112)	40 (19/93)	47 (6/121)
Pain (now) Peptide n=13 Control n=13	3.0 (1.0/8.8)	4.0 (2.4/7.2)	4.0 (2.0/7.6)	4.0 (1.4/7.2)
Pain (average over last week) Peptide n=13 Control n=13	4.0 (1.4/6.6)	5.0 (2.4/8.0)	5.0 (2.0/7.6)	4.0 (1.4/7.6)
Pain (worst during last week) Peptide n=13 Control n=13	5.0 (1.0/7.5)	6.0 (3.0/8.0)	6.0 (2.4/9.6)	6.0 (1.8/9.6)
Duration of morning stiffness (levels of 15 min) Peptide n=13 Control n=14	2.0 (1.0/8.4)	5.0 (1.0/7.5)	3.0 (1.0/6.0)	2.5 (1.0/6.0)
HAQ Peptide n=13 Control n=14	0.88 (0.5/1.88)	1.32 (0.13/2.25)	1.00 (0.5/2.20)	1.19 (0.00/2.19)

General assessment of health on the day of examination Peptide n=13 Control n=14	3.0 (1.4/3.6)	2.0 (2.0/4.0)	3.0 (1.0/3.0)	2.0 (1.5/4.0)
General assessment of health, average during the last week Peptide n=13 Control n=14	3.0 (1.0/3.0)	2.0 (1.0/3.0)	2.0 (1.0/3.0)	2.0 (1.5/4.0)

Table 13:

Grasstener diet vs no change in diet

Author: Hansen 1996

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Grasstener diet	control	Relative (95% CI)	Absolute (95% CI)		

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Grasstener diet	control	Relative (95% CI)	Absolute (95% CI)		
HAQ change long term (6 months)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	36	45	-	MD 0.01 lower (0.23 lower to 0.21 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant
Disease activity as inferred from Duration of morning stiffness change (min) long term (6 months)												
1	randomised trials	not serious	not serious	Serious[b]	Serious[a]	none	36	45	-	MD 3 lower (23.47 lower to 17.47 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant
Disease activity as inferred from swollen joints change (1 - 3 scale) long term (6 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Grasstener diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	Serious[b]	Serious[a]	none	36	45	-	MD 2 lower (2.52 lower to 1.48 lower)	⊕⊕◯◯ Low	CRITICAL Statistically Significant Favors Grasstener Diet

Pain VAS scale change long term (6 months)

1	randomised trials	not serious	not serious	not serious	Serious[a]	none	36	45	-	MD 0.4 lower (0.89 lower to 0.09 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Disease activity as inferred from ESR change long term (6 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Grasstener diet	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	Serious[b]	Serious[a]	none	36	45	-	MD 1 higher (8.58 lower to 10.58 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Disease activity as inferred from physician global assessment change long term (6 months)

1	randomised trials	not serious	not serious	Serious[b]	Serious[a]	none	36	45	-	MD 0 (0.51 lower to 0.51 higher)	⊕⊕◯◯ Low	IMPORTANT Not statistically significant

CI: confidence interval; MD: mean difference

[a]Benefit and harm included – imprecision

[b]Surrogate measure – indirectness

Table 14: High SatFat/Low UnsatFat/Hypoallergenic vs Well-controlled Diet for RA

Author: Sarzi-Puttini 2000

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive High SatFat/Low UnsatFat/Hypoallergenic	Well-controlled Diet	Relative (95% CI)	Absolute (95% CI)		
Disease activity as inferred from Duration morning stiffness 24 weeks												
1	randomised trials	not serious	not serious	Serious[c]	serious[a]	none	21	22	-	MD 5.2 lower (27.51 lower to 17.11 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant
Disease activity as inferred from Ritchie's index 24 weeks												
1	randomised trials	not serious	not serious	serious	very serious[a]	none	21	22	-	MD 0.9 lower (3.39 lower to 1.59 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive High SatFat/Low UnsatFat/Hypoallergenic	Well-controlled Diet	Relative (95% CI)	Absolute (95% CI)		
Disease activity as inferred from Tender joint count 24 weeks												
1	randomised trials	not serious	not serious	Serious[c]	Serious[a]	none	21	22	-	MD 1.6 lower (3.86 lower to 0.66 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant
Disease activity as inferred from Swollen joint count 24 weeks												
1	randomised trials	not serious	not serious	Serious[c]	Serious[a]	none	21	22	-	MD 0.4 lower (1.99 lower to 1.19 higher)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive High SatFat/Low UnsatFat/Hypoallergenic	Well-controlled Diet	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS 24 weeks												
1	randomised trials	not serious	not serious	not serious	Serious[c]	none	21	22	-	MD 2.8 lower (13.21 lower to 7.61 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not statistically significant

CI: confidence interval; MD: mean difference

Explanations

a. Wide CI, on both sides of effect line
b. Contains benefit and harm – imprecision
c. Surrogate measure - indirectness

Table 15:

"Arthritis Diet" vs no change in diet

Author: Panush 1983

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive "Arthritis Diet"	Control	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from "Improvement" 10 weeks												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	5/11 (45.5%)	6/15 (40.0%)	RR 1.14 (0.46 to 2.78)	56 more per 1,000 (from 216 fewer to 712 more)	⊕⊕◯◯ Low	CRITICAL Not statistically significant

CI: confidence interval; RR: risk ratio

Explanations

a. Vague outcome - "improvement". Not sure how it was assessed from the mixture of outcomes in the word file - cannot be reproduced.

b. Wide CI on both sides of effect line

Table 16:

Additional data on “Arthritis” Diet vs no change in diet, Panush 1983

Notes: Data presented here as there are no confidence intervals, standard deviation, p-values, etc provided.

Outcomes	Arthritis diet (n=11)	Placebo diet (n=15)
Morning stiffness_10W (minutes)—mean	91	91
Grip strength_10W—mean	87	111
Walk time_10W – mean	14.7	14.8
Tender joints_10W—mean	23	17
Swollen joints_10W—mean	9	10
Patient assessment_10W—mean 1-5 scale, 5 excellent, 1 poor	3.1	2.7
Examiner assessment_10W—mean 1-5 scale, 5 excellent, 1 poor	3.4	3.0
ESR_10W	35	39

Table 17:

Low dose of food sensitivities vs no change in diet

Author: Gianfranceschi 1996

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Low dose of food sensitivities	control	Relative (95% CI)	Absolute (95% CI)		
Pain as inferred from number of Painful joints 12 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	12	12	-	MD 0.65 lower (1.9 lower to 0.6 higher)	⊕◯◯◯ Very low	CRITICAL Not statistically significant
Disease activity as inferred from Morning stiffness (min) 12 weeks												
1	randomised trials	serious[a,d]	not serious	serious[b]	serious[c]	none	12	12	-	MD 9.1 higher (14.32 lower to 32.52 higher)	⊕◯◯◯ Very low	CRITICAL Not statistically significant
Disease activity as inferred from Swollen joints 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inactive Low dose of food sensitivities	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	12	12	-	MD 0.4 lower (1.7 lower to 0.9 higher)	⊕◯◯◯ Very low	CRITICAL Not statistically significant

CI: confidence interval; MD: mean difference

Explanations

a. high risk for blinding of participants

b. diet may be difficult to administer in practice

c. sample size of 12 participants

References PICO 1

Adam, O., Beringer, C., Kless, T., Lemmen, C., Adam, A., Wiseman, M., . . . Forth, W. (2003). Anti-inflammatory effects of a low arachidonic acid diet and fish oil in patients with rheumatoid arthritis. *Rheumatol Int, 23*(1), 27-36. doi:10.1007/s00296-002-0234-7

Darlington, L. G., Ramsey, N. W., & Mansfield, J. R. (1986). Placebo-controlled, blind study of dietary manipulation therapy in rheumatoid arthritis. *Lancet, 1*(8475), 236-238. doi:10.1016/s0140-6736(86)90774-9

Elkan, A. C., Sjöberg, B., Kolsrud, B., Ringertz, B., Hafström, I., & Frostegård, J. (2008). Gluten-free vegan diet induces decreased LDL and oxidized LDL levels and raised atheroprotective natural antibodies against phosphorylcholine in patients with rheumatoid arthritis: a randomized study. *Arthritis Res Ther, 10*(2), R34. doi:10.1186/ar2388

Fraser, D. A., Thoen, J., Djøseland, O., Førre, O., & Kjeldsen-Kragh, J. (2000). Serum levels of interleukin-6 and dehydroepiandrosterone sulphate in response to either fasting or a ketogenic diet in rheumatoid arthritis patients. *Clin Exp Rheumatol, 18*(3), 357-362.

Garcia-Morales, J. M., Lozada-Mellado, M., Hinojosa-Azaola, A., Llorente, L., Ogata-Medel, M., Pineda-Juarez, J. A., . . . Castillo-Martinez, L. (2020). Effect of a Dynamic Exercise Program in Combination With Mediterranean Diet on Quality of Life in Women With Rheumatoid Arthritis. *J Clin Rheumatol, 26*(7S Suppl 2), S116-S122. doi:10.1097/RHU.0000000000001064

Ghaseminasab-Parizi, M., Nazarinia, M. A., & Akhlaghi, M. (2022). The effect of flaxseed with or without anti-inflammatory diet in patients with rheumatoid arthritis, a randomized controlled trial. *Eur J Nutr, 61*(3), 1377-1389. doi:10.1007/s00394-021-02707-9

Gianfranceschi, P., Fasani, G., & Speciani, A. F. (1996). Rheumatoid arthritis and the drop in tolerance to foods: elimination diets and the reestablishment of

tolerance by low-dose diluted food. *Ann N Y Acad Sci, 778*, 379-381. doi:10.1111/j.1749-6632.1996.tb21149.x

Guagnano, M. T., D'Angelo, C., Caniglia, D., Di Giovanni, P., Celletti, E., Sabatini, E., . . . Paganelli, R. (2021). Improvement of Inflammation and Pain after Three Months' Exclusion Diet in Rheumatoid Arthritis Patients. *Nutrients, 13*(10). doi:10.3390/nu13103535

Hafström, I., Ringertz, B., Spångberg, A., von Zweigbergk, L., Brannemark, S., Nylander, I., . . . Klareskog, L. (2001). A vegan diet free of gluten improves the signs and symptoms of rheumatoid arthritis: the effects on arthritis correlate with a reduction in antibodies to food antigens. *Rheumatology (Oxford), 40*(10), 1175-1179. doi:10.1093/rheumatology/40.10.1175

Hagfors, L., Nilsson, I., Sköldstam, L., & Johansson, G. (2005). Fat intake and composition of fatty acids in serum phospholipids in a randomized, controlled, Mediterranean dietary intervention study on patients with rheumatoid arthritis. *Nutr Metab (Lond), 2*, 26. doi:10.1186/1743-7075-2-26

Hänninen, Kaartinen, K., Rauma, A. L., Nenonen, M., Törrönen, R., Häkkinen, A. S., . . . Laakso, J. (2000). Antioxidants in vegan diet and rheumatic disorders. *Toxicology, 155*(1-3), 45-53. doi:10.1016/s0300-483x(00)00276-6

Hansen, G. V., Nielsen, L., Kluger, E., Thysen, M., Emmertsen, H., Stengaard-Pedersen, K., . . . Andersen, P. W. (1996). Nutritional status of Danish rheumatoid arthritis patients and effects of a diet adjusted in energy intake, fish-meal, and antioxidants. *Scand J Rheumatol, 25*(5), 325-330. doi:10.3109/03009749609104066

Holst-Jensen, S. E., Pfeiffer-Jensen, M., Monsrud, M., Tarp, U., Buus, A., Hessov, I., . . . Stengaard-Pedersen, K. (1998). Treatment of rheumatoid arthritis with a peptide diet: a randomized, controlled trial. *Scand J Rheumatol, 27*(5), 329-336. doi:10.1080/03009749850154339

Nenonen, M. T., Helve, T. A., Rauma, A. L., & Hänninen, O. O. (1998). Uncooked, lactobacilli-rich, vegan food and rheumatoid arthritis. *Br J Rheumatol, 37*(3), 274-281. doi:10.1093/rheumatology/37.3.274

Panush, R. S., Carter, R. L., Katz, P., Kowsari, B., Longley, S., & Finnie, S. (1983). Diet therapy for rheumatoid arthritis. *Arthritis Rheum, 26*(4), 462-471. doi:10.1002/art.1780260403

Peltonen, R., Nenonen, M., Helve, T., Hänninen, O., Toivanen, P., & Eerola, E. (1997). Faecal microbial flora and disease activity in rheumatoid arthritis during a vegan diet. *Br J Rheumatol, 36*(1), 64-68. doi:10.1093/rheumatology/36.1.64

Pineda-Juárez, J. A., Lozada-Mellado, M., Hinojosa-Azaola, A., García-Morales, J. M., Ogata-Medel, M., Llorente, L., . . . Castillo-Martínez, L. (2022). Changes in hand grip strength and body weight after a dynamic exercise program and Mediterranean diet in women with rheumatoid arthritis: a randomized clinical trial. *Physiother Theory Pract, 38*(4), 504-512. doi:10.1080/09593985.2020.1777605

Sarzi-Puttini, P., Comi, D., Boccassini, L., Muzzupappa, S., Turiel, M., Panni, B., & Salvaggio, A. (2000). Diet therapy for rheumatoid arthritis. A controlled double-blind study of two different dietary regimens. *Scand J Rheumatol, 29*(5), 302-307. doi:10.1080/030097400447688

Siddique, S., Imran, Y., Afzal, M. N., & Malik, U. (2020). Effect of Ramadan fasting on disease activity in patients with rheumatoid arthritis presenting in tertiary care hospital. *Pak J Med Sci, 36*(5), 1032-1035. doi:10.12669/pjms.36.5.2099

Sköldstam, L., Hagfors, L., & Johansson, G. (2003). An experimental study of a Mediterranean diet intervention for patients with rheumatoid arthritis. *Ann Rheum Dis, 62*(3), 208-214. doi:10.1136/ard.62.3.208

Turesson Wadell, A., Bärebring, L., Hulander, E., Gjertsson, I., Hagberg, L., Lindqvist, H. M., & Winkvist, A. (2021). Effects on health-related quality of life in the randomized, controlled crossover trial ADIRA (Anti-inflammatory Diet In Rheumatoid Arthritis). *PLoS One, 16*(10), e0258716. doi:10.1371/journal.pone.0258716

Vadell, A. K. E., Bärebring, L., Hulander, E., Gjertsson, I., Lindqvist, H. M., & Winkvist, A. (2020). Anti-inflammatory Diet In Rheumatoid Arthritis (ADIRA)-a randomized, controlled crossover trial indicating effects on disease activity. *Am J Clin Nutr, 111*(6), 1203-1213. doi:10.1093/ajcn/nqaa019

PICO 2: Should patients with RA use a commercially available dietary supplement?

Summary: Literature searches identified 37 randomized controlled trials (RCTs) and 1 observational study addressing this PICO question. Studies encompassed many different categories of supplementation approaches, with large variability in doses and formulations.

- Vitamin D (vs placebo, or usual care, or Calcitriol, or the effect of adding vitamin D to calcium carbonate)
    - Table 1. Vitamin D vs. Placebo (1-5)
    - Table 2. Additional data on Vitamin D vs. Placebo (1, 2)
    - Table 3. Vitamin D vs. Usual Care (6)
    - Table 4. Vitamin D vs. Calcitrol (4)
    - Table 5. Vitamin D + Calcium carbonate vs. Calcium carbonate (7)
- Selenium (vs placebo)
    - Table 6. Selenium vs. Placebo (8, 9)
    - Table 7. Additional data on Selenium vs. Placebo (8)
- Ginger (vs placebo)
    - Table 8. Ginger vs. Placebo (10)
- Probiotics
    - Table 9. Lactobacillus Rhamnosus vs Placebo (11)
    - Table 10. Lactobacillus Rhamnosus + Lactobacillus reuteri vs Placebo (12)
    - Table 11. Bacillus Coagulans vs Placebo (13)
    - Table 12. Additional data on probiotic supplementation (14, 15)
- Glucosamine (vs placebo)
    - Table 13. Glucosamine vs. Placebo (16)
    - Table 14. Additional data on Glucosamine vs. Placebo (16)
- Vitamin E (vs placebo)
    - Table 15. Vitamin E vs. Placebo (17)
- Conjugated linoleic acid + Vitamin E (vs placebo)
    - Table 16. Conjugated linoleic acid + Vitamin E vs. Placebo (17)
- Fatty acid vs. Placebo
    - Omega 6 (vs placebo)
        - Table 17. Borage oil (Omega 6 fatty acid) vs. Placebo (18)
        - Table 18. Conjugated linoleic acid (Omega 6 fatty acid) vs. Placebo (17)

    - Table 45. Additional data on Fish oil vs. usual diet (41)
    - Table 46. Fatty acid vs. fatty acid + g-linolenic acid (30)
    - Table 47. Flaxseed oil vs. Safflower oil (42)
    - Table 48. Flaxseed vs. Wheat (43)
    - Table 49. Primrose oil versus stinging nettle(19)
- N-actylcysteine vs placebo
    - Table 50. N-actylecysteine vs. placebo (44)
- Stinging nettle versus placebo
    - Table 51. Stinging nettle versus placebo(19)

Below, we separately discuss the evidence in each of these categories, along with separate certainty of evidence grades.

Vitamin D

Comparison: Vitamin D vs. Placebo

Evidence Summary: Four randomized control trials and one non-radnomized interventional trial looked at the use of vitamin D vs placebo in patients with RA. Li et al looked at 246 patients after 6 weeks and found a lower relative risk of having 9-13 swollen joints in the vitamin D group vs. a lower relative risk of 4-8 swollen joints; however, there was no difference between the groups for 14+ swollen joints. Other surrogates of disease activity (morning stiffness and CRP) were lower in the vitamin D group, although there was no difference in ESR. Salesi et al looked at 117 patients at 12 weeks and found a lower tender joint count in those on vitamin D, but no difference in swollen joint count or DAS 28, and a higher pain VAS score in those on vitamin D. Chawla et al also found no difference in pain scores. Soubrier et al and Yang et al followed patients with vitamin D deficiency and RA. Soubrier found no difference in HAQ, RAPID3, SF36, pain, fatigue, or activity at 6 months, although the vitamin D group did have a lower ESR and CRP at 6 months. Yang found that there was no difference in flare rate between those on vitamin D and those on placebo. Overall, while there are some surrogate markers of improvement in disease activity, there is no improvement in direct measures such as DAS28 and worsened pain scores. There is no evidence of improvement in functional status.

Quality of evidence: Very Low

Table 1. Vitamin D vs. Placebo (1-5)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	Placebo	Relative (95% CI)	Absolute (95% CI)		
Tender joint count, 12 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	60	57	-	MD **2.1 lower** (3.88 lower to 0.32 lower)	⨁⨁◯◯ Low	CRITICAL **Tender joint count significantly lower in the Vitamin D group.**
Swollen joint count, 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	60	57	-	MD **1 lower** (2.08 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL No difference in swollen joint count.

DAS28, 12 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	60	57	-	MD **0.5 lower** (1.12 lower to 0.12 higher)	⨁⨁◯◯ Low	CRITICAL No difference in DAS28.

Pain VAS, 12 weeks

2	randomised trials	Very serious[d]	serious	not serious	Very serious[c]	none	60	57	-	MD **0.1 higher** (0.99 lower to 1.2 higher)	⨁◯◯◯ Very low	CRITICAL No difference in pain VAS at 12 weeks.

ESR, 12 weeks

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	60	57	-	SMD **0.08 lower** (0.44 lower to 0.28 higher)	⨁◯◯◯ Very low	CRITICAL No difference in ESR at 12 weeks.

Change in HAQ, 6 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	29	30	-	MD **0.11 lower** (0.23 lower to 0.01 higher)	⨁◯◯◯ Very low	CRITICAL No difference in HAQ at 6 months.

Number of Patients with Flares, 2 years

1	randomised trials	very serious[d]	not serious	serious[b]	serious[c]	none	16/84 (19.0%)	26/88 (29.5%)	**RR 0.64** (0.37 to 1.11)	**106 fewer per 1,000** (from 186 fewer to 33 more)	⨁◯◯◯ Very low	CRITICAL No difference in proportion of patients flaring at 2 years.

Proportion with 4 - 8 swollen joints, 6 weeks

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	45/123 (36.6%)	12/123 (9.8%)	**RR 3.75** (2.09 to 6.74)	**268 more per 1,000** (from 106 more to 560 more)	⨁⨁◯◯ Low	CRITICAL **Significantly higher risk of having 4-8 swollen joints at 6 weeks in the vitamin D group.**

Proportion with 9 - 13 swollen joints, 6 weeks

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	45/123 (36.6%)	65/123 (52.8%)	**RR 0.69** (0.52 to 0.92)	**164 fewer per 1,000** (from 254 fewer to 42 fewer)	⨁⨁◯◯ Low	CRITICAL **Significantly higher risk of having 9-13 swollen joints at 6 weeks in the placebo group.**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	Placebo	Relative (95% CI)	Absolute (95% CI)		
Proportion with 14+ swollen joints, 6 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	33/123 (26.8%)	46/123 (37.4%)	**RR 0.72** (0.50 to 1.04)	**105 fewer per 1,000** (from 187 fewer to 15 more)	⨁◯◯◯ Very low	CRITICAL No difference in the risk of having 14+ swollen joints.
Duration Morning Stiffness in minutes, 6 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	123	123	-	MD **15 lower** (19.01 lower to 10.99 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly fewer minutes of morning stiffness for those on vitamin D at 6 weeks.**
CRP, 6 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	123	123	-	MD **0.18 lower** (0.31 lower to 0.05 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower CRP in the vitamin D group at 6 weeks.**
HAQ - Disease Activity Subscale, 6 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	246	246	-	MD **0.05 lower** (0.1 lower to 0.01 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower HAQ (disease activity subscale) in the vitamin D group at 6 weeks.**

Pain VAS, 6 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	123	123	-	MD **0.11 higher** (0.06 lower to 0.28 higher)	⨁⨁◯◯ Low	CRITICAL No difference in pain VAS at 6 weeks.

ESR, 6 weeks

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	123	123	-	MD **0.28 higher** (0.02 higher to 0.53 higher)	⨁⨁◯◯ Low	CRITICAL **Significantly higher ESR in the vitamin D group at 6 weeks.**

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

a. Many unclear risk of bias categorizations

b. Indirect measure of disease activity

c. Crosses no effect threshold

d. 2 types of bias flagged as high risk- both blinding categories

Table 2. Additional data on Vitamin D vs. Placebo (1, 2)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
5123, Soubrier, 2018	RCT	6 months	Total n: 59 patients w RA and vitamin D deficiency Mean age: 59.8 y +/- 10.9; 83.1% female; mean disease duration: 17.0 y +/- 9.7	Intervention: vitamin D (cholecalciferol 100,000 IU) - for 24 weeks Control: placebo	After adjusting for age, gender, season, initial vitamin D status, improvements in the intervention group were observed at 6-months follow up for: · ESR (p = 0.002) · CRP (p=0.04) · DAS-28-ESR (p value was not reported, but it said it was not significant) No differences at follow up for (no p values reported): · RAPID Score · SF36 · VAS Pain · VAS Fatigue · VAS Activity Patient Global Assessment
7366, Salesi, 2012	RCT	12 weeks	117 active RA patients	Vit D supplementation (50,000 IU weekly) vs placebo	Primary outcome was proportion of patients with a 0.6 point improvement in DAS28 after 12 weeks: Odds ratio for Vit D Supp: 2.1, 95% CI 0.77-6.2, p-value=0.139

Comparison: Vitamin D vs. usual care

Evidence Summary: One single-center retrospective cohort study (Wu et al 2020) evaluated Vitamin D versus usual care in a cohort of 1180 patients with RA treated with csDMARDs only. The study was designed as a retrospective chart review of individuals with RA treated at a single medical center. Individuals were followed for 4 months, with the time-zero point starting at the date of first Vitamin D supplementation. Controls were selected from patients who never received Vitamin D, although the selection process for controls was not clearly described. The study found significant improvements in HAQ and swollen joint count in patients who received Vitamin D compared to those who did not, however there was no difference in DAS28 response or tender joint count. There are also significant concerns about attrition and control selection.

Quality of Evidence: Very Low

Table 3. Vitamin D vs. Usual Care (6)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vitamin D Supplementation	the Control Group (usual DMARD medication), OS	Relative (95% CI)	Absolute (95% CI)		
DAS Response (Good versus Non-Response)												
1	observational studies	serious[a]	not serious	not serious	serious[b]	none	54/263 (20.5%)	28/141 (19.9%)	RR 1.03 (0.69 to 1.55)	6 more per 1,000 (from 62 fewer to 109 more)	⨁◯◯◯ Very low	CRITICAL No significant difference
HAQ												

1	observational studies	serious[a]	not serious	not serious	not serious	none	263	141	-	MD 0.9 lower (1.54 lower to 0.26 lower)	⊕◯◯◯ Very low	CRITICAL Significant difference in favor of Vitamin D

SJC28

1	observational studies	serious[a]	not serious	not serious	not serious	none	263	141	-	MD 1.31 lower (2.22 lower to 0.4 lower)	⊕◯◯◯ Very low	CRITICAL Significant difference in favor of Vitamin D

TJC28

1	observational studies	serious[a]	not serious	not serious	not serious	none	263	141	-	MD 0.6 lower (1.74 lower to 0.54 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference

CI: confidence interval; MD: mean difference; RR: risk ratio

a. Very high attrition (50%) and unclear selection process for controls

b. Wide CI that crosses 1

Comparison: Vitamin D vs. Calcitriol

Evidence summary: One RTC (Li et al) looked at the difference between vitamin D supplementation and calcitriol. There was no difference in relative risk of number of swollen joints. Duration of morning stiffness and CRP was higher in those on Vitamin D compared to calcitriol.

Quality of evidence: Very low

Table 4. Vitamin D vs. Calcitrol (4)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	calcitrol	Relative (95% CI)	Absolute (95% CI)		
9 - 13 swollen joints, 6 weeks												
1	randomised trials	not serious	not serious	serious[a]	very serious[b]	none	45/123 (36.6%)	39/123 (31.7%)	**RR 1.15** (0.81 to 1.63)	**48 more per 1,000** (from 60 fewer to 200 more)	⨁◯◯◯ Very low	CRITICAL No difference in the risk of having 9-13 swollen joints.
4 - 8 swollen joints, 6 weeks												
1	randomised trials	not serious	not serious	serious[a]	very serious[b]	none	45/123 (36.6%)	50/123 (40.7%)	**RR 0.90** (0.66 to 1.23)	**41 fewer per 1,000** (from 138 fewer to 93 more)	⨁◯◯◯ Very low	CRITICAL No difference in the risk of having 4-8 swollen joints.
14+ swollen joints, 6 weeks												
1	randomised trials	not serious	not serious	serious[a]	very serious[b]	none	33/123 (26.8%)	35/123 (28.5%)	**RR 0.94** (0.63 to 1.41)	**17 fewer per 1,000** (from 105 fewer to 117 more)	⨁◯◯◯ Very low	CRITICAL No difference in the risk of having 14+ swollen joints.
Duration Morning Stiffness in Minutes, 6 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	22-oxa-calcitrol (Vit D)	calcitrol	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	not serious	none	123	123	-	MD **10 higher** (6.37 higher to 13.63 higher)	⨁⨁◯◯ Low	CRITICAL **Significantly longer duration of morning stiffness in those on vitamin D compared to calcitriol.**

CRP, 6 weeks

1	randomised trials	not serious	not serious	serious[a]	not serious	none	123	123	-	MD **0.09 higher** (0.06 higher to 0.12 higher)	⨁⨁◯◯ Low	CRITICAL **CRP higher on those on vitamin D compared to calcitriol.**

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio
a. Surrogate marker for disease activity

b. Crosses no effect threshold

Comparison: Vitamin D + Calcium carbonate vs. Calcium carbonate

Evidence summary: One RCT examined the combination of Vitamin D with calcium compared to calcium alone. The study has a very high risk of bias due to no mention of blinding and unclear methods for randomization and allocation concealment. There was a significant reduction in VAS Pain score at the end of 3 months in the vitamin D group compared to the calcium group. There was not a significant difference between time to achieve pain relief between vitamin D and control groups or reduction in VAS scores at the onset of pain relief between vitamin D and control groups.

Quality of evidence: Very Low

Table 5. Vitamin D + Calcium carbonate vs. Calcium carbonate (7)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
6021, Gopinath, 2011	open-labeled randomized controlled trial	3 months	110 patients newly diagnosed with RA who had not previously been treated with anything other analgesics completed the study. 55 patients in each arm.	Treatment: combination of 500 IU vitamin D and 1,000 mg calcium carbonate Control: 1,000 mg calcium carbonate Both groups: 10 mg single dose of methotrexate once per week, 5 mg folic acid twice per week, sulphasalazine (escalating doses up to 1 g twice daily), hydroxychloroquine (200 mg/day), and naproxen (275 mg twice daily)	Primary outcome: Time to achieve pain relief for the first time (median days) Vitamin D group (n=59) = 21 days (range 7-90) Calcium group (n=62) = 21 days (range 7-90) p-value = 0.415 Secondary outcomes: Reduction in VAS score at the onset of pain relief (median %) Vitamin D group (n=59) = 10 (range 0-30) Calcium group (n=62) = 10 (range 0-50) p-value = 0.150 Reduction in VAS score at the end of 3 months (median %)

					Vitamin D group (n=59) = 50 (range 0-100) **Calcium group (n=62) = 30 (range 0-100)** **p-value = 0.006 – statistically significant**

Selenium

Comparison: Selenium vs. Placebo

Evidence Summary: Two RCTs (Tarp 1985, Peretz 2001) evaluated selenium supplementation compared to placebo over 3 months (Peretz) or 6 months (Tarp). Only swollen joints, morning stiffness, and pain VAS were reported in both studies, with Ritchie index, grip strength, pain relief, motion limitation, fatigue onset, and number of painful joints being reported in one study. There was no significant difference between selenium and placebo in any of the outcome measures above.

Quality of Evidence: Moderate

Table 6. Selenium vs. Placebo (8, 9)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Selenium supplementation	placebo	Relative (95% CI)	Absolute (95% CI)		

Articular index (Ritchie modified)

1	randomised trials	not serious	not serious	not serious	serious[a]	none	20	20	-	SMD 0.14 lower (0.76 lower to 0.48 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

Number of joints with limitation of motion

1	randomised trials	not serious	not serious	serious[b]	very serious[c]	none	20	20	-	MD 3.2 higher (0.12 lower to 6.52 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference

Number of swollen joints

2	randomised trials	not serious	not serious	not serious	not serious	none	48	47	-	MD 0.02 higher (1.49 lower to 1.52 higher)	⊕⊕⊕⊕ High	CRITICAL No significant difference

Grip strength (mmHg)

1	randomised trials	not serious	not serious	serious[d]	serious[a]	none	20	20	-	MD 32 lower (88 lower to 24 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Ring size of PIP joints (mm)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	20	20	-	MD 4 lower (37.67 lower to 29.67 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Pain (VAS)

2	randomised trials	not serious	not serious	not serious	not serious	none	48	47	-	MD 0.81 higher (0.97 lower to 2.6 higher)	⊕⊕⊕⊕ High	CRITICAL No significant difference

Pain relief (VAS)

1	randomised trials	not serious	not serious	not serious	serious[a]	none	20	20	-	MD 0.4 lower (2.26 lower to 1.46 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

Morning stiffness (hours)

2	randomised trials	not serious	not serious	serious[b]	serious[a]	none	48	47	-	MD 0.6 lower (2.17 lower to 0.97 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Time until onset of fatique (hours)

1	randomised trials	not serious	not serious	serious[d]	serious[a]	none	20	20	-	MD 0.9 lower (2.67 lower to	⊕⊕◯◯ Low	CRITICAL No significant difference

										0.87 higher)		

Number of painful joints

1	randomised trials	not serious	not serious	not serious	serious[a]	none	28	27	-	MD 1 lower (3.71 lower to 1.71 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

CI: confidence interval; MD: mean difference; SMD: standardised mean difference

a. Wide CI that crosses 0

b. Indirect measure of disease activity

c. Wide CI that crosses 0 and high effect threshold

d. Indirect measure of functional status

Table 7. Additional data on Selenium vs. Placebo (8)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results

<table>
<tr>
<td>5751 Peretz 2001</td>
<td>RCT</td>
<td>90 days</td>
<td>RA patients = 55

Selenium supplementation n = 28, Age: 61 ± 13, Male: 7, Female: 21

Placebo n = 27, Age: 60 ± 13,

Male: 7, Female: 20</td>
<td>Selenium group: 200mg (2 ´ 100 mg/d) selenium-enriched yeast capsules

Control group: placebo of identical aspect</td>
<td>
<table>
<tr><th>Morning stiffness</th><th>Median</th><th>Range</th><th>P value</th></tr>
<tr><td>Selenium</td><td>60</td><td>0 - 480</td><td>NS</td></tr>
<tr><td>Control</td><td>60</td><td>0 – 360</td><td><0.01</td></tr>
</table>
<table>
<tr><th>Arm movements</th><th>Mean</th><th>P value – btw groups</th></tr>
<tr><td>Selenium</td><td>1.4</td><td rowspan="2"><0.005</td></tr>
<tr><td>Control</td><td>2.9</td></tr>
</table>
</td>
</tr>
</table>

Ginger

Comparison: Ginger vs. Placebo

Evidence Summary: One double-blind RCT (Aryaeian 2019) compared supplementation with 1500mg of ginger to placebo. 33 patients were assigned to ginger and 30 to placebo. There was no significant difference in the primary outcome of DAS28-ESR at 12 weeks. The study suffered from some imprecision.

Quality of Evidence: Moderate

Table 8. Ginger vs. Placebo (10)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ginger	control	Relative (95% CI)	Absolute (95% CI)		
DAS28, 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	33	30	-	MD **0.86 lower** (1.73 lower to 0.01 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

a. Wide CI that crosses 0

Probiotics

Comparison: Probiotic supplementation vs. Placebo

Evidence Summary: Five RCTs examined probiotic supplementation in patients with RA. Hatakka et al examined the use of lactobacillus rhamnosus vs placebo in a small trial of 26 patients with RA for 12 months. The study was limited by high (>25%) drop out rates and there was no difference in functional status (HAQ) or RA disease activity as assessed by swollen and tender joint counts in either treatment arm. De los Angeles Pineda et al examined supplementation with lactobacillus rhamnosus and lactobacillus reuteri for 3 months in 29 patients with RA and found no significant difference in the proportion of patients reaching ACR20 criteria between treatment arms, though mean reduction in DAS was slightly greater (0.8 pts 95% CI 0.16-1.44) in the probiotic treatment arm. Mandel et al examined the effect bacillus coagulans vs placebo in 45 patients with RA. While there was a slight reduction in pain scores in the treatment arm vs placebo (16/100 pts lower, 95% CI 4.05-27.95), there was no difference in the proportion of patients meeting an ACR20 response and no significant difference in the DAS, TJC, or SJC at the end of the study. Additionally, there were no observed differences in several measures of functional status between treatment arms. Cannarella et al found no difference in DAS28 with supplementing a probiotic containing lactobacillus acidophilus, Lactobacillus casei, Lactococcus lactis, Bifidobacterium lactis, and and B. bifidum. Vaghef-Mehrabany et al observed a significant decrease in the DAS28, TJC, and SJC with L. Casei supplementation, though this study was limited by high drop out rate (25%) and selective reporting of data in the trial publication.

Quality of evidence: Low to very low

Table 9. Lactobacillus Rhamnosus vs Placebo (11)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Probiotic supplement (Lactobacillus Rhamnosus)	placebo for 12 months	Relative (95% CI)	Absolute (95% CI)		
HAQ-index												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	8	13	-	MD 0.2 lower (0.67 lower to 0.27 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Swollen Joint Count												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	8	13	-	MD 0.1 lower (2.16 lower to 1.96 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Tender Joint Count												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	8	13	-	MD 0.1 lower (1.86 lower to 1.66 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference

CI: confidence interval; MD: mean difference

a. High drop out

b. Single study

c. CI crosses zero

Table 10. Lactobacillus Rhamnosus + Lactobacillus reuteri vs Placebo (12)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Probiotic Supplement (Lactobacillus rhamnosus + Lactobacillus reuteri)	Placebo for 3 months	Relative (95% CI)	Absolute (95% CI)		
ACR20 Criteria, 3 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	3/15 (20.0%)	1/14 (7.1%)	RR 2.80 (0.33 to 23.86)	129 more per 1,000 (from 48 fewer to 1,000 more)	⨁⨁◯◯ Low	CRITICAL No sig difference in proportion of patients meeting ACR criteria
Change in HAQ score, 3 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	15	14	-	MD 0.13 higher (0.02 lower to 0.28 higher)	⨁⨁◯◯ Low	CRITICAL No sig difference
DAS change												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	15	14	-	MD 0.8 higher (0.16 higher to 1.44 higher)	⨁⨁⨁◯ Moderate	CRITICAL Greater reduction in DAS in probiotic group

SJC change

1	randomised trials	not serious	not serious	not serious	serious[a]	none	15	14	-	MD 0.6 higher (1.92 lower to 3.12 higher)	⊕⊕⊕◯ Moderate	CRITICAL No sig difference

TJC change

1	randomised trials	not serious	not serious	not serious	serious[a]	none	15	14	-	MD 0.75 higher (3.9 lower to 5.4 higher)	⊕⊕⊕◯ Moderate	CRITICAL No sig difference

Phy global change

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	15	14	-	MD 0.6 higher (0.4 lower to 1.6 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

Pt global change

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	15	14	-	MD 0.37 lower (1.44 lower to 0.7 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

Morning stiffness change

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	15	14	-	MD 0.95 higher (32 lower to 33.9 higher)	⊕⊕◯◯ Low	IMPORTANT No sig difference

Pain change

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	15	14	-	MD 0.27 lower (1.81 lower to 1.27 higher)	⨁⨁◯◯ Low	CRITICAL No sig difference

Fatigue change

1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	15	14	-	MD 1.72 lower (3.13 lower to 0.31 lower)	⨁⨁◯◯ Low	CRITICAL Lower level of fatigue in probiotic group

CI: confidence interval; MD: mean difference; RR: risk ratio

a. Single study

b. Wide CI, crosses zero

c. Functional status surrogate

Table 11. Bacillus Coagulans vs Placebo (13)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Probiotic supplementation (bacillus coagulans)	placebo for 2 months	Relative (95% CI)	Absolute (95% CI)		
Individual Function: Improvement in Arising at 2 months												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	3/22 (13.6%)	4/21 (19.0%)	RR 0.72 (0.18 to 2.82)	53 fewer per 1,000 (from 156 fewer to 347 more)	⨁◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement inWalking 2 miles at 2 months

1	randomised trials	very serious [a]	not serious	not serious	very serious[b,c]	none	2/22 (9.1%)	7/22 (31.8%)	RR 0.29 (0.07 to 1.23)	226 fewer per 1,000 (from 296 fewer to 73 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement in Daily Activities at 2 months

1	randomised trials	very serious [a]	not serious	not serious	very serious[b,c]	none	2/22 (9.1%)	4/22 (18.2%)	RR 0.50 (0.10 to 2.45)	91 fewer per 1,000 (from 164 fewer to 264 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement in dressing and grooming at 2 months

1	randomised trials	very serious [a]	not serious	not serious	very serious[b,c]	none	4/22 (18.2%)	4/22 (18.2%)	RR 1.00 (0.29 to 3.50)	0 fewer per 1,000 (from 129 fewer to 455 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement in eating at 2 months

1	randomised trials	very serious [a]	not serious	not serious	very serious[b,c]	none	6/22 (27.3%)	4/22 (18.2%)	RR 1.50 (0.49 to 4.59)	91 more per 1,000 (from 93 fewer to 653 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement in Hygiene at 2 months

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	2/22 (9.1%)	2/22 (9.1%)	RR 1.00 (0.15 to 6.48)	0 fewer per 1,000 (from 77 fewer to 498 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement in Reach at 2 months

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	4/22 (18.2%)	9/22 (40.9%)	RR 0.44 (0.16 to 1.23)	229 fewer per 1,000 (from 344 fewer to 94 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Individual Function: Improvement in Grip at 2 months

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	5/22 (22.7%)	4/22 (18.2%)	RR 1.25 (0.39 to 4.05)	45 more per 1,000 (from 111 fewer to 555 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

Met ACR20 Criteria at 2 months

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	8/22 (36.4%)	6/22 (27.3%)	RR 1.33 (0.55 to 3.21)	90 more per 1,000 (from 123 fewer to 603 more)	⊕◯◯◯ Very low	CRITICAL No sig difference

DAS-28 8 weeks

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	30	30	-	MD 0.3 lower (0.65 lower to 0.05 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

Tender joint count (0-28) 8 weeks

1	randomised trials	very serious [a]	not serious	not serious	very serious[b,c]	none	30	30	-	MD 0.1 higher (1.17 lower to 1.37 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

Swollen joint count (0-28) 8 weeks

1	randomised trials	very serious [a]	not serious	not serious	very serious[b,c]	none	30	30	-	MD 0.7 lower (2.19 lower to 0.79 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

VAS pain (0-100) 8 weeks

1	randomised trials	very serious [a]	not serious	not serious	serious[b,d]	none	30	30	-	MD 16 lower (27.95 lower to 4.05 lower)	⊕◯◯◯ Very low	CRITICAL Lower pain in probx group

CI: confidence interval; MD: mean difference; RR: risk ratio

a. Selective reporting of results

b. single study

c. wide CI, crosses zero

d. wide CI

Table 12. Additional data on probiotic supplementation (14, 15)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
3871, Cannarella, 2021	randomized, double-blind, placebo-controlled study	60 days	47 RA patients enrolled in the study, 42 patients completed the study (21 in each group)	Intervention group: daily ingestion of probiotics for 60 days (5 freeze-dried strains; Lactobacillus acidophilus La-14, Lactobacillus casei Lc-11, Lactococcus lactis Ll-23, Bifidobacterium lactis BI-04, and B. bifidum Bb-06) Placebo group: daily ingestion of maltodextrin for 60 days	Reported as median (interquartile range) at baseline and 60 days Placebo *Baseline* ESR 23.00 (9.00-48.50) DAS-28 3.83 (2.75-4.69) *60 days* ESR 29 (12-39) DAS-28 3.88 (2.29-4.45) *p-value* ESR 0.717 DAS-28 0.411 Probiotics *Baseline* ESR 19.50 (14.50-33.00) DAS-28 3.20 (2.47-4.21) *60 days*

					ESR 25.00 (16.00-42.00) DAS-28 3.18 (2.49-3.96) *p-value* 0.197 0.526
4926, Vaghef-Mehrabany, 2013	A double-blind, randomized, placebo-controlled trial	8 weeks	RA patients = 46 Probiotic supplementation n = 22 Age, mean: 41.14 ± 12.65 Female: 22 Placebo n = 24	Probiotic supplements given to treatment group were given L. casei, the active agent of the probiotic capsules, and maltodextrin was used as the excipient. The placebo capsules contained only maltodextrin.	Tender and swollen joint counts decreased significantly in the probiotic group by the end of study (P = 0.003 and P = 0.003, respectively) compared to the placebo group. VAS score decreased significantly in the probiotic group compared to the placebo group (P<0.001). DAS28 score also significantly decreased in the probiotic group compared to the control group (0.039). Physical activity scores between groups did not differ significantly by the end of the study (p = 0.602).

			Age, mean: 44.29 ± 9.77 Female: 24		

Glucosamine

Comparison: Glucosamine vs. Placebo

Evidence Summary: One RTC (Nakamura et al) looked at glucosamine vs. placebo. They found that pain, painful joint count, and patient and physician globals were improved with glucosamine, but there was no difference in swollen joint count.

Quality of evidence: Low

Table 13. Glucosamine vs. Placebo (16)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain VAS 0-10, 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	not serious	none	25	26	-	MD **2.89 lower** (5.49 lower to 0.29 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Significantly lower pain score in those taking glucosamine.**

Painful Joint Count, 12 weeks

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	25	25	-	MD **1.3 lower** (2.56 lower to 0.04 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower painful joint count in those taking glucosamine.**

Swollen Joint Count, 12 weeks

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	25	26	-	MD **1.24 lower** (2.7 lower to 0.22 higher)	⨁◯◯◯ Very low	CRITICAL No difference in SJC.

CI: confidence interval; **MD:** mean difference
a. 4 unclear risk and 1 high risk (reporting bias)

b. surrogate for disease activity

c. Crosses 0 (no effect threshold)

Table 14. Additional data on Glucosamine vs. Placebo (16)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
2004_Nakamura	RCT	12 weeks	51 RA patients	Glucosamine 1500 mg/day for 12 weeks vs placebo	2 = improvement 1 = slight improvement 0 = no improvement -1 = worsening -2 = extreme worsening Patient global assessment post intervention: Gluco median = 1 (10^{th} percentile = -1 90^{th} percentile = 2) Placebo median = 0 (10^{th} percentile = -1 90^{th} percentile = 1) P <0.05 Physicians global assessment post intervention: Gluco median = 0 (10^{th} percentile = 0 90^{th} percentile = 2) Placebo median = 0 (10^{th} percentile = -1 90^{th} percentile = 0) P <0.05

Vitamin E

Comparison: Vitamin E vs. placebo

Evidence Summary: One double-blind, RCT (Aryaeian et al), randomized 87 RA patients to receive conjugated linoleic acid (CLA), vitamin E, CLA + Vit E, or placebo for 12 weeks. Amongst subjects in the Vit E treatment arm, there was no difference in pain or RA disease activity as assessed by the DAS28, TJC, or SJC compared to placebo.

Quality of evidence: Low

Table 15. Vitamin E vs. Placebo (17)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vit E	Placebo	Relative (95% CI)	Absolute (95% CI)		
Morning pain reduction (mm), 3 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	21	22	-	MD 1.3 lower (16.78 lower to 14.18 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference
Night pain reduction (mm), 3 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	21	22	-	MD 10.43 higher (4.27 lower to 25.13 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference
After activity of pain reduction, (mm) 3 months												

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	21	22	-	MD 10.75 higher (2.67 lower to 24.17 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

Morning stiffness reduction (hour), 3 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	21	22	-	MD 0.61 lower (1.21 lower to 0.01 lower)	⊕⊕⊕◯ Moderate	IMPORTANT Less reduction in AM stiffness in Vit E

Swollen joint count reduction, 3 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	21	22	-	MD 1.57 higher (3.77 lower to 6.91 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

Tender joint count reduction, 3 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	21	22	-	MD 0.61 higher (2.07 lower to 3.29 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

DAS 28 reduction, 3 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	21	22	-	MD 0.46 higher (0.1 lower to 1.02 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

CRP (IU/mL) 3 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	21	22	-	MD 5.24 higher (7.37 lower to 17.85 higher)	⊕◯◯◯ Very low	IMPORTANT No sig difference

ESR (mm/h) 3 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	21	22	-	MD 5.24 higher (7.37 lower to 17.85 higher)	⊕◯◯◯ Very low	IMPORTANT No sig difference

CI: confidence interval; MD: mean difference

a. Single study

b. Wide CI, crosses zero

c. Nonspecific lab measure of disease activity

Omega 6+ vitamin E

Comparison: Conjugated linoleic acid + Vitamin E vs. placebo

Evidence Summary: One double-blind, RCT (Aryaeian et al), randomized 87 RA patients to receive conjugated linoleic acid (CLA, an Omega-6 fatty acid), vitamin E, CLA + Vit E, or placebo for 12 weeks. Amongst subjects who received CLA + Vit E, a modest but statistically significant greater reduction in DAS28 (1.49 point greater reduction) was seen in patients on CLA + Vit E vs placebo, likely driven primarily by patient/provider global scores as no significant difference was observed in tender or swollen joint counts between treatment arms. Reported pain reduction was greater (range 32-35/100 point greater reduction) in patients on CLA + Vit E. Of note, there were no significant differences in disease activity amongst patients in the Vit E only treatment arm in this study

Quality of evidence: Low

Table 16. Conjugated linoleic acid + Vitamin E vs. Placebo (17)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	conjugated linoleic acid + vitamin E	Placebo	Relative (95% CI)	Absolute (95% CI)		
Morning pain reduction (mm), 3 months												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	22	22	-	MD 32.5 higher (12.64 higher to 52.36 higher)	⨁⨁⨁◯ Moderate	CRITICAL Greater reduction in pain on CLA + Vit E
Night pain reduction (mm), 3 months												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	22	22	-	MD 35.69 higher (18.56 higher to 52.82 higher)	⨁⨁⨁◯ Moderate	CRITICAL Greater reduction in pain on CLA + Vit E
After activity pain reduction (mm), 3 months												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	22	22	-	MD 34.55 higher (18.89 higher to 50.21 higher)	⨁⨁⨁◯ Moderate	CRITICAL Greater reduction in Pain on CLA + Vit E

Morning stiffness reduction (hour) 3 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	22	22	-	MD 0.87 higher (0.3 higher to 1.44 higher)	⊕⊕⊕◯ Moderate	IMPORTANT Greater AM stiffness reduction on CLA + VitE

Swollen joint count reduction 3 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	22	22	-	MD 3.5 higher (0.44 lower to 7.44 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

Tender joint count reduction 3 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	22	22	-	MD 1.78 higher (0.46 lower to 4.02 higher)	⊕⊕◯◯ Low	CRITICAL No sig difference

DAS 28 reduction 3 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	22	22	-	MD 1.49 higher (0.9 higher to 2.08 higher)	⊕⊕⊕◯ Moderate	CRITICAL Greater reduction in DAS28 on CLA+VitE

CRP (IU/mL) 3 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	22	22	-	MD 2.31 lower (5.16 lower to 0.54 higher)	⊕◯◯◯ Very low	IMPORTANT No sig difference

ESR (mm/h) 3 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	22	22	-	MD 9.27 lower (18.69 lower to 0.15 higher)	⊕◯◯◯ Very low	IMPORTANT No sig difference

CI: confidence interval; MD: mean difference

a. Single study

b. CI crosses zero

c. Nonspecific lab measure of disease activity

Omega 6

Comparison: Omega 6 vs. Placebo

Evidence Summary: Three RCTs examined the effects of Omega 6 fatty acid supplementation in patients with RA. Kumar et al compared Borage oil supplements (containing 1320 mg gamma linolenic acid) to placebo amongst 28 RA patients randomized to Borage oil or control. There was no significant difference between groups in the pain VAS score. Notably, fewer patients in the borage oil group reported feeling worse overall at the end of the study, however there was no difference in the number of patients that felt better. The study also suffers from significant imprecision, attrition bias (50% of subjects not used for analysis), and reporting bias, as Ritchie articular index, morning stiffness, and grip strength were measured but were only reported as "not significant" without values being reported. Aryaeian et al randomized 87 RA patients 1:1:1:1 to receive conjugated linoleic acid (CLA, an omega 6 fatty acid), vitamin E, CLA + Vit E, or placebo. Amongst patients in the CLA arm, a modest but statistically significant greater reduction in morning stiffness, tender and swollen joint counts, as well as DAS28 (1.62 pt greater reduction, 95% CI 0.95-2.29) was seen in the CLA group relative to placebo. Abd-Nikfarjam et al compared evening primrose oil containing 420 mg of gamma linolenic acid with

stinging nettle or placebo. The study found a significantly lower DAS-28-ESR and CRP in the primrose oil group compared to placebo at 3 months but there was no difference in the patient global VAS or the ESR at 3 months. The study suffered from significant attrition and lack of intent-to-treat analysis of the patients lost to followup, as well as potential unblinding as the stinging nettle and primrose oil supplements were formulated differently and participants may have been able to recognize the difference.

Quality of evidence: Low to very low

Table 17. Borage oil (Omega 6 fatty acid) vs. Placebo (18)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Borage Oil	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain VAS 0-10												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	14	14	-	MD **1.29 lower** (2.99 lower to 0.41 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
GROC Better												
1	randomised trials	very serious[a]	not serious	serious[c]	very serious[d]	none	5/13 (38.5%)	2/12 (16.7%)	**RR 2.31** (0.55 to 9.74)	**218 more per 1,000** (from 75 fewer to 1,000 more)	⨁◯◯◯ Very low	NOT IMPORTANT No significant difference
GROC Worse												

1	randomised trials	very serious[a]	not serious	serious[c]	serious[e]	none	2/13 (15.4%)	8/12 (66.7%)	**RR 0.23** (0.06 to 0.88)	**513 fewer per 1,000** (from 627 fewer to 80 fewer)	⨁◯◯◯ Very low	NOT IMPORTANT **Statistically significant in favor of borage oil**

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

a. High risk of attrition bias (50%), non-significant disease activity measures not reported

b. Very wide CI that crosses 0 and high effect threshold

c. Indirect measure of health status

d. Very wide CI that crosses 1 and high effect threshold

e. Very wide CI that does not cross 1

Table 18. Conjugated linoleic acid (Omega 6 fatty acid) vs. Placebo (17)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**conjugated linoleic acid (CLA)**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Morning pain reduction (mm), 3 months												
1	**randomised trials**	**not serious**	**not serious**	**not serious**	**very serious[a]**	**none[b]**	**22**	**22**	**-**	**MD 15.66 higher (0.54 lower to 31.86 higher)**	⨁⨁◯◯ **Low**	**CRITICAL** **No sig difference**
Night pain reduction (mm), 3 months												

1	randomised trials	not serious	not serious	not serious	serious[b]	none[b]	22	22	-	MD 24.21 higher (6.68 higher to 41.74 higher)	⊕⊕⊕◯ Moderate	CRITICAL Greater PM pain reduction in CLA group

After activity pain reduction (mm) 3 months

1	randomised trials	not serious	not serious	not serious	serious[b]	none[b]	22	22	-	MD 33.73 higher (17.52 higher to 49.94 higher)	⊕⊕⊕◯ Moderate	CRITICAL Greater post activity pain reduction in CLA group

Morning stiffness reduction (hour) 3 months

1	randomised trials	not serious	not serious	not serious	serious[b]	none[b]	22	22	-	MD 0.73 higher (0.21 higher to 1.25 higher)	⊕⊕⊕◯ Moderate	IMPORTANT Greater morning stiffness reduction in CLA group

Swollen joint count reduction, 3 months

1	randomised trials	not serious	not serious	not serious	serious[b,c]	none[b]	22	22	-	MD 5.58 higher (1.01 higher to 10.15 higher)	⊕⊕⊕◯ Moderate	CRITICAL Greater SJC reduction in CLA group (wide CI)

Tender joint count reduction, 3 months

1	randomised trials	not serious	not serious	not serious	serious[b]	none[b]	22	22	-	MD 4.45 higher (1.63 higher to 7.27 higher)	⊕⊕⊕◯ Moderate	CRITICAL Greater TJC reduction in CLA group

DAS 28 Reduction, 3 months

1	randomised trials	not serious	not serious	not serious	serious[b]	none[b]	22	22	-	MD 1.62 higher (0.95 higher to 2.29 higher)	⊕⊕⊕◯ Moderate	CRITICAL Greater DAS28 reduction in CLA group

CRP (IU/ml) - 3 months

1	randomised trials	not serious	not serious	serious[d]	very serious[a]	none[b]	22	22	-	MD 0.02 lower (3.31 lower to 3.27 higher)	⊕◯◯◯ Very low	IMPORTANT No difference in CRP

ESR (mm/h) 3 months

1	randomised trials	not serious	not serious	serious[d]	very serious[a]	none[b]	22	22	-	MD 7.9 lower (17.04 lower to 1.24 higher)	⊕◯◯◯ Very low	IMPORTANT No difference in ESR

CI: confidence interval; MD: mean difference

a. Only one study, CI crosses zero

b. Only one study

c. Wide CI

d. Nonspecific laboratory measure of disease activity

Table 19. Evening Primrose Oil (gamma-linolenic acid) vs. Placebo(19)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Primrose oil	placebo	Relative (95% CI)	Absolute (95% CI)		
DAS-28-ESR (3 months)												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	29	30	-	MD **0.75 lower** (1.23 lower to 0.27 lower)	⨁⨁⨁◯ Moderate	CRITICAL Significant difference in favor of primrose oil
Patient global VAS (3 months)												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	29	30	-	MD **0.11 higher** (1.36 lower to 1.58 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
CRP (3 months)												
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	29	30	-	MD **3.75 lower** (5.91 lower to 1.59 lower)	⨁⨁◯◯ Low	CRITICAL Significant difference in favor of primrose oil
ESR (3 months)												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	29	30	-	MD **5.14 lower** (10.41 lower to 0.13 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Risk of participant unblinding due to different supplement formulations. High attrition with no intent-to-treat analysis (only 90 patients left at final endpoint were analyzed)

b. Not a direct measure of disease activity

c. Wide CI crossing zero

Omega 3 and Omega 3+6

Comparison: Primrose oil + Fish oil/Omega 3 vs. placebo

Evidence Summary: Two randomized control trials (Belch, Veselinovic) looked at primrose oil + fish oil or omega 3. One double blind RCT (1955 Belch) compared evening primrose oil with and without fish oil to placebo. Veselinovic showed that patients on omega 3 and primrose oil showed lower DAS28 scores and lower pain scores at 12 weeks, but otherwise showed no difference. Belch reported changes in outcomes from baseline between the groups was inconsistent, with no clear indication of superiority of the supplements compared to placebo.

Quality of Evidence: Very low

Table 20. Primrose oil + Fish oil/Omega 3 vs. Placebo (20)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Omega 3 FA + Primrose Oil	Control	Relative (95% CI)	Absolute (95% CI)		
DAS28, 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	20	20	-	MD **0.44 lower** (0.87 lower to 0.01 lower)	⨁⨁⨁◯ Moderate	CRITICAL **DAS28 lower in patients on omega 3 and primrose oil at 12 weeks.**
CRP, 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Omega 3 FA + Primrose Oil	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	20	-	MD **0.2 higher** (2.66 lower to 3.06 higher)	⨁◯◯◯ Very low	CRITICAL No difference in CRP.

Tender joint count, 12 weeks

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	20	-	MD **0.6 lower** (1.53 lower to 0.33 higher)	⨁◯◯◯ Very low	CRITICAL No difference in TJC.

Swollen joint count, 12 weeks

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	20	-	MD **0.1 lower** (0.46 lower to 0.26 higher)	⨁◯◯◯ Very low	CRITICAL No difference in SJC.

Pain VAS, 12 weeks

1	randomised trials	serious[a]	not serious	not serious	not serious	none	20	20	-	MD **8.8 lower** (13.11 lower to 4.49 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Significantly lower pain score in patients on omega 3+primrose oil compared to control at 12 weeks.**

ESR, 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Omega 3 FA + Primrose Oil	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	very serious[b]	serious[c]	none	20	20	-	MD **4.2 lower** (11.91 lower to 3.51 higher)	⨁◯◯◯ Very low	CRITICAL No difference in ESR.

CI: confidence interval; **MD:** mean difference
a. Study with 2 unclear and one high risk

b. Surrogate for disease activity

c. Crosses 0 (no-effect threshold)

Table 21. Additional data on Primrose oil + Fish oil/Omega 3 vs. Placebo (21)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results

<table>
<tr>
<td>1955 Belch 1988</td>
<td>RCT, double blinded</td>
<td>12 months</td>
<td>49 patients with RA, all of whom were on NSAIDs for disease but did not require any DMARDs</td>
<td>Primrose oil (EPO): total dose of 540 mg of gamma linolenic acid (GLA)/day

Primrose oil + Fish oil: 450 mg GLA and 240mg Eicosapentaenoic acid(EPA)/day

Placebo

all capsules contained vitamin E (dose 120 mg/day)</td>
<td>The following outcomes are presented as %change from baseline at 12M (with baseline considered 100%)
<table>
<tr><th>Outcome</th><th>EPO (n= 16)</th><th>EPO/fishoil (n=15)</th><th>Placebo (n=18)</th></tr>
<tr><td>AM stiffness_12M</td><td>39%</td><td>189%</td><td>128%</td></tr>
<tr><td>Pain VAS_12M</td><td>62%</td><td>116%</td><td>17%</td></tr>
<tr><td>Grip Strength_12M</td><td>100%</td><td>71%</td><td>57%</td></tr>
<tr><td>Articular Index_12M</td><td>135%</td><td>103%</td><td>97%</td></tr>
<tr><td>ESR_12M</td><td>134%</td><td>73%</td><td>96%</td></tr>
<tr><td>CRP_12M</td><td>118%</td><td>78%</td><td>130%</td></tr>
</table>
</td>
</tr>
</table>

Comparison: Fish oil vs. Placebo

Evidence Summary: There were 5 RCTs that looked at fish oil vs placebo. In terms of pain, Skoldstam et al and Tulleken et al showed slightly lower pain scores in those receiving fish oil, but Kremer, Berbert, and Nielsen did not show any improvement in pain scores, and overall analysis did not show improvement in pain score. Fish oil did not show any improvement in the following surrogate measures of disease activity: Ritchie articular index (Skoldstam and Berbert), patient and physician global (Kremer), ESR (Skoldstam and Tulleken), and swollen joint count (Skoldstam and Tulleken). While Berbert showed decreased morning stiffness in those on fish oil, this was not seen and Kremer and overall, there was no

difference. There was no difference in functional status (Berbert) or surrogates of fatigue (Kremer and Berbert) or grip strength (overall for Skoldstam, Kremer, Berbert).

Quality of evidence: Very low

Table 22. Fish oil vs. Placebo (22-26)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish oil	placebo	Relative (95% CI)	Absolute (95% CI)		
Global arthritis activity change, 6 months												
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	22	21	-	MD **0.39 lower** (0.49 lower to 0.29 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower global activity score in those on fish oil.**
Pain scale VAS change, 3-6 months												
3	randomised trials	very serious[c]	not serious	not serious	not serious	none	45	44	-	MD **0.15 lower** (0.24 lower to 0.06 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower pain score in fish oil group.**
Richie's index change, 3-6 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish oil	placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[c]	not serious	serious[b]	serious[d]	none	30	30	-	MD **0.17 lower** (0.87 lower to 0.54 higher)	⨁◯◯◯ Very low	CRITICAL No difference in Ritchie's index.

ESR (mmHg) change, 6 months

1	randomised trials	serious[a]	not serious	serious[b]	serious[d]	none	22	21	-	MD **0** (2.12 lower to 2.12 higher)	⨁◯◯◯ Very low	CRITICAL No difference in ESR.

CRP (umol/l) change, 6 months

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	22	21	-	MD **5 lower** (8.81 lower to 1.19 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower CRP in the fish oil group.**

NSAID Consumption change, 6 months

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	22	21	-	MD **0.16 lower** (0.23 lower to 0.09 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower amount of NSAID use in the fish oil group.**

Morning stiffness (minutes), 24-30 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish oil	placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[c]	not serious	serious[b]	not serious	none	23	23	-	MD 49.07 **lower** (80.53 lower to 17.61 lower)	⨁◯◯◯ Very low	CRITICAL **Significantly lower minutes of morning stiffness in the fish oil group.**

Onset of Fatigue (minutes), 24-30 weeks

2	randomised trials	very serious[c]	very serious[e]	serious[f]	serious[d]	none	23	23	-	MD **0.07 higher** (7.47 lower to 7.61 higher)	⨁◯◯◯ Very low	CRITICAL No difference in fatigue.

Patient Global Assessment, 24-30 weeks

2	randomised trials	very serious[c]	not serious	serious[b]	serious[d]	none	23	23	-	MD **0.01 lower** (0.54 lower to 0.53 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global.

Functional Status, 24 weeks

1	randomised trials	very serious[c]	not serious	not serious	serious[d]	none	8	9	-	MD **0.15 higher** (0.8 lower to 1.1 higher)	⨁◯◯◯ Very low	CRITICAL No difference in functional status.

Change in Tender Joint Count, week 26-30

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish oil	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[d]	none	15	14	-	MD **1.4 lower** (8.08 lower to 5.28 higher)	⨁◯◯◯ Very low	CRITICAL No difference in tender joint count.

Change in Swollen Joint Count, week 26-30

1	randomised trials	serious[a]	not serious	serious[b]	serious[d]	none	15	14	-	MD **0.9 higher** (5.35 lower to 7.15 higher)	⨁◯◯◯ Very low	CRITICAL No difference in swollen joint count.

Change in Physician Assessment of Pain, week 26-30

1	randomised trials	serious[a]	not serious	serious[g]	serious[d]	none	15	14	-	MD **0.48 lower** (1.15 lower to 0.19 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician assessment of pain.

Grip Strength, 3-6 months

3	randomised trials	very serious[c]	very serious[e]	serious[f]	not serious	none	75	74	-	MD **6.02 lower** (8.57 lower to 3.47 lower)	⨁◯◯◯ Very low	CRITICAL **Significantly lower grip strength in fish oil group.**

Change in Physician Global Assessment of Arth Activity, week 24-30

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish oil	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[d]	none	15	14	-	MD **0.23 lower** (0.75 lower to 0.29 higher)	⊕◯◯◯ Very low	CRITICAL No difference in physician global.

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference
a. High level of unclear risk

b. Surrogate measure for disease activity

c. At least one study with 2 high risk categories for bias

d. Crosses no effect threshold

e. I2 70-100

f. Surrogate for functional status

g. Surrogate for pain

Table 23. Additional data on Fish oil vs. Placebo (24, 26)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results

1986 Nielsen 1992	**RCT**	**12 weeks**	**51 RA patients** **n = 27 intervention group; n = 24 control group**	**intervention group: daily supplement of 6 capsules of fish oil** **control group: daily supplementation of 6 capsules of fat comparable to the average Danish diet (flavored like fish oil)** **both groups: all other diet and medications held constant**	**Median (25^{th}-75^{th}quartile)** **Timepoint for all = baseline to 12 week change** **RA Disease Activity** · **ESR (negative) (mm H^{-1})** o **Interv: 34 (17-55) to 34 (21-59)** o **Control: 33 (20-40) to 33 (19-44)** · **CRP (negative) (mg l^{-1})** o **Interv: 21 (9-41) to 17 (9-26)** o **Control: 17 (11-26) to 18 (13-25)** · **joint swelling (negative) (index)** o **interv: 8 (6-10) to 8 (5-9)** o **control: 8 (6-10) to 8 (6-11)** · **morning stiffness (negative) (min)** o **interv: 120 (60-180) to 75 (30-120)** o **control: 120 (90-120) to 120 (53-180)** **Pain:** · **joint tenderness (negative) (index)** o **interv: 10 (8-13) to 8 (5-11)** o **control: 12 (10-15) to 10 (6-16)** · **global pain (negative) (arbitrary units)** o **interv: 120 (90-143) to 104 (78-143)**

					o control: 118 (81-142) to 136 (86-170) **Functional status:** **· grip strength (positive) (mmHg)** **o interv: 63 (42-140) to 78 (50-118)** **o control: 130 (80-146) to 120 (72-159)** **Negative = lower scores are better; Positive = higher scores are better** **Over the 12 weeks, morning stiffness and CRP decreased significantly in the fish oil group but not the placebo group. Joint tenderness decreased in both groups. Neither group had improvement in joint swelling, pain VAS, grip strength, or daily activity score.**

<table>
<tr>
<td>6724 Tulleken 1990</td>
<td>Randomized controlled trial</td>
<td>3 months</td>
<td>28 patients with RA</td>
<td>Daily fish oil supplement compared with placebo for 3 months</td>
<td>Results are given as median values with ranges.
<table>
<tr><th></th><th colspan="2">Fish Oil Group</th><th colspan="2">Placebo</th></tr>
<tr><th></th><th>Before</th><th>After</th><th>Before</th><th>After</th></tr>
<tr><td>Joint Pain index</td><td>27 (3-103)</td><td>6 (0-49)</td><td>15.5 (5-27)</td><td>11.5 (4-29)</td></tr>
<tr><td>Ritchie Articular Index</td><td>18 (3-49)</td><td>6 (0-49)</td><td>27 (5-52)</td><td>20 (4-48)</td></tr>
<tr><td>Joint swelling Index</td><td>7 (0-26)</td><td>4 (1-16)</td><td>6 (2-14)</td><td>4 (1-16)</td></tr>
<tr><td>Swollen Joints</td><td>6 (0-24)</td><td>3 (1-16)</td><td>5 (2-13)</td><td>4 (1-16)</td></tr>
<tr><td>AM stiffness (minutes)</td><td>60 (0-60)</td><td>30 (0-120)</td><td>45 (0-120)</td><td>60 (0-180)</td></tr>
<tr><td>Pain (10 cm VAS)</td><td>4 (0.5-6.1)</td><td>2.4 (0-7.4)</td><td>4.4 (1.4-8.0)</td><td>3.8 (0.5-8.1)</td></tr>
</table>
After treatment, patients on fish oil had improvement in joint pain index and Ritchie articular inde, but no differences in ESR, CRP, painful joints, and swollen joint count.</td>
</tr>
</table>

Comparison: Fish oil + olive oil vs. Placebo

Evidence Summary: One RTC (Berbert) looked at fish oil + olive oil vs. placebo at 24 weeks. They found improvement in pain. Functional status showed no improvement, although a surrogate (grip strength) did. Surrogates for disease activity (morning stiffness, Ritchie articular index) did improvement, while overall patient global assessment had no difference. This study was limited by a very low number of participants (19 total).

Quality of evidence: Very low

Table 24. Fish oil + olive oil vs. Placebo (22)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish Oil and Olive Oil	Placebo at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
Disease activity as inferred from Morning stiffness, 24 weeks												
1	randomised trials	very serious[a]	not serious	serious[b]	serious[c]	none	10	9	-	MD **40 lower** (77.01 lower to 2.99 lower)	⨁◯◯◯ Very low	CRITICAL **Significantly less morning stiffness in the fish oil+ olive oil group vs. placebo.**
Joint Pain Intensity, 24 weeks												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	10	9	-	MD **1.32 lower** (2.25 lower to 0.39 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower joint pain intensity in the fish oil+ olive oil group vs. placebo.**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish Oil and Olive Oil	Placebo at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
Onset of Fatigue, 24 weeks												
1	randomised trials	very serious[a]	not serious	serious[d]	serious[e]	none	10	9	-	MD **2.3 lower** (9.7 lower to 5.1 higher)	⨁◯◯◯ Very low	CRITICAL No difference in fatigue.
Ritchie Articular Index, 24 weeks												
1	randomised trials	very serious[a]	not serious	serious[b]	not serious	none	10	9	-	MD **4 lower** (7.21 lower to 0.79 lower)	⨁◯◯◯ Very low	CRITICAL **Significantly lower Ritchie articular index in the fish oil+ oil olive group + placebo.**
Grip Strength (right hand), 24 weeks												
1	randomised trials	very serious[a]	not serious	very serious[d]	serious[c]	none	20	18	-	MD **42.9 higher** (4.07 higher to 81.73 higher)	⨁◯◯◯ Very low	CRITICAL **Significantly stronger grip strength in the fish oil+ olive oil group vs. placebo.**
Patient Global Assessment, 24 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish Oil and Olive Oil	Placebo at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	serious[b]	serious[f]	none	10	9	-	MD **0.43 lower** (1.19 lower to 0.33 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global.

Functional Status, 24 weeks

1	randomised trials	very serious[a]	not serious	not serious	serious[f]	none	10	9	-	MD **0** (0.99 lower to 0.99 higher)	⨁◯◯◯ Very low	CRITICAL No difference in functional status.

CI: confidence interval; **MD:** mean difference
a. two high risk ratings, 3 unclear

b. Surrogate for disease activity

c. Very wide confidence interval

d. surrogate for functional status

e. Crosses 0 and wide confidence interval

f. Crosses 0

Comparison: EPA (omega 3 fatty acid) vs. Placebo

Evidence Summary: One RCT (Kremer et al) examined the effect of eicosapentanoic acid (EPA) supplementation compared to placebo in a sample that consisted initially of 66 RA patients. This study was subject to significant bias, as only 49 patients completed any clinical evaluation in follow up and only 39 completed the study. Additionally there were significantly more patients on mtx (39 vs 13%) or prednisone (47% vs 23%) in the fish oil group at baseline. A greater reduction in morning stiffness was seen in the EPA group, whereas there was no difference in fatigue, functional status (assessed by grip strength and 50ft walk time), or tender/swollen joint counts.

Quality of evidence: Very low

Table 25. EPA (Omega 3 Fatty Acid) vs. Placebo (27)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Eicosapentanoic acid	Placebo	Relative (95% CI)	Absolute (95% CI)		
Morning stiffness (min) 12 wk												
1	randomised trials	very serious[D]	not serious	not serious	serious[a]	none	17	21	-	MD 75 lower (139.88 lower to 10.12 lower)	⨁◯◯◯ Very low	IMPORTANT Greater reduction in AM stiffness in EPA group
Time to fatigue (min) 12 wk												
1	randomised trials	very serious[D]	not serious	serious[b]	very serious[a,c]	none	17	21	-	MD 35 lower (156.23 lower to 86.23 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Grip strength 12 wk												
1	randomised trials	very serious[D]	not serious	serious[b]	very serious[a,c]	none	17	21	-	MD 21 higher (17.3 lower to 59.3 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
50 ft walk (s)												

1	randomised trials	very serious[D]	not serious	not serious	very serious[a,c]	none	17	21	-	MD 0 (2.26 lower to 2.26 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

Tender joints

1	randomised trials	very serious[D]	not serious	not serious	very serious[a,c]	none	17	21	-	MD 2.6 lower (7.04 lower to 1.84 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

Swollen joints

1	randomised trials	very serious[D]	not serious	not serious	very serious[a,c]	none	17	21	-	MD 0.1 lower (3.77 lower to 3.57 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

CI: confidence interval; MD: mean difference

a. Single study

b. Functional status surrogate

c. Wide CI, crosses zero

D. High drop out rate (>25%)

Table 26. Additional data on EPA (Omega 3 Fatty Acid) vs. Placebo (27)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results

2021, Kremer, 1985	**Double-blind, controlled, randomized trial**	**Follow-up at 4 weeks, 8 weeks, and 12 weeks**	**44 RA patients**	**Treatment group: 10 capsules daily with a total of 1.8 g EPA, diet with a ratio of polyunsaturated to saturated fats of 1/4** **Control group: 10 capsules daily with non-digestible paraffin wax, diet with random manipulations and messaging about avoiding foods high in polyunsaturates** **Both groups received instruction on how to eat a balanced diet and daily multivitamin tablets**	**There was a significant difference between groups at 12-weeks for morning stiffness. Between the 12 week evaluation and the 1-2 month follow-up, there were significant decreases in health found in patients' rating of pain and overall condition, physicians' pain rating, and physicians' overall evaluations.**

Comparison: EPA + DHA (omega 3 fatty acid) vs. Placebo

Evidence Summary: Two RCTs examined the effect of combination EPA + DHA compared to placebo. Park et al conducted a multicenter, double blind study of 109 patients randomized to EPA/DHA or placebo. No significant differences were seen in pain, morning stiffness, physician global, or functional status as assessed by the HAQ. Rajaei et al conducted a double blind RCT of 60 patients with lower dropout rate (~18%) but where significantly ($p<0.05$) greater reductions in pain, morning stiffness, tender and swollen joints, and ESR were seen though some concerns were raised due to selective reporting of quantitative data, where specific means and SDs as reported in the methods were not provided.

Quality of evidence: Very low

Table 27. EPA + DHA (Omega 3 Fatty Acid) vs. Placebo (28)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Eicosapentanoic + docosahexaenoic acid	Placebo	Relative (95% CI)	Absolute (95% CI)		
PhyGA (0-10) 16 wk												
1	randomised trials	serious[d]	serious[c]	not serious	very serious[a,b]	none	41	40	-	MD 0.28 higher (0.5 lower to 1.06 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
PatGA												
1	randomised trials	serious[d]	serious[c]	not serious	very serious[a,b]	none	41	40	-	MD 0.01 lower (1.06 lower to 1.04 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Morning stiffness (mins)												
1	randomised trials	serious[d]	serious[c]	not serious	very serious[a,b]	none	41	40	-	MD 22.55 higher (3.35 lower to 48.45 higher)	⨁◯◯◯ Very low	IMPORTANT No sig difference
Pain scale (0-100)												

1	randomised trials	serious[d]	serious[c]	not serious	very serious[a,b]	none	41	40	-	MD 7.1 higher (3.26 lower to 17.46 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference

KHAQ

1	randomised trials	serious[d]	not serious	not serious	very serious[a,b]	none	41	40	-	MD 0.07 higher (0.17 lower to 0.31 higher)	⨁⨁◯◯ Low	CRITICAL No sig difference

CI: confidence interval; MD: mean difference

a. Single study

b. CI crosses zero

C. Conflicting results between rcts

D. Selective reporting bias

Table 28. Additional data on EPA + DHA (Omega 3 Fatty Acid) vs. Placebo (29)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results

1820 Rajaei 2016	**RCT – double blind**	**3 months**	**60 participants (49 females, 11 males) who attended the rheumatology clinic of Ahvaz Golestan Hospital with active RA.**	**Omega-3 group (n=25) - patients consumed two omega-3 capsules daily which contained 1.8 and 2.1 g of EPA and DHA respectively** **Placebo group (n=24) - consumed two placebo Capsules daily containing starch**	

	Placebo		Omega-3	
	Base	End	Base	End
Morning Stiffness	116	94	128	40
Number of tender joints	24	20	21	5
Number of swollen joints	7	5	10	3
ESR	35	33	39	16
CRP	2+	2+ to 3+	2+	0 to 1+
Patient's pain assessment	8	8	9	4
Doctor's pain assessment	4	5	4	2

Comparison: Fatty acid vs. Placebo

Evidence Summary: One RCT (Dawcynski 2011) compared 1575 mg n-3 LC-PUFA to placebo, as part of a comparison of multiple fatty acid formulations. 14 patients were randomized to the fatty acid group and 12 patients were in the control group. There was no significant difference between the fatty acid and the control group in either the DAS28 at 12 weeks or the pain VAS score at 12 weeks.

Quality of Evidence: Very Low

Table 29. Fatty acid vs. Placebo (30)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fatty acid	placebo	Relative (95% CI)	Absolute (95% CI)		
Disease activity (DAS28) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	14	12	-	MD 0.1 higher (0.56 lower to 0.76 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
VAS 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	14	12	-	MD 5.4 lower (23.35 lower to 12.55 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

CI: confidence interval; MD: mean difference

a. Significant differential attrition between supplement and control groups

b. Wide CI that crosses 0

c. Wide CI that crosses 0 and high effect threshold

Comparison: 5.2 mg of omega 3 vs. Placebo

Evidence Summary: Four RCTs looked at the use of 5.2 mg of omega 3 vs placebo (Volker, Cleland, Veselinovic, and Galarraga). All studies were relatively small and Volker, Cleland, and Galarraga had a high dropout rate. Overall, there were no differences in any measures of disease activity, function, or pain.

Quality of evidence: Very low

Table 30. 5.2 mg of omega 3 vs. Placebo (20, 31-33)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	5.2 gm Omega 3 FA	control	Relative (95% CI)	Absolute (95% CI)		
Grip strength, 12- 16 weeks												
2	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	121	119	-	MD **16.74 higher** (7.95 lower to 41.42 higher)	⨁◯◯◯ Very low	CRITICAL No difference in grip strength.
Pain score, 12-16 weeks												
4	randomised trials	serious[a]	very serious[d]	not serious	serious[c]	none	105	104	-	MD **6.85 lower** (15.66 lower to 1.96 higher)	⨁◯◯◯ Very low	CRITICAL No difference in pain score.
DAS28, 12 weeks												
2	randomised trials	serious[a]	not serious	not serious	serious[c]	none	69	68	-	MD **0.12 lower** (0.53 lower to 0.29 higher)	⨁⨁◯◯ Low	CRITICAL No difference in DAS 28.
Tender joint count, 12-15 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	5.2 gm Omega 3 FA	control	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[e]	not serious	serious[f]	not serious	none	33	33	-	MD **1.3 lower** (2.26 lower to 0.34 lower)	⨁◯◯◯ Very low	CRITICAL **Significantly lower TJC in those on 5.2 g of omega 3 compared to controls.**

Swollen joint count, 12-15 weeks

2	randomised trials	very serious[e]	not serious	serious[f]	not serious	none	33	33	-	MD **0.4 higher** (0.24 higher to 0.56 higher)	⨁◯◯◯ Very low	CRITICAL **Significantly higher SJC in those on omega 3 vs controls.**

ESR, 12-15 weeks

2	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	33	33	-	MD **1.65 lower** (11 lower to 7.69 higher)	⨁◯◯◯ Very low	CRITICAL No difference in ESR.

HAQ, 12-15 weeks

2	randomised trials	very serious[e]	very serious[d]	serious[f]	serious[c]	none	62	61	-	MD **8.55 lower** (27.84 lower to 10.74 higher)	⨁◯◯◯ Very low	CRITICAL No difference in HAQ.

Morning stiffness (min), 12-15 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	5.2 gm Omega 3 FA	control	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[e]	very serious[d]	serious[f]	serious[c]	none	62	61	-	MD **52.9 lower** (173.38 lower to 67.58 higher)	⨁◯◯◯ Very low	CRITICAL No difference in morning stiffness.

Change in patient global, 15 weeks

1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	13	13	-	MD **28.1 lower** (73.85 lower to 17.65 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global.

CRP (mg/L), 12 -15 weeks

3	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	82	81	-	MD **0.41 higher** (1.54 lower to 2.37 higher)	⨁◯◯◯ Very low	CRITICAL No difference in CRP.

Achievement of ACR20, 15 weeks

1	randomised trials	very serious[e]	not serious	not serious	serious[c]	none	5/13 (38.5%)	3/13 (23.1%)	**RR 1.67** (0.50 to 5.57)	**155 more per 1,000** (from 115 fewer to 1,000 more)	⨁◯◯◯ Very low	CRITICAL No difference in risk of ACR20.

Change in Physician global, 15 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	5.2 gm Omega 3 FA	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	13	13	-	MD **55.2 lower** (113.04 lower to 2.64 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician global.

Grip strength, 10 weeks

1	randomised trials	very serious[e]	not serious	serious[b]	serious[c]	none	46	48	-	MD **0.82 higher** (9.21 lower to 10.85 higher)	⨁◯◯◯ Very low	CRITICAL No difference in grip strength.

Pain, 10 weeks

1	randomised trials	very serious[e]	not serious	not serious	serious[c]	none	23	24	-	MD **3.13 higher** (10.24 lower to 16.51 higher)	⨁◯◯◯ Very low	CRITICAL No difference in pain.

Early morning stiffness, 10 weeks

1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **0.38 lower** (6.04 lower to 5.27 higher)	⨁◯◯◯ Very low	CRITICAL No difference in early morning stiffness.

HAQ, 10 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	5.2 gm Omega 3 FA	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **0.05 higher** (0.24 lower to 0.34 higher)	⨁◯◯◯ Very low	CRITICAL No difference in HAQ.

ESR, 10 weeks

1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **9.64 higher** (3.02 lower to 22.31 higher)	⨁◯◯◯ Very low	CRITICAL No difference in ESR.

Ritchie Articular Index, 10 weeks

1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **0.22 higher** (3.94 lower to 4.38 higher)	⨁◯◯◯ Very low	CRITICAL No difference in Ritchie index.

CRP, 10 weeks

1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **7.88 higher** (10.49 lower to 26.25 higher)	⨁◯◯◯ Very low	CRITICAL No difference in CRP.

Swollen joint count, 10 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	5.2 gm Omega 3 FA	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **1.07 higher** (3.52 lower to 5.67 higher)	⨁◯◯◯ Very low	CRITICAL No difference in SJC.

Tender joint count, 10 weeks

1	randomised trials	very serious[e]	not serious	serious[f]	serious[c]	none	23	24	-	MD **2.05 higher** (4.78 lower to 8.87 higher)	⨁◯◯◯ Very low	CRITICAL No difference in TJC.

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio
a. Unclear risk and less than 2 high risk categories of bias

b. Surrogate measure of functional status

c. Crosses no effect threshold

d. I2 70-100%

e. At least one study with 2 or greater high risk categories of bias

f. Surrogate measure of disease activity

Comparison: 2.6 g of omega 3 vs. Placebo

Evidence Summary: One double-blind, placebo controlled RCT (Geusens et al) randomized 90 RA patients to 2.6 g of Omega-3, 1.3 g of omega 3, or olive oil placebo. All patients received a recommended prescription diet consisting of 30% fat, 12-15% protein, and 50-58% carbohydrates and were instructed to consume fish once weekly. At the end of the 12 month study, patient global assessment of disease activity was slightly lower (1.52 pts lower, 95% CI 0.16-2.88) in the 2.6g omega 3 group, otherwise there were no differences in RA disease activity as assessed by physician global, tender joint count, and Ritchie articular index. Additionally, there were no differences in pain or functional status (assessed by grip strength) between treatment arms. This study was particularly limited by a high (>30%) drop out rate.

Quality of evidence: Very low

Table 31. 2.6 g of omega 3 vs. Placebo (34)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	2.6 omega-3	olive oil (placebo)	Relative (95% CI)	Absolute (95% CI)		
Physician global assessment												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	19	20	-	MD **0.25 lower** (0.77 lower to 0.27 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Patient global assessment												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	19	20	-	MD **1.52 lower** (2.88 lower to 0.16 lower)	⨁◯◯◯ Very low	CRITICAL Pt global lower in 2.6g Omega 3 group
Patient pain score												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,d]	none	19	20	-	MD **0.36 lower** (0.89 lower to 0.17 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Ritchie articular pain index												

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	19	20	-	MD **1 higher** (10.09 lower to 12.09 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

No of painful joints

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	19	20	-	MD **1 lower** (8.07 lower to 6.07 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

Grip strength

1	randomised trials	very serious[a]	not serious	serious[e]	very serious[b,c]	none	19	20	-	MD **26 higher** (8.68 lower to 60.68 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference

CI: confidence interval; **MD:** mean difference

a. High drop out

b. single study

c. wide CI, crosses zero

d. CI crosses zero

e. surrogate for functional status

Comparison: 1.3g of omega 3 vs. Placebo

Evidence Summary: One double-blind, placebo controlled RCT (Geusens et al) randomized 90 RA patients to 2.6 gm of Omega-3, 1.3 gm of omega 3, or olive oil placebo. All patients received a recommended prescription diet consisting of 30% fat, 12-15% protein, and 50-58% carbohydrates and were instructed to consume fish once weekly. At the end of the 12 month study, there were no differences in RA disease activity as assessed by patient/physician global, tender joint count, and Ritchie articular index in the 1.3g Omega 3 arm vs placebo. Additionally, there were no differences in pain or functional status (assessed by grip strength) between treatment arms. This study was particularly limited by a high (>30%) drop out rate.

Quality of evidence: Very low

Table 32. 1.3g of omega 3 vs. Placebo (34)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	1.3 omega-3	olive oil (placebo)	Relative (95% CI)	Absolute (95% CI)		
Physician global assessment												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	21	20	-	MD **0.02 lower** (0.51 lower to 0.47 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Patient global assessment												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	21	20	-	MD **0.09 lower** (1.6 lower to 1.42 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Patient pain score												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	21	20	-	MD **0.21 lower** (0.74 lower to 0.32 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference
Ritchie articular pain index												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,d]	none	21	20	-	MD **6 higher** (3.8 lower to 15.8 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference

No of painful joints

1	randomised trials	very serious[a]	not serious	not serious	very serious[b,d]	none	21	20	-	MD **1 higher** (6.07 lower to 8.07 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference

Grip strength

1	randomised trials	very serious[a]	not serious	serious[e]	very serious[b,d]	none	21	20	-	MD **9 higher** (17.66 lower to 35.66 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference

CI: confidence interval; **MD:** mean difference

a. high drop out

b. single study

c. CI crosses zero

d. Wide CI, crosses zero

e. Surrogate for functional status

Comparison: 0.82g Omega 3 vs. Placebo

Evidence Summary: An RCT (Kjeldsen-Kragh 1992) evaluated the use of 0.82g of Omega 3 (0.54g EPA and 0.28g DHA) compared to placebo for 10 weeks. 44 subject in total were randomized to omega 3 or placebo. There was no significant difference between groups in Ritchie articular index, SJC, TJC, grip strength, pain, morning stiffness, or HAQ at 10 weeks. The study suffered from high attrition and there was a significant difference in baseline inflammatory markers between groups, suggesting of possible differential baseline disease activity.

Quality of Evidence: Very Low

Table 33. 0.82g of omega 3 vs. Placebo (35)

Certainty assessment	№ of patients	Effect	Certainty	Importance

№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	0.82 gm Omega 3 (2)	Placebo	Relative (95% CI)	Absolute (95% CI)		
RAI, 10 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	20	24	-	MD 0.76 higher (3.57 lower to 5.1 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
SJC, 10 weeks												
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	20	24	-	MD 1.92 higher (3.13 lower to 6.96 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
TJC, 10 weeks												
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	20	24	-	MD 1.55 higher (5.48 lower to 8.57 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
Right grip strength, 10 weeks												
1	randomised trials	serious[a]	not serious	serious[d]	serious[b]	none	20	24	-	MD 4.96 lower (18.23 lower to 8.32 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
Left grip strength, 10 weeks												

1	randomised trials	serious[a]	not serious	serious[d]	serious[b]	none	20	24	-	MD 6.73 lower (18.59 lower to 5.14 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference

Pain, 10 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	20	24	-	MD 7.05 higher (8.45 lower to 22.55 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Early morning stiffness, 10 weeks

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	20	24	-	MD 1.23 lower (9.12 lower to 6.66 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference

HAQ, 10 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	20	24	-	MD 0.03 higher (0.27 lower to 0.34 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

CI: confidence interval; MD: mean difference

a. High attrition bias, significant imbalance between baseline inflammatory markers between groups

b. Wide CI crossing 0

c. Indirect marker of disease activity

d. Indirect marker of functional status

Comparison: N-3 long-chain PUFA compared to Placebo

Evidence Summary: One RCT (Dawczynski 2009) compared n-3 long-chain PUFA supplements in dairy products to unsupplemented dairy products with a 12-week follow-up period. The study was designed as a crossover with a 12 week intervention, an 8 week washout period, and then another 12 week intervention. There was no significant difference in DAS28, swollen joint count, tender joint count, or duration of morning stiffness between the intervention and control groups.

Quality of Evidence: Low

Table 34. N-3 long-chain PUFA compared to Placebo (36)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	n-3 long-chain PUFA	Placebo	Relative (95% CI)	Absolute (95% CI)		
Duration of Morning Stiffness (12 weeks)												
1	randomised trials	serious	not serious	serious[a]	serious[b]	none	39	39	-	MD **5 higher** (6.81 lower to 16.81 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference
Tender Joint Count (12 weeks)												
1	randomised trials	serious[c]	not serious	not serious	serious[b]	none	39	39	-	MD **0.19 lower** (2.9 lower to 2.52 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Swollen Joint Count (12 weeks)												
1	randomised trials	serious[c]	not serious	not serious	serious[b]	none	39	39	-	MD **0.53 higher** (0.59 lower to 1.65 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

DAS28 (12 weeks)

1	randomised trials	serious[c]	not serious	not serious	serious[b]	none	39	39	-	MD **0.08 higher** (0.35 lower to 0.51 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

a. Indirect marker of disease activity

b. Wide CI crossing 0

c. Based on single study with high risk of attrition bias

Comparison: Fatty Acid vs. Placebo

Evidence Summary: (Fatty acid supplementation resulted in a significant improvement in disease activity at 12 weeks; those receiving placebo did not have a significant improvement.)

Quality of evidence: Very low

Table 35. Additional data on Fatty Acid vs. Placebo (37)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1357_Espersen	**RCT**	**12 weeks**	**32 RA patients with low disease activity**	**3.6 g n-3 polyunsaturated fatty acids "Pikasol" or placebo pill**	**Ritchie Arthritis index** **Pikasol mean before/after 10.5/7.5 P<0.02** **Placebo mean before/after 12.6/10.6 n.s.**

Comparison: Fatty acid + g-linolenic acid vs. Placebo

Evidence Summary: One RCT (Dawcynski 2011) compared 1575 mg n-3 LC-PUFA plus 1800 mg GLA/d to placebo, as part of a comparison of multiple fatty acid formulations. 13 patients were randomized to the combined fatty acid group and 12 patients were in the placebo group. There was no significant difference between groups in the pain VAS score at 12 weeks.

Quality of Evidence: Very Low

Table 36. Fatty acid + g-linolenic acid vs. Placebo (30)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fatty acid + g-linolenic acid	placebo	Relative (95% CI)	Absolute (95% CI)		
VAS 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	13	12	-	MD **3.7 higher** (16.33 lower to 23.73 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

a. Significant differential attrition between supplement

Comparison: Nutritional Supplement (Omega-3, Omega-6, micronutrients) vs. Placebo

Evidence Summary: One RCT (Remans 2004) evaluated a nutritional supplement containing EPA, DHA, GLA and micronutrients compared to placebo. 26 patients received the nutritional supplement and 29 received a placebo capsule. There was no significant difference between groups in the change in DAS28, SJC, TJC, overall patient/physician VAS scores, grip strength, HAQ, or AIMS scores.

Quality of Evidence: Moderate

Table 37. Nutritional Supplement (Omega-3, Omega-6, micronutrients) vs. Placebo (38)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nutritional Supplement	Placebo at 4 months	Relative (95% CI)	Absolute (95% CI)		
Change in Tender Joint Count												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **1 lower** (3.43 lower to 1.43 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference
Change in Swollen Joint Count												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **1.2 higher** (1.12 lower to 3.52 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference
Change in DAS28												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **0.01 higher** (0.44 lower to 0.46 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference
Change in VAS overall health												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **4 higher** (5.43 lower to 13.43 higher)	⨁⨁⨁◯ Moderate	NOT IMPORTANT No significant difference
Change in VAS patient												

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	26	29	-	MD **9 higher** (0.28 lower to 18.28 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Change in VAS physician

1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **4 lower** (11.08 lower to 3.08 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

Change in Grip Strength (right hand)

1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	26	29	-	MD **13 lower** (37.51 lower to 11.51 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Change in Grip Strength (left hand)

1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	26	29	-	MD **0** (20.48 lower to 20.48 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Change in HAQ

1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **0.12 lower** (0.27 lower to 0.03 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

Change in AIMS

1	randomised trials	not serious	not serious	not serious	serious[a]	none	26	29	-	MD **0.45 higher** (0.31 lower to 1.21 higher)	⊕⊕⊕◯ Moderate	NOT IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

a. Wide CI that crosses 0

b. Very wide CI that crosses both 0 and a high effect threshold

c. Indirect measure of functional status

Comparison: 2.6 g of omega 3 vs. 1.3g Omega 3

Evidence Summary: One double-blind, placebo controlled RCT (Geusens et al) randomized 90 RA patients to 2.6 gm of Omega-3, 1.3 gm of omega 3, or olive oil placebo. All patients received a recommended prescription diet consisting of 30% fat, 12-15% protein, and 50-58% carbohydrates and were instructed to consume fish once weekly. At the end of the 12 month study, pt global assessment of disease activity was slightly lower (1.43 pts lower, 95% CI 0.16-2.88) in the 2.6g omega 3 group, otherwise there were no differences in RA disease activity as assessed by physician global, tender joint count, and Ritchie articular index. Additionally, there were no differences in pain or functional status (assessed by grip strength) between treatment arms. This study was particularly limited by a high (>30%) drop out rate.

Quality of evidence: Very low

Table 38. 2.6 g of omega 3 vs. 1.3g Omega 3 (34)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	2.6 gm omega-3	1.3 gm omega-3	Relative (95% CI)	Absolute (95% CI)		
Physician global assessment												

1	randomised trials	very serious[a,b]	not serious	not serious	very serious[c,d]	none	19	21	-	MD **0.23 lower** (0.79 lower to 0.33 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference in Physician global

Patient global assessment

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	19	21	-	MD **1.43 lower** (2.77 lower to 0.09 lower)	⊕◯◯◯ Very low	CRITICAL **Lower patient global in 2.6 vs 1.3g group**

Patient pain score

1	randomised trials	very serious[a]	not serious	not serious	very serious[c,d]	none	19	21	-	MD **0.15 lower** (0.7 lower to 0.4 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference in pain

Ritchie articular pain index

1	randomised trials	very serious[a]	not serious	not serious	very serious[c,e]	none	19	21	-	MD **5 lower** (14.8 lower to 4.8 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference in RAI

No of painful joints

1	randomised trials	very serious[a]	not serious	not serious	very serious[c,e]	none	19	21	-	MD **2 lower** (7.54 lower to 3.54 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference in TJC

Grip strength

1	randomised trials	very serious[a]	not serious	serious[b]	very serious[c,d]	none	19	21	-	MD **17 higher** (7.79 lower to 41.79 higher)	⊕◯◯◯ Very low	CRITICAL No sig difference in grip strength

CI: confidence interval; **MD:** mean difference

a. high drop out

b. surrogate for functional status

c. single study

d. CI crosses zero

e. Wide CI, crosses zero

Comparison: Omega 3 + Primrose Oil vs. Omega 3

Evidence Summary: One RCT (Veselinovic et al) looked at omega 3 + primrose oil vs. omega 3 alone. While those on omega 3+ primrose oil had a slightly (but statistically significant) lower number of swollen joints, all other measures of disease activity and pain showed no difference. No measures of functional status were reported.

Quality of evidence: Very low

Table 39. Omega 3 + Primrose Oil vs. Omega 3 (20)

.

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Omega 3 FA + Primrose Oil	Omega 3	Relative (95% CI)	Absolute (95% CI)		
DAS28, 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	20	20	-	MD **0.12 lower** (0.59 lower to 0.35 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference in DAS28.
CRP, 12 weeks												
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	20	20	-	MD **0.2 lower** (2.92 lower to 2.52 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference in CRP.
Tender joint count, 12 weeks												
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	20	20	-	MD **0.7 higher** (0.2 lower to 1.6 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference in TJC.
Swollen joint count, 12 weeks												
1	randomised trials	serious[a]	not serious	serious[c]	not serious	none	20	20	-	MD **0.5 lower** (0.87 lower to 0.13 lower)	⨁⨁◯◯ Low	CRITICAL No significant difference in SJC.
Pain VAS, 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Omega 3 FA + Primrose Oil	Omega 3	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	20	20	-	MD **3.8 higher** (0.57 lower to 8.17 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference in pain VAS.

ESR, 12 weeks

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	20	20	-	MD **3.3 lower** (11.98 lower to 5.38 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference in ESR.

CI: confidence interval; **MD:** mean difference
a. 2 unclear risk, one high risk

b. Crosses 0 (no effect threshold)

c. surrogate for disease activity

Comparison: Fish oil vs. Olive oil

Evidence Summary: One RCT (Kremer) looked at high fish oil vs olive oil and low fish oil vs. olive oil. While grip strength improved for patients on high fish oil vs olive oil, there was no difference in any of the other measures including tender joint count, swollen joint count, morning stiffness, fatigue, patient global health, patient pain, physician global and physician pain. There was no difference in any measures of pain, disease activity or function for low fish oil vs. olive oil. This study had a low sample size. An additional RCT (Cleland) also found that fish oil supplementation resulted in lower tender joint count, but not swollen joint count, walk time, or morning stiffness compared to olive oil.

Quality of evidence: Very low

Table 40. High fish oil vs. Olive oil (39)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High fish oil	olive oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
Tender joint count 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	17	12	-	MD **2.1 lower** (4.83 lower to 0.63 higher)	⨁◯◯◯ Very low	CRITICAL No difference in TJC.
Joint swelling count 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	17	12	-	MD **0.4 lower** (4.05 lower to 3.25 higher)	⨁◯◯◯ Very low	CRITICAL No difference in sJC.
Morning stiffness (mins) 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	17	12	-	MD **45.3 lower** (91.61 lower to 1.01 higher)	⨁◯◯◯ Very low	CRITICAL No difference in morning stiffness.
Interval to onset of fatigue (hrs) 24 weeks												
1	randomised trials	serious[a]	not serious	serious[d]	serious[c]	none	17	12	-	MD **1.1 higher** (1 lower to 3.2 higher)	⨁◯◯◯ Very low	CRITICAL No difference in fatigue.

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High fish oil	olive oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
Grip strength 24 weeks												
1	randomised trials	serious[a]	not serious	serious[e]	serious[f]	none	17	12	-	MD **18.2 higher** (1.03 higher to 35.37 higher)	⨁◯◯◯ Very low	CRITICAL **Significantly higher grip strength in high fish oil vs. olive oil.**
Patient evaluation of pain 24 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	17	12	-	MD **0.1 higher** (0.41 lower to 0.61 higher)	⨁⨁◯◯ Low	CRITICAL No difference in pain score.
Patient evaluation of global disease 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	17	12	-	MD **0.3 higher** (0.16 lower to 0.76 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global score.
Physician evaluation of pain 24 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High fish oil	olive oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[g]	not serious	none	17	12	-	MD **0.5 lower** (0.95 lower to 0.05 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower physician pain score in the high fish oil group.**

Physician evaluation of global disease 24 weeks

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	17	12	-	MD **0.1 lower** (0.68 lower to 0.48 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician global.

CI: confidence interval; **MD:** mean difference

a. 4 unclear ratings

b. Surrogate for disease activity

c. Crosses 0

d. Surrogate for functional status

e. Surrogate for functional status

f. Very wide CI

g. Surrogate for pain

Table 41. Low fish oil vs. Olive oil (39)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Low fish oil	olive oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
Tender joint count 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	12	-	MD **2.3 lower** (5.24 lower to 0.64 higher)	⨁◯◯◯ Very low	CRITICAL No difference in TJC.
Joint swelling count 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	12	-	MD **1.7 lower** (5.71 lower to 2.31 higher)	⨁◯◯◯ Very low	CRITICAL No difference in SJC.
Morning stiffness (mins) 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	very serious[d]	none	20	12	-	MD **3.6 lower** (53.07 lower to 45.87 higher)	⨁◯◯◯ Very low	CRITICAL No difference in morning stiffness.
Interval to onset of fatigue (hrs) 24 weeks												
1	randomised trials	serious[a]	not serious	serious[e]	serious[c]	none	20	12	-	MD **0.9 higher** (1.21 lower to 3.01 higher)	⨁◯◯◯ Very low	CRITICAL No difference in fatigue.
Grip strength 24 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Low fish oil	olive oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	very serious[e]	serious[c]	none	20	12	-	MD **9.1 higher** (5.49 lower to 23.69 higher)	⨁◯◯◯ Very low	CRITICAL No difference in grip strength.

Patient evaluation of pain 24 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	20	12	-	MD **0.1 higher** (0.42 lower to 0.62 higher)	⨁⨁◯◯ Low	CRITICAL No difference in pain score.

Patient evaluation of global disease 24 weeks

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	12	-	MD **0** (0.39 lower to 0.39 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global.

Physician evaluation of pain 24 weeks

1	randomised trials	serious[a]	not serious	serious[f]	serious[c]	none	20	12	-	MD **0.4 lower** (0.85 lower to 0.05 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician pain.

Physician evaluation of global disease 24 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Low fish oil	olive oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	20	12	-	MD **0** (0.46 lower to 0.46 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician global.

CI: confidence interval; **MD:** mean difference
a. 4 categories of unclear risk

b. Surrogate for disease activity

c. Cross 0

d. Crosses 0 and wide CI

e. Surrogate for functional status

f. Surrogate for pain

Table 42. Additional data on Fish oil vs. Olive oil (31)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results

884, Cleland 1988	RCT	12 weeks	46 RA patients, 23 in each arm (14 dropped out)	Fish oil supplementation vs olive oil control	Tender joint count, 12 weeks: -Fish oil (mean 9.5, range 1-31; *p=0.01 for paired t-test vs baseline TJC*) -Control (mean 12, range 0-41) Swollen joint count, 12 weeks: -Fish oil (mean 3.6, range 0-9) -Control (mean 3.5, range 0-12) 15-meter walk time (sec), 12 weeks: -Fish oil (mean 17, range 9-28) -Control (mean 17, range 11-33) Morning stiffness (min), 12 weeks: -Fish oil (mean 25, range 0-120) -Control (mean 38, range 0-180)

Comparison: High fish oil vs. Low fish oil

Evidence Summary: One RTC (Kremer) looked at high fish oil vs low fish oil. There was no difference in any measures of pain, disease activity or function.

Quality of evidence: Very low

Table 43. High fish oil vs. Low fish oil (39)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High fish oil	low fish oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
Tender joint count 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	very serious[c]	none	17	20	-	MD **0.2 higher** (1.93 lower to 2.33 higher)	⨁◯◯◯ Very low	CRITICAL No difference in TJC.
Joint swelling count 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	very serious[d]	none	17	20	-	MD **1.3 higher** (2.02 lower to 4.62 higher)	⨁◯◯◯ Very low	CRITICAL No difference in SJC.
Morning stiffness (mins) 24 weeks												
1	randomised trials	serious[a]	not serious	serious[b]	extremely serious[d]	none	17	20	-	MD **41.7 lower** (84.37 lower to 0.97 higher)	⨁◯◯◯ Very low	CRITICAL No difference in morning stiffness.
Interval to onset of fatigue (hrs) 24 weeks												
1	randomised trials	serious[a]	not serious	serious[e]	very serious[d]	none	17	20	-	MD **0.2 higher** (1.38 lower to 1.78 higher)	⨁◯◯◯ Very low	CRITICAL No difference in fatigue.
Grip strength 24 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High fish oil	low fish oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	very serious[a]	very serious[c]	none	17	20	-	MD **9.1 higher** (8.16 lower to 26.36 higher)	⨁◯◯◯ Very low	CRITICAL No difference in grip strength.

Patient evaluation of pain 24 weeks

1	randomised trials	serious[a]	not serious	not serious	extremely serious[c]	none	17	20	-	MD **0** (0.53 lower to 0.53 higher)	⨁◯◯◯ Very low	CRITICAL No difference in pain.

Patient evaluation of global disease 24 weeks

1	randomised trials	serious[a]	not serious	serious[b]	extremely serious[c]	none	17	20	-	MD **0.3 higher** (0.16 lower to 0.76 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global.

Physician evaluation of pain 24 weeks

1	randomised trials	serious[a]	not serious	serious[b]	extremely serious[c]	none	17	20	-	MD **0.1 lower** (0.49 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician pain score.

Physician evaluation of global disease 24 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High fish oil	low fish oil at 24 weeks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	extremely serious[c]	none	17	20	-	MD **0.1 lower** (0.63 lower to 0.43 higher)	⨁◯◯◯ Very low	CRITICAL No difference in physician global.

CI: confidence interval; **MD:** mean difference
a. A number of "unclear" biases, and "high risk" for attrition rate

b. Surrogate for disease activity

c. Crosses 0 (no-effect threshold)

d. Crosses 0 (no-effect threshold) and large CI

e. Surrogate for functional status

Comparison: Fish oil vs. usual diet

Evidence Summary: One randomized controlled trial (Magaro et al), compared fish oil supplementation to usual diet (with no placebo) in a small sample of 20 women with rheumatoid arthritis. This was an unblinded study and final assessments were made after 45 days. No difference was seen in RA disease activity as assessed by the ritchie articular index, morning stiffness, or ESR, nor was any difference observed in grip strength or pain.

Quality of evidence: Very low

Table 44. Fish oil vs. usual diet (40)

Certainty assessment	№ of patients	Effect	Certainty	Importance

№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fish oil supplementation	Usual diet	Relative (95% CI)	Absolute (95% CI)		
Ritchie index 45 days												
1	randomised trials	Very serious[a]	not serious	not serious	very serious[b,c]	none	10	10	-	MD 3.8 lower (8.74 lower to 1.14 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference in RAI
Grip Strength (mmHg) 45 days												
1	randomised trials	Very serious[a]	not serious	serious[d]	very serious[b,c]	none	10	10	-	MD 3.9 lower (33.6 lower to 25.8 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference in grip strength
Morning Stiffness (min) 45 days												
1	randomised trials	Very serious[a]	not serious	not serious	very serious[b,c]	none	10	10	-	MD 1.9 higher (10.95 lower to 14.75 higher)	⨁◯◯◯ Very low	IMPORTANT No sig difference in AM stiffness
Pain VAS (cm) 45 days												
1	randomised trials	Very serious[a]	not serious	not serious	very serious[b,c]	none	10	10	-	MD 0.6 higher (0.64 lower to 1.84 higher)	⨁◯◯◯ Very low	CRITICAL No sig difference in pain
Erythrocyte Sedimentation Rate (mm/1st hour)												

1	randomised trials	Very serious[a]	not serious	serious[e]	very serious[b,c]	none	10	10	-	MD 6.5 lower (40.68 lower to 27.68 higher)	⊕◯◯◯ Very low	IMPORTANT No sig difference in ESR

CI: confidence interval; MD: mean difference

a. No blinding

b. Single study

c. Wide CI, crosses zero

d. Surrogate for functional status

e. Nonspecific lab surrogate for disease activity

Table 45. Additional data on Fish oil vs. usual diet (41)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
577, Fatel et al., 2021	A randomized, single-blind intervention study	90 days	RA patients: 62 Control group: n =21, Age, median (IQR): 52 (43-63), 3M/18F Fish oil: n = 21, Age, median (IQR): 58 (47-64), 4M/17F Fish oil and cranberry: n = 20, Age, median (IQR): 58 (47-65), 5M/15F	Control group: regular diet, the second group Fish oil group: 3 g/d of fish oil v-3 fatty acids (10 capsules) Cranberry juice + fish oil group: 3 g/d of fish oil v-3 fatty acids and 500 mL/d of reduced-calorie cranberry juice Each fish oil capsule contained 180 mg of eicosapentaenoic acid and 120 mg of docosahexaenoic acid, originating from sardines.	DAS28-CRP: median for fish oil treatment after 90 days was 2.98 (IQR = 2.47 - 3.53, p = 0.045) vs 2.77 (IQR = 2.42 – 3.52, p = NS) for control, indicating a reduction in disease activity in the fish oil group vs the control group.

Comparison: Fatty acid vs. fatty acid + g-linolenic acid

Evidence Summary: One RCT (Dawcynski 2011) compared 1575 mg n-3 LC-PUFA plus 1800 mg GLA/d to 1575 mg n-3 LC-PUFA, as part of a comparison of multiple fatty acid formulations. 13 patients were randomized to the combined fatty acid group and 14 patients were in the n-3 LC-PUFA group. There was no significant difference between groups in the pain VAS score at 12 weeks.

Quality of Evidence: Very Low

Table 46. Fatty acid vs. fatty acid + g-linolenic acid (30)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fatty acid	fatty acid + g-linolenic acid	Relative (95% CI)	Absolute (95% CI)		
VAS 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	14	13	-	MD **9.1 lower** (27.47 lower to 9.27 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

a. Significant differential attrition between supplement and control groups

b. Wide CI that crosses 0 and high effect threshold

Comparison: Flaxseed oil vs. Safflower oil

Evidence Summary: One RCT (Nordstrom 1995) compared flaxseed oil (containing 32% alpha-linolenic acid) to safflower oil (containing 33% linolenic acid). 11 patients were randomized to each of the intervention groups. At 3 months, there was no significant difference between groups in the patient or physician global assessment, pain VAS, functional class, or joint index. The study had a significant limitation in imprecision due to the low number of participants in each group.

Quality of Evidence: Low

Table 47. Flaxseed oil vs. Safflower oil (42)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Flaxseed oil	safflower oil	Relative (95% CI)	Absolute (95% CI)		
Patient global assessment 3 months												
1	randomised trials	not serious	not serious	serious[a]	very serious[b]	none	11	11	-	MD **0.2 higher** (0.41 lower to 0.81 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
Global assessment physician 3 months												
1	randomised trials	not serious	not serious	serious[a]	very serious[b]	none	11	11	-	MD **0.4 higher** (0.21 lower to 1.01 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
Functional class												
1	randomised trials	not serious	not serious	not serious	serious[c]	none	11	11	-	MD **0** (0.45 lower to 0.45 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference
Joint score index 3 months												

1	randomised trials	not serious	not serious	not serious	serious[c]	none	11	11	-	MD **0.4 lower** (5.51 lower to 4.71 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

Pain VAS 3 months

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	11	11	-	MD **0.6 lower** (2.94 lower to 1.74 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

a. Indirect measure of disease activity

b. Very wide CI that crosses 0 and high effect threshold

c. Wide CI that crosses 0

Comparison: Flaxseed vs. wheat (control)

Evidence Summary: One RCT (Ghaseminasab et al) compared flaxseed supplementation to wheat (control) in a 12 week, single-blind randomized controlled trial. They found beneficial effects in several outcomes, including a significantly lower DAS28, HAQ-DI, pain, and several quality of life measures included in the SF-36 questionairre. The trial was limited in that it was small (N=40 pts in either arm), was single-blind which in particular would affect interpretability of the DAS28 outcome measure, and there were significant differential changes in several food groups between the treatment arms making the true effect of flaxseed supplementation less interpretable.

Quality of Evidence: Low

Table 48. Flaxseed vs. Wheat (43)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Flaxseed	wheat (control)	Relative (95% CI)	Absolute (95% CI)		

Change in DAS28, 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Flaxseed	wheat (control)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	40	40	-	MD **0.63 lower** (1.05 lower to 0.21 lower)	⨁⨁◯◯ Low	CRITICAL **Significantly lower DAS28 with flaxseed**

Change in HAQ-DI, 12 weeks

1	randomised trials	serious[b]	not serious	not serious	not serious	none	40	40	-	MD **0.53 lower** (0.76 lower to 0.3 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Significantly lower HAQ-DI with flaxseed**

Change in Pain, 12 weeks

1	randomised trials	serious[b]	not serious	not serious	not serious	none	40	40	-	MD **2.42 lower** (3.39 lower to 1.45 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Significantly lower pain with flaxseed**

Change in AM stiffness (min), 12 weeks

1	randomised trials	serious[b]	not serious	not serious	serious[c]	none	40	40	-	MD **19.66 lower** (42.27 lower to 2.95 higher)	⨁⨁◯◯ Low	IMPORTANT No Significant diffference

Change in Overall health (SF-36), 12 weeks

1	randomised trials	serious[b]	not serious	not serious	not serious	none	40	40	-	MD **23.3 higher** (13.71 higher to 32.89 higher)	⨁⨁⨁◯ Moderate	IMPORTANT **Significantly higher overall health with flaxseed**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Flaxseed	wheat (control)	Relative (95% CI)	Absolute (95% CI)		
Change in Physical Function (SF-36), 12 weeks												
1	randomised trials	serious[b]	not serious	not serious	not serious	none	40	40	-	MD **29.8 higher** (19.04 higher to 40.56 higher)	⨁⨁⨁◯ Moderate	IMPORTANT **Significantly higher physical function with flaxseed**
Change in Emotional Well-Being (SF-36), 12 weeks												
1	randomised trials	serious[b]	not serious	not serious	not serious	none	40	40	-	MD **13.2 higher** (6.17 higher to 20.23 higher)	⨁⨁⨁◯ Moderate	IMPORTANT **Significantly higher emotional well being with flaxseed**
Change in Mental Health (SF-36), 12 weeks												
1	randomised trials	serious[b]	not serious	not serious	not serious	none	40	40	-	MD **18 higher** (9.32 higher to 26.68 higher)	⨁⨁⨁◯ Moderate	IMPORTANT **Significantly higher mental health with flaxseed**

CI: confidence interval; **MD:** mean difference

Explanations

a. Two high risk categories (unblinded assessors and other)

b. One applicable high risk category (other)

c. CI crosses zero

Comparison: Primrose oil versus stinging nettle

Evidence Summary: One RCT (Abd-Nikfarjam et al) compared primrose oil containing 420 mg of gamma-linolenic acid to stinging nettle or placebo. There was no significant difference between primrose oil and stinging nettle in DAS-28-ESR, patient global VAS, CRP, or ESR at 3 months. The study suffered from significant attrition and lack of intent-to-treat analysis of the patients lost to followup, as well as potential unblinding as the stinging nettle and primrose oil supplements were formulated differently and participants may have been able to recognize the difference.

Quality of Evidence: Very Low

Table 49. Primrose oil versus stinging nettle (19)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Primrose oil	stinging nettle	Relative (95% CI)	Absolute (95% CI)		
DAS-28-ESR (3 months)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	29	31	-	MD **0.07 lower** (0.52 lower to 0.38 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
Patient global VAS (3 months)												
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	29	31	-	MD **0.8 lower** (2.2 lower to 0.6 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
CRP (3 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Primrose oil	stinging nettle	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	29	31	-	MD **0.9 lower** (2.85 lower to 1.05 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

ESR (3 months)

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	29	29	-	MD **1.11 lower** (5.82 lower to 3.6 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Risk of participant unblinding due to different supplement formulations. High attrition with no intent-to-treat analysis (only 90 patients left at final endpoint were analyzed)

b. Wide CI crossing zero

c. Not a direct measure of disease activity

Comparison: N-acetylcysteine vs. placebo

Evidence Summary: One RCT (Jamali et al) compared NAC to placebo. Although the study found lower tender joint count, ESR, and DAS28-ESR at week 8, the study was very small, there was a high rate of attrition, and only those who completed the study were analzyzed.

Quality of Evidence: Low

Table 50. NAC vs. placebo(44)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	N-acetylcysteine	placebo	Relative (95% CI)	Absolute (95% CI)		
Number of tender joints, 8 weeks												
1	randomised trials	Very serious[a]	not serious	serious[b]	not serious	none	22	19	-	MD **2.52 lower** (4.26 lower to 0.78 lower)	⨁⨁◯◯ Low	CRITICAL **Number of tender joints lower in the NAC group vs. placebo.**
Number of swollen joints, 8 weeks												
1	randomised trials	Very serious[a]	not serious	serious[b]	serious[c]	none	22	19	-	MD **0.34 lower** (0.83 lower to 0.15 higher)	⨁◯◯◯ Very low	CRITIAL No difference in swollen joints.
Patient global, 8 weeks												
1	randomised trials	Very serious[a]	not serious	serious[b]	serious[c]	none	22	19	-	MD **9.02 lower** (18.88 lower to 0.84 higher)	⨁◯◯◯ Very low	CRITICAL No difference in patient global.
ESR, 8 week												
1	randomised trials	Very serious[a]	not serious	serious[b]	not serious	none	22	19	-	MD **8.29 lower** (14.87 lower to 1.71 lower)	⨁⨁◯◯ Low	CRITICAL **Lower ESR in the NAC group vs placebo.**
DAS28-ESR, 8 week												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	N-acetylcysteine	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	Very serious[a]	not serious	not serious	not serious	none	22	19	-	MD **1.06 lower** (1.53 lower to 0.59 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Lower DAS28-ESR in the NAC group.**

CI: confidence interval; **MD:** mean difference

Explanations

a. High rate of attirtion with per protocol analysis

b. Surrogate for disease activity

c. Crosses no effect threshold

Comparison: Stinging nettle versus placebo

Evidence Summary: One RCT (Abd-Nikfarjam et al) compared stinging nettle to primrose oil containing 420 mg of gamma-linolenic acid or placebo. The study found a significantly lower DAS-28-ESR and CRP at 3 months, but no significant difference in patient global VAS or ESR. The study suffered from significant attrition and lack of intent-to-treat analysis of the patients lost to followup, as well as potential unblinding as the stinging nettle and primrose oil supplements were formulated differently and participants may have been able to recognize the difference.

Quality of Evidence: Low

Table 51. stinging nettle vs. placebo(19)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stinging nettle	placebo	Relative (95% CI)	Absolute (95% CI)		
DAS-28-ESR (3 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stinging nettle	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	not serious	none	31	30	-	MD **0.68 lower** (1.19 lower to 0.17 lower)	⨁⨁⨁◯ Moderate	CRITICAL Significant difference in favor of stinging nettle

Patient global VAS (3 months)

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	31	30	-	MD **0.91 higher** (0.58 lower to 2.4 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

CRP (3 months)

1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	31	30	-	MD **2.85 lower** (5.55 lower to 0.15 lower)	⨁⨁◯◯ Low	CRITICAL Significant difference in favor of stinging nettle

ESR (3 months)

1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	29	30	-	MD **4.03 lower** (9.02 lower to 0.96 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Risk of participant unblinding due to different supplement formulations. High attrition with no intent-to-treat analysis (only 90 patients left at final endpoint were analyzed)

b. Not a direct measure of disease activity

c. Wide CI crossing zero

Bibliography

1. Salesi M, Farajzadegan Z. Efficacy of vitamin D in patients with active rheumatoid arthritis receiving methotrexate therapy. Rheumatol Int. 2012;32(7):2129-33.
2. Soubrier M, Lambert C, Combe B, Gaudin P, Thomas T, Sibilia J, et al. A randomised, double-blind, placebo-controlled study assessing the efficacy of high doses of vitamin D on functional disability in patients with rheumatoid arthritis. Clin Exp Rheumatol. 2018;36(6):1056-60.
3. Yang J, Liu L, Zhang Q, Li M, Wang J. Effect of vitamin D on the recurrence rate of rheumatoid arthritis. Exp Ther Med. 2015;10(5):1812-6.
4. Li C, Yin S, Yin H, Cao L, Zhang T, Wang Y. Efficacy and Safety of 22-Oxa-Calcitriol in Patients with Rheumatoid Arthritis: A Phase II Trial. Med Sci Monit. 2018;24:9127-35.
5. Chawla H, Gupta G, HO A, G S, A K. Correlation of vitamin D levels and pain in rheumatoid arthritis. . Natl J Physiol Pharm Pharmacol. 2022;12:210-4.
6. Wu J, Dong J, Li S, Luo J, Zhang Y, Liu H, et al. The Role of Vitamin D in Combination Treatment for Patients With Rheumatoid Arthritis. Front Med (Lausanne). 2020;7:312.
7. Gopinath K, Danda D. Supplementation of 1,25 dihydroxy vitamin D3 in patients with treatment naive early rheumatoid arthritis: a randomised controlled trial. Int J Rheum Dis. 2011;14(4):332-9.
8. Peretz A, Siderova V, Neve J. Selenium supplementation in rheumatoid arthritis investigated in a double blind, placebo-controlled trial. Scand J Rheumatol. 2001;30(4):208-12.
9. Tarp U, Overvad K, Thorling EB, Graudal H, Hansen JC. Selenium treatment in rheumatoid arthritis. Scand J Rheumatol. 1985;14(4):364-8.
10. Aryaeian N, Shahram F, Mahmoudi M, Tavakoli H, Yousefi B, Arablou T, et al. The effect of ginger supplementation on some immunity and inflammation intermediate genes expression in patients with active Rheumatoid Arthritis. Gene. 2019;698:179-85.
11. Hatakka K, Martio J, Korpela M, Herranen M, Poussa T, Laasanen T, et al. Effects of probiotic therapy on the activity and activation of mild rheumatoid arthritis--a pilot study. Scand J Rheumatol. 2003;32(4):211-5.
12. Pineda Mde L, Thompson SF, Summers K, de Leon F, Pope J, Reid G. A randomized, double-blinded, placebo-controlled pilot study of probiotics in active rheumatoid arthritis. Med Sci Monit. 2011;17(6):CR347-54.
13. Mandel DR, Eichas K, Holmes J. Bacillus coagulans: a viable adjunct therapy for relieving symptoms of rheumatoid arthritis according to a randomized, controlled trial. BMC Complement Altern Med. 2010;10:1.
14. Cannarella LAT, Mari NL, Alcantara CC, Iryioda TMV, Costa NT, Oliveira SR, et al. Mixture of probiotics reduces inflammatory biomarkers and improves the oxidative/nitrosative profile in people with rheumatoid arthritis. Nutrition. 2021;89:111282.
15. Vaghef-Mehrabany E, Alipour B, Homayouni-Rad A, Sharif SK, Asghari-Jafarabadi M, Zavvari S. Probiotic supplementation improves inflammatory status in patients with rheumatoid arthritis. Nutrition. 2014;30(4):430-5.

16. Nakamura H, Masuko K, Yudoh K, Kato T, Kamada T, Kawahara T. Effects of glucosamine administration on patients with rheumatoid arthritis. Rheumatol Int. 2007;27(3):213-8.
17. Aryaeian N, Shahram F, Djalali M, Eshragian MR, Djazayeri A, Sarrafnejad A, et al. Effect of conjugated linoleic acids, vitamin E and their combination on the clinical outcome of Iranian adults with active rheumatoid arthritis. Int J Rheum Dis. 2009;12(1):20-8.
18. Kumar P SA, Ho M, Maple C, Radederstoff D, Morley K, et al. The Effects of Borage Oil Supplementation on Non-Steroidal Anti-Inflammatory Drug Requirements in Patients with Rheumatoid Arthritis. Journal of Complementary and Integrative Medicine. 2008;5.
19. Abd-Nikfarjam B, Abbasi M, Memarzadeh M, Farzam S, Jamshidian A, Dolati-Somarin A. Therapeutic Efficacy of Urtica dioica and Evening Primrose in Patients with Rheumatoid Arthritis: A Randomized Double-Blind, Placebo-Controlled Clinical Trial [PREPRINT Version 1]. Rsearch Square. 2021.
20. Veselinovic M, Vasiljevic D, Vucic V, Arsic A, Petrovic S, Tomic-Lucic A, et al. Clinical Benefits of n-3 PUFA and -Linolenic Acid in Patients with Rheumatoid Arthritis. Nutrients. 2017;9(4).
21. Belch JJ, Ansell D, Madhok R, O'Dowd A, Sturrock RD. Effects of altering dietary essential fatty acids on requirements for non-steroidal anti-inflammatory drugs in patients with rheumatoid arthritis: a double blind placebo controlled study. Ann Rheum Dis. 1988;47(2):96-104.
22. Berbert AA, Kondo CR, Almendra CL, Matsuo T, Dichi I. Supplementation of fish oil and olive oil in patients with rheumatoid arthritis. Nutrition. 2005;21(2):131-6.
23. Kremer JM, Lawrence DA, Petrillo GF, Litts LL, Mullaly PM, Rynes RI, et al. Effects of high-dose fish oil on rheumatoid arthritis after stopping nonsteroidal antiinflammatory drugs. Clinical and immune correlates. Arthritis Rheum. 1995;38(8):1107-14.
24. Nielsen GL, Faarvang KL, Thomsen BS, Teglbjaerg KL, Jensen LT, Hansen TM, et al. The effects of dietary supplementation with n-3 polyunsaturated fatty acids in patients with rheumatoid arthritis: a randomized, double blind trial. Eur J Clin Invest. 1992;22(10):687-91.
25. Skoldstam L, Borjesson O, Kjallman A, Seiving B, Akesson B. Effect of six months of fish oil supplementation in stable rheumatoid arthritis. A double-blind, controlled study. Scand J Rheumatol. 1992;21(4):178-85.
26. Tulleken JE, Limburg PC, Muskiet FA, van Rijswijk MH. Vitamin E status during dietary fish oil supplementation in rheumatoid arthritis. Arthritis Rheum. 1990;33(9):1416-9.
27. Kremer JM, Bigauoette J, Michalek AV, Timchalk MA, Lininger L, Rynes RI, et al. Effects of manipulation of dietary fatty acids on clinical manifestations of rheumatoid arthritis. Lancet. 1985;1(8422):184-7.
28. Park Y, Lee A, Shim SC, Lee JH, Choe JY, Ahn H, et al. Effect of n-3 polyunsaturated fatty acid supplementation in patients with rheumatoid arthritis: a 16-week randomized, double-blind, placebo-controlled, parallel-design multicenter study in Korea. J Nutr Biochem. 2013;24(7):1367-72.
29. Rajaei E, Mowla K, Ghorbani A, Bahadoram S, Bahadoram M, Dargahi-Malamir M. The Effect of Omega-3 Fatty Acids in Patients With Active Rheumatoid Arthritis Receiving DMARDs Therapy: Double-Blind Randomized Controlled Trial. Glob J Health Sci. 2015;8(7):18-25.
30. Dawczynski C, Hackermeier U, Viehweger M, Stange R, Springer M, Jahreis G. Incorporation of n-3 PUFA and gamma-linolenic acid in blood lipids and red blood cell lipids together with their influence on disease activity in patients with chronic inflammatory arthritis--a randomized controlled human intervention trial. Lipids Health Dis. 2011;10:130.

31. Cleland LG, French JK, Betts WH, Murphy GA, Elliott MJ. Clinical and biochemical effects of dietary fish oil supplements in rheumatoid arthritis. J Rheumatol. 1988;15(10):1471-5.
32. Galarraga B, Ho M, Youssef HM, Hill A, McMahon H, Hall C, et al. Cod liver oil (n-3 fatty acids) as an non-steroidal anti-inflammatory drug sparing agent in rheumatoid arthritis. Rheumatology (Oxford). 2008;47(5):665-9.
33. Volker D, Fitzgerald P, Major G, Garg M. Efficacy of fish oil concentrate in the treatment of rheumatoid arthritis. J Rheumatol. 2000;27(10):2343-6.
34. Geusens P, Wouters C, Nijs J, Jiang Y, Dequeker J. Long-term effect of omega-3 fatty acid supplementation in active rheumatoid arthritis. A 12-month, double-blind, controlled study. Arthritis Rheum. 1994;37(6):824-9.
35. Kjeldsen-Kragh J, Lund JA, Riise T, Finnanger B, Haaland K, Finstad R, et al. Dietary omega-3 fatty acid supplementation and naproxen treatment in patients with rheumatoid arthritis. J Rheumatol. 1992;19(10):1531-6.
36. Dawczynski C, Schubert R, Hein G, Muller A, Eidner T, Vogelsang H, et al. Long-term moderate intervention with n-3 long-chain PUFA-supplemented dairy products: effects on pathophysiological biomarkers in patients with rheumatoid arthritis. Br J Nutr. 2009;101(10):1517-26.
37. Espersen GT, Grunnet N, Lervang HH, Nielsen GL, Thomsen BS, Faarvang KL, et al. Decreased interleukin-1 beta levels in plasma from rheumatoid arthritis patients after dietary supplementation with n-3 polyunsaturated fatty acids. Clin Rheumatol. 1992;11(3):393-5.
38. Remans PH, Sont JK, Wagenaar LW, Wouters-Wesseling W, Zuijderduin WM, Jongma A, et al. Nutrient supplementation with polyunsaturated fatty acids and micronutrients in rheumatoid arthritis: clinical and biochemical effects. Eur J Clin Nutr. 2004;58(6):839-45.
39. Kremer JM, Lawrence DA, Jubiz W, DiGiacomo R, Rynes R, Bartholomew LE, et al. Dietary fish oil and olive oil supplementation in patients with rheumatoid arthritis. Clinical and immunologic effects. Arthritis Rheum. 1990;33(6):810-20.
40. Magaro M, Zoli A, Altomonte L, Mirone L, De Sole P, Di Mario G, et al. Effect of fish oil on neutrophil chemiluminescence induced by different stimuli in patients with rheumatoid arthritis. Ann Rheum Dis. 1992;51(7):877-80.
41. Fatel ECS, Rosa FT, Alfieri DF, Flauzino T, Scavuzzi BM, Lozovoy MAB, et al. Beneficial effects of fish oil and cranberry juice on disease activity and inflammatory biomarkers in people with rheumatoid arthritis. Nutrition. 2021;86:111183.
42. Nordstrom DC, Honkanen VE, Nasu Y, Antila E, Friman C, Konttinen YT. Alpha-linolenic acid in the treatment of rheumatoid arthritis. A double-blind, placebo-controlled and randomized study: flaxseed vs. safflower seed. Rheumatol Int. 1995;14(6):231-4.
43. Ghaseminasab-Parizi M, Nazarinia MA, Akhlaghi M. The effect of flaxseed with or without anti-inflammatory diet in patients with rheumatoid arthritis, a randomized controlled trial. Eur J Nutr. 2022;61(3):1377-89.
44. Jamali F, Ahmadzadeh A, Sahraei Z, Salamzadeh J. Study of the Effects of N-acetylcysteine on Inflammatory Biomarkers and Disease Activity Score in Patients with Rheumatoid Arthritis. Iran J Allergy Asthma Immunol. 2021;20(5):574-83.

PICO 3: Should patients with RA who are overweight or obese receive a weight loss intervention?

Summary: Literature searches identified one randomized controlled trial (RCT)[**Error! Reference source not found.**]

Somers et al.[**Error! Reference source not found.**] randomized 50 participants to receive enhanced lifestyle behavioral weight management (n = 29) or standard care of RA (n = 21). Patients had to (1) have met the American College of Rheumatology criteria for RA, (2) have obesity (defined as BMI >28 kg/m2), (3) have had RA for at least 2 years, and (3) have self-reported RA pain in the last 2 weeks. Treatment group participated in an enhanced lifestyle behavioral weight management intervention that included instruction in pain coping skills and traditional behavioral weight management strategies. This program was delivered weekly over 12 weeks in 90-minute group sessions. The intervention followed a session-by-session, manualized protocol and was delivered by clinical psychologists with prior experience in PCST. The control group was designed to serve as a standard care/usual treatment comparison group. Patients assigned to this condition continued to receive their routine RA care, including regular appointments with their rheumatologist and appointments as needed for symptom flares or other RA-related problems.

There was no difference in mean of the outcomes at follow-up (after the end of the program which occurred at 12 weeks) when comparing the intervention and the control for disease activity as measured by disease severity (VAS), physical functioning as measured by AIMS), physical functioning as measured by the 6MWT, self-efficacy for weight loss, and self-efficacy for arthritis (Table 1).

Quality of evidence across all critical outcomes: Very low

Table 1. Data from Randomized Controlled Trials

Certainty							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Enhanced Lifestyle Behavioral WeightManagement Group	Standard Care Group	Relative (95% CI)	Absolute (95% CI)		
Disease activity as measured by disease severity (VAS) (0 to 100 scale), 12 weeks												
1	randomised trials	not serious	not serious	very serious[a,b]	serious[c]	none	26	14	-	MD **1.01 lower** (8.65 lower to 6.63 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Physical Functioning: AIMS2 (0 to 10 scale), 12 weeks												

Certainty							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Enhanced Lifestyle Behavioral WeightManagement Group	Standard Care Group	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[c]	none	24	14	-	MD **0.73 lower** (1.96 lower to 0.5 higher)	⨁⨁◯◯ low	CRITICAL No statistically significant difference

Function as inferred from 6MWT (meters walked in six minutes), 12 weeks

1	randomised trials	not serious	not serious	very serious[a,b]	serious[c]	none	24	14	-	MD **48.02 higher** (2.86 lower to 98.9 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference

Self-efficacy for Weight Loss (0 to 9 scale), 12 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[c]	none	24	14	-	MD **0.73 higher** (0.43 lower to 1.89 higher)	⨁⨁◯◯ low	IMPORTANT No statistically significant difference

Self-efficacy Arthritis (10 to 100 scale), 12 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[c]	none	24	14	-	MD **4.59 higher** (7.97 lower to 17.15 higher)	⨁⨁◯◯ low	IMPORTANT No statistically significant difference

CI: confidence interval; MD: mean difference, 6MWT: six-minute walk test

Explanations

a. Enrolled patients are not typical, due to three patient enrollment criteria ((1) have met the American College of Rheumatology criteria for RA, (2) have obesity (BMI >28 kg/m2), (3) have had RA for at least 2 years, and (3) have self-reported RA pain in the last 2 weeks).

b. Surrogate measure of disease activity (disease severity (VAS)) or function (6MWT).

c. Small sample size

References:

1. Somers TJ, Blumenthal JA, Dorfman CS, Huffman KM, Edmond SN, Miller SN, Wren AA, Caldwell D, Keefe FJ. Effects of a Weight and Pain Management Program in Patients With Rheumatoid Arthritis With Obesity: A Randomized Controlled Pilot Investigation. J Clin Rheumatol. 2022 Jan 1;28(1):7-13. doi: 10.1097/RHU.0000000000001793. PMID: 34670994.

Exercise

PICO 4: Should patients with RA consistently engage in an aerobic exercise program?

Summary: Literature searches identified 26 controlled trials (1-29) and 4 observational studies (30-34) addressing this PICO question. All observational studies were prospective cohorts, two studies specified some level of matching (31, 32, 34), whereas no matching was specified for one (33), and groups were allocated by proximity to the intervention for the other (30). Studies included a wide range of outcome measures including generic QOL (e.g., SF-36, EQ-5D), disease-related QOL (HAQ, AIMS/AIMS2), domains of health (pain, fatigue, function, mental health/depression), disease activity (DAS4, ESR), and performance-based measures (e.g. step-up, walk test, etc.). Individual studies of aerobic exercise reported significant improvements across multiple outcomes including HAQ (4, 14, 34), RADAI (14), DAS-28(17), CES-D (29), fatigue severity scale (8), Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaire (29), pain (2, 4, 8), and stiffness (8).

Studies comparing aerobic treatment versus inactive comparators using disease-related QOL outcomes showed mixed results. Overall merged effect significantly favored the aerobic arm using the HAQ for studies with >12-week follow-up, but not for studies with <12-week follow-up.(2, 29) Studies comparing aerobic treatment versus an active treatment arm showed no significant improvement using HAQ

RCT discussion:

10 RCTs assessed active aerobic physical activity treatment arms versus an inactive control. They were equally subdivided into studies (or treatment arms) with a largely self-management focus (8, 12, 14, 19, 22) and studies (or treatment arms) using a more traditional structured exercise format (3, 6, 7, 9, 19, 24, 28).

<u>Infrequent in-person contact/accountability (self-management focus)</u>

Home exercise versus UC (8), Pedometer, Pedometer + (14), online individualized training versus access to web pages(12, 23), home exercise using video versus control (19)

<u>Frequent in-person contact/accountability (traditional structured exercise format)</u>

Sensorimotor exercise(4), aerobic exercise(24), high intensity weight bearing program – including aerobic training (6, 7), dynamic exercise versus general physical activity recommendations(9), multidisciplinary group rehab versus waitlist(3), class exercise versus control (19), personalized exercise versus standard of care(28)

Few outcomes were assessed by enough studies to separately sub-group based on intervention format.

Inactive comparators >12 weeks:

Disease-related QOL –

HAQ was used by 9 studies assessing aerobic interventions versus inactive comparators (control, education, etc.)(4, 6-8, 12, 14, 22, 24, 28).

Studies produced mean estimates that universally favored the aerobic exercise intervention arm however half of the studies reported confidence intervals that crossed the line of no effect. Test for overall effect showed a significant improvement favoring the aerobic arm across studies (p=0.002). Studies were further subdivided based on self-management versus traditional exercise format focus. This showed that studies/treatment arms classified as having a self-management focus(8, 12, 14, 22, 35) showed significant overall effect whereas the studies/treatment arms classified as having a traditional exercise structure(4, 6, 7, 24) did not. Brodin, et al. also reported improvements in HAQ for the aerobic exercise arm (p-value 0.026), however they reported point estimates and variance in a manner that did not allow inclusion into the meta-analysis (28).

Function – Self report

SF-36 function domain

Four RCTs(4, 9, 12, 22) reported physical function using the SF-36 physical function summary measure, two studies with a self-management focus(12, 22) and two with a traditional exercise format(4, 9). All mean estimates reported improvements over inactive comparators but with

levels of uncertainty that crossed the line of no effect for all but one study (4). All merged estimates showed no significant improvements relative to control.

Other studies assessed function using different instruments and showing mixed results including ASES function(3) in favor of aerobic exercise (P=0.05) and MACTAR(6, 7, 22) reporting overall effect favoring aerobic activity (p=0.007), but no effect shown using AIMS2(3) Physical health (95% CI crossed the line of no effect).

Function – Performance-based outcomes

Sensorimotor rehabilitation produced a significant improvement in the TUG (4), and 50 feet walk time was improved relative to control in both class-based and home-based aerobic exercise but it was not statistically significant(19).

Pain

Three studies reported on some aspect of pain intensity, using VAS(4, 8) and McGill pain intensity scale (19), and another study assessed pain interference (14). Merged effects on pain intensity show a small (<0.5 SMD) but significant improvement with aerobic training relative to inactive controls p=(0.03). Katz, et al. reported mean improvement in pain interference with both arms but neither were significantly improved over control (14).

Inactive comparators <12 weeks

Two studies assessed aerobic exercise versus a control. A community walking program was compared against a control with a 6-week follow-up. Baxter, et al. reported no significant improvements across outcomes assessed. No adverse events were reported.(2) Loeppenthin, et al. compared aerobic exercise versus control and found between-group differences in fatigue -16.1 (95% CI -25.1 to -7.0 , P=0.001) and depressive symptoms -6.8 (95% CI: -12.4 to -1.1, P=0.02) in favor of the aerobic exercise group.

Active comparators >12 weeks

Disease-related QOL

Studies/treatment arms of aerobic exercise versus an active comparator showed mixed results. Lange, et al. reporting mean, but not statistically significant improvement in HAQ at 1 year(16), but significant improvement in HAQ at 4-year follow-up(17). Van den End, et al. reported mean improvement in HAQ when comparing high intensity dynamic group exercise versus home exercise but showed losses in HAQ scores when comparing high intensity dynamic group exercise versus low intensity individual or low intensity group exercise.(23)

Function – Self report

SF-36 – a single study

Dynamic exercise + diet had a small but significant improvement in SF-36 when compared with diet alone, 0.74 [0.32, 1.16](9). Yang assessed a group aerobic exercise program meeting at a frequency of 4 days/week and reported no change in self-reported physical function, 0.02 [-0.35, 0.40] using QOLRA-Physical Function scale(25).

Function – performance

No performance-based outcome estimates of function were significant for any aerobic exercise intervention when assessed against an active comparator. Lange, et al. found a mean improvement in short term TUG(16) but reversal in four year outcome(17). 50ft walk (Van den ende, et al.) walking speed improved in the high intensity group exercise arm relative to home exercise and low intensity-individual arm but showed less improvement when compared with the low intensity-group arm (23). Sit-to-stand (Lange, et al.) showed small mean improvements for both timepoints in the moderate-high intensity exercise arm.(16, 17)

Pain

There were mixed results between high versus lower intensity exercise interventions. Lange, et al. found a high intensity intervention to reduce pain relative to lower intensity comparators (16, 17), whereas Van Den Ende, et al. found the opposite across all treatment arms (23). The high intensity exercise arm reported higher pain than lower intensity comparisons including significantly less pain with low intensity group exercise using VAS (0-100) 13.00 [1.65, 24.35] (23).

Active comparators <12 weeks

Disease-related QOL –

Hsieh, et al (11) and Sanford-Smith, et al (21) both reported mean improvements in HAQ, but Melikoglu, et al. (18) reported a mean decrease in HAQ when assessing aerobic exercise against an active comparator. No differences were significant.

Pain –

Mean, non-significant improvements in pain were found when comparing dynamic exercise versus range of motion exercises (18), and aerobic exercises versus the combination of Pilates and aerobic exercises (26). Significant improvements in pain were found when Pilates only was compared with aerobic exercise using the McGill Pain questionnaire VAS 0.40 [0.24, 0.56] (26).

No self-reported function measures were assessed but performance-based function was assessed using walking time over 50ft and walking distance using the 6 min walk test. No differences in performance-based function were significant across groups. 50 ft Walking time was less for the home exercise versus supervised exercise group (11) and walking distance was greater in the aerobic exercise group over the pilates only group, and in the pilates+aerobic group over the aerobic exercise only group (26).

Cohort discussion is limited to inactive comparators > 12 weeks:

Disease-related QOL

Disease-related QOL was measured in one cohort which reported significant improvement in HAQ for aerobic exercise versus control -0.98 [-1.67, -0.28] (34).

Function

Self reported function was measured in one cohort using AIMS-physical activity, but no significant improvements were reported(33), likewise performance-based outcomes, including walk time, showed improvements that did not pass statistical significance.(31-33)

Pain -

Noreau et al. assessed AIMS-pain and found mean improvements in dance-based therapy that was not significant.(33)

Data table from RCTs aerobic exercise versus inactive comparators >12 week outcomes

Author(s):

Question: RCT aerobic ex compared to inactive comparator > 12wks for health problem or population

Setting:

Bibliography: . PICO4 6019 7169 1064 1778. Cochrane Database of Systematic Reviews [Year], Issue [Issue].

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
HAQ >12 weeks												
7	randomised trials	serious[a]	not serious	not serious	serious[b]	none	432	424	-	MD **0.17 lower** (0.28 lower to 0.06 lower)	⨁⨁◯◯ Low	CRITICAL Overall effect favoring aerobic arm P=0.002
HAQ >12 weeks - low accountability/contact												
4	randomised trials	serious[c]	not serious	not serious	serious[d]	none	234	218	-	MD **0.14 lower** (0.25 lower to 0.02 lower)	⨁⨁◯◯ Low	CRITICAL Overall effect favoring aerobic arm P=0.02
HAQ >12 weeks - High accountability/contact												
3	randomised trials	serious[e]	not serious	not serious	very serious[f]	none	198	206	-	MD **0.24 lower** (0.53 lower to 0.05 higher)	⨁◯◯◯ Very low	CRITICAL
SF-36 physical function												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
4	randomised trials	serious[g]	not serious	not serious	very serious[h]	none	219	207	-	MD **5.95 higher** (1.17 lower to 13.07 higher)	⨁◯◯◯ Very low	CRITICAL

SF-36 physical function - low accountabiilty/contact

2	randomised trials	serious[i]	not serious	not serious	very serious[j]	none	132	129	-	MD **1.84 higher** (2.49 lower to 6.18 higher)	⨁◯◯◯ Very low	CRITICAL

SF-36 physical function - high accountability/contact

2	randomised trials	serious[i]	not serious	not serious	very serious[k]	none	87	78	-	MD **16.46 higher** (16.43 lower to 49.34 higher)	⨁◯◯◯ Very low	CRITICAL

Stiffness (VASo-100) (12 months)

1	randomised trials	serious[a]	not serious	not serious	serious[l]	none	40	38	-	MD **18.4 lower** (31.05 lower to 5.75 lower)	⨁⨁◯◯ Low	CRITICAL

Pittsburgh Sleep Quality Index (12 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	40	38	-	MD **0.8 higher** (0.82 lower to 2.42 higher)	⨁⨁◯◯ Low	IMPORTANT

Fatigue Severity Scale (FSS) (12 months)

1	randomised trials	serious[a]	not serious	serious[n]	serious[l]	none	40	38	-	MD **9.2 lower** (17.1 lower to 1.3 lower)	⨁◯◯◯ Very low	IMPORTANT Overall effect favoring aerobic arm P=0.02

SF-36 mental health

4	randomised trials	serious[g]	serious[o]	not serious	very serious[p]	none	219	207	-	MD **2.72 higher** (1.88 lower to 7.31 higher)	⨁◯◯◯ Very low	IMPORTANT

SF-36 mental health - low accountability/contact

2	randomised trials	serious[g]	serious	not serious	very serious[q]	none	132	129	-	MD **0.29 higher** (4.2 lower to 4.78 higher)	⨁◯◯◯ Very low	IMPORTANT

SF-36 mental health - high accountability/contact

2	randomised trials	serious[g]	serious[o]	not serious	serious[r]	none	87	78	-	MD **6.7 higher** (6.67 lower to 20.06 higher)	⨁◯◯◯ Very low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
Pain (VAS 0-10)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	91	89	-	MD **1.92 lower** (1.2 lower to 2.7 lower)	⨁⨁⨁◯ Moderate	CRITICAL Overall effect favoring aerobic aarm P=0.03
SF-36 global health												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	36	27	-	MD **0.72 higher** (0.23 higher to 1.21 higher)	⨁⨁◯◯ Low	CRITICAL
DAS28												
2	randomised trials	serious[a]	not serious	not serious	serious[m]	none	94	90	-	MD **0.3 higher** (0.22 lower to 0.82 higher)	⨁⨁◯◯ Low	IMPORTANT
Dutch-AIMS2 physical health												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	17	15	-	MD **0.54 lower** (1.08 lower to 0)	⨁⨁◯◯ Low	CRITICAL (P=0.05)
Dutch-AIMS2 pyschological health												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	17	15	-	MD **0.42 lower** (1.29 lower to 0.45 higher)	⨁⨁◯◯ Low	IMPORTANT
Dutch-AIMS2 social interaction												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	17	15	-	MD **0.4 higher** (0.97 lower to 1.77 higher)	⨁⨁◯◯ Low	IMPORTANT
ASES pain and other symptoms												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	17	15	-	MD **0.14 higher** (0.41 lower to 0.69 higher)	⨁⨁◯◯ Low	CRITICAL
ASES Function												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	17	15	-	MD **0.19 higher** (0.14 lower to 0.52 higher)	⨁⨁◯◯ Low	CRITICAL
Timed Up And Go (16 weeks)												
1	randomised trials	serious[a]	not serious	serious[s]	not serious	none	51	51	-	SMD **0.68 lower** (1.08 lower to 0.28 lower)	⨁⨁◯◯ Low	IMPORTANT P=0.0005)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
Functional status: MACTAR												
2	randomised trials	serious[a]	not serious	not serious	serious[m]	none	213	220	-	MD **2.43 higher** (0.68 higher to 4.19 higher)	⨁⨁◯◯ Low	CRITICAL Overall effect favoring aerobic arm (P=0.007)
Radiographic damage: Larsen score for large joints, 24 months												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	136	145	-	MD **0** (0.23 lower to 0.23 higher)	⨁⨁◯◯ Low	IMPORTANT
Radiographic progression: Number with relevant progression, 24 months												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	20/136 (14.7%)	15/145 (10.3%)	**OR 1.49** (0.73 to 3.05)	**43 more per 1,000** (from 26 fewer to 157 more)	⨁⨁◯◯ Low	IMPORTANT
Mental health: HADS, 24 months												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	136	145	-	MD **1.3 lower** (2.25 lower to 0.35 lower)	⨁⨁⨁◯ Moderate	IMPORTANT P=0.007
Disease activity: DAS4 (Ritchie index + number swollen joints), 24 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	136	145	-	MD **0.2 lower** (0.47 lower to 0.07 higher)	⨁⨁◯◯ Low	IMPORTANT

Radiographic damage: Feet only, 24 months

1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	136	145	-	MD **0.8 lower** (1.6 lower to 0)	⨁⨁◯◯ Low	IMPORTANT

Radiographic damage: Hands only, 24 months

1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	136	145	-	MD **1.3 lower** (3.1 lower to 0.5 higher)	⨁⨁◯◯ Low	IMPORTANT

Left grip strength

1	randomised trials	serious[a]	not serious	serious[s]	serious[m]	none	147	146	-	MD **3.56 higher** (10.02 lower to 17.14 higher)	⨁◯◯◯ Very low	IMPORTANT

Right grip strength

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[s]	serious[m]	none	147	146	-	MD **0.33 higher** (13.53 lower to 14.18 higher)	⨁◯◯◯ Very low	IMPORTANT

Walk time

1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	147	146	-	MD **0.61 lower** (1.46 lower to 0.24 higher)	⨁⨁◯◯ Low	IMPORTANT

Total joint count

2	randomised trials	serious[a]	serious[t]	not serious	serious[m]	none	158	156	-	MD **6.79 lower** (12.05 lower to 1.53 lower)	⨁◯◯◯ Very low	IMPORTANT Overall effect favoring aerobic arm (P=0.001)

McGill pain intensity

1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	147	146	-	MD **0.23 lower** (0.73 lower to 0.26 higher)	⨁⨁◯◯ Low	CRITICAL

Self-efficacy

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	147	146	-	MD **0.42 higher** (0.15 higher to 0.68 higher)	⨁⨁◯◯ Low	IMPORTANT

CES-D depression

1	randomised trials	serious[a]	serious[u]	not serious	serious[m]	none	147	146	-	MD **0.27 higher** (1.77 lower to 2.31 higher)	⨁◯◯◯ Very low	IMPORTANT

Change in ESR, 12 months

1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	10	10	-	MD **5.8 lower** (15.15 lower to 3.55 higher)	⨁⨁◯◯ Low	IMPORTANT

Change in fitness score, 12 months

1	randomised trials	serious[a]	not serious	not serious	not serious	none	11	10	-	MD **26.8 higher** (12.8 higher to 40.8 higher)	⨁⨁⨁◯ Moderate	P=0.0002

RAQol score

2	randomised trials	serious[a]	not serious	not serious	serious[m]	none	132	128	-	MD **0.94 lower** (2.01 lower to 0.13 higher)	⨁⨁◯◯ Low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
Fatigue: PROMIS Fatigue												
1	randomised trials	serious[a]	not serious	serious	serious[m]	none	62	52	-	MD **2.38 lower** (5.26 lower to 0.5 higher)	⨁◯◯◯ Very low	IMPORTANT
Disease Activity: RADAI												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	62	52	-	MD **0.86 lower** (1.36 lower to 0.35 lower)	⨁⨁⨁◯ Moderate	IMPORTANT P=0.0009
Pain: PROMIS Pain Interference												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	62	52	-	MD **1.18 lower** (3.83 lower to 1.47 higher)	⨁⨁◯◯ Low	CRITICAL
Mental health: PHQ-8												
1	randomised trials	serious[a]	not serious	not serious	serious[m]	none	62	52	-	MD **0.52 lower** (2.13 lower to 1.08 higher)	⨁⨁◯◯ Low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. All aerobic exercise studies are unable to blind participants and personnel who deliver treatment. A large potential effect with exercise is the trainer and it is unclear across studies whether the trainer was the same person across groups (per study) in order to mitigate trainer effects. As participants are unblinded and they self-report, technically outcome assessors are unblinded. Other reasons to downgrade include a lack of clarity in many risk of bias categories across studies.

b. Imprecision for HAQ in full sample is downgraded because four studies have confidence intervals that cross the line of no difference however mean estimates are in the same direction.

c. RoB in subgroups is same as above

d. In low accountability subgroup imprecision is downgraded because two studies have CIs that cross line of no difference, but mean estimates are in same direction

e. RoB in subgroups is as above

f. In high accountability subgroup imprecision is downgraded because 2/3 studies have CIs that cross the line of no difference but mean estimate are in the same direction.

g. RoB for SF-36 scales is serious because of unblinding of treatment groups to exercise, as this is self-report. All assessors (participants) were unblinded to having been allocated to exercise versus control.

h. Imprecision in SF-36 physical function outcome is consistently manifested by all CIs but one crossing the line of no difference. Mean estimates/differences are in the same direction

i. RoB for SF-36 subgroups is as above

j. Imprecision in low accountability/contact subgroup for SF-36 physical function outcome is consistently manifested by all CIs crossing the line of no difference.

k. Imprecision in the high accountability/contact subgroup for SF-36 physical function. Only two studies reported with different magnitudes. One estimate crosses line of no difference.

l. single study with small sample, not powered on this outcome

m. imprecision because CI crossed line of no difference

n. fatigue is surrogate of functional status

o. inconsistency in direction, uncertainty around inconsistency in magnitude.

p. imprecision in SF-36 mental health summary score shows all studies but one crossing line of no difference. Study estimates are not in the same direction.

q. imprecision in SF-36 mental health summary score for low accountability/contact shows both studies crossing line of no difference. Study estimates are not in the same direction.

r. imprecision in SF-36 mental health summary score for high accountability/contact shows 1/2 studies crossing line of no difference. Study estimates are in the same direction.

s. surrogate measure of functional status as per GRADE instruction

t. inconsistency in magnitude

u. inconsistency in direction

Data table from RCTs aerobic exercise versus active comparators >12 week outcomes

Author(s):

Question: RCT- aerobic v active comparator >12wks compared to placebo for health problem or population

Setting:

Bibliography: . PICO4 6019 7169 1064 1778. Cochrane Database of Systematic Reviews [Year], Issue [Issue].

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic v active comparator >12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
SF-36 physical function												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	32	35	-	MD **0.74 higher** (0.32 higher to 1.16 higher)	⨁⨁◯◯ Low	CRITICAL Dynamic ex + diet versus diet alone P=0.0006)
50ft walk test												
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	72	69	-	MD **0.57 lower** (1.2 lower to 0.06 higher)	⨁◯◯◯ Very low	IMPORTANT
Swollen joint count												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	72	69	-	MD **1.49 lower** (2.37 lower to 0.6 lower)	⨁⨁◯◯ Low	IMPORTANT High intensity versus Low intensity ex Test for overall effect: (P = 0.0009)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic v active comparator >12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
Richie index												
1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	72	69	-	MD **0.24 lower** (1.91 lower to 1.44 higher)	⨁⨁◯◯ Low	IMPORTANT
SF-36 mental health												
1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	32	35	-	MD **0.41 higher** (0.3 lower to 1.12 higher)	⨁⨁◯◯ Low	IMPORTANT
Global assessment of disease activity												
1	randomised trials	serious[d]	not serious	not serious	serious[a]	none	72	69	-	MD **0.73 higher** (0.32 lower to 1.78 higher)	⨁⨁◯◯ Low	IMPORTANT
ESR												
1	randomised trials	serious[d]	not serious	not serious	very serious[e]	none	72	69	-	MD **2.63 higher** (4.62 lower to 9.88 higher)	⨁◯◯◯ Very low	IMPORTANT
SF-36 global health												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic v active comparator >12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[d]	not serious	not serious	serious[e]	none	32	35	-	MD **0.21 higher** (0.25 lower to 0.67 higher)	⨁⨁◯◯ Low	CRITICAL
HAQ												
3	randomised trials	serious[d]	serious[f]	not serious	serious[e]	none	132	129	-	MD **0.03 lower** (0.11 lower to 0.05 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (VAS 0-100)												
2	randomised trials	serious[d]	serious[f]	not serious	serious[e]	none	96	92	-	MD **3.72 higher** (6.7 lower to 14.14 higher)	⨁◯◯◯ Very low	CRITICAL
Disease Activity - DAS28												
2	randomised trials	serious[d]	not serious	not serious	serious[b]	none	48	47	-	MD **0.45 lower** (0.87 lower to 0.04 lower)	⨁⨁◯◯ Low	IMPORTANT Overall effect favoring aerobic arm P=0.03
Disease Activity - CDAI (20 weeks)												
1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	24	24	-	MD **1.6 lower** (7.83 lower to 4.63 higher)	⨁⨁◯◯ Low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic v active comparator >12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
Functional Status: Performance Measure - VO2/kg/min, ml (Baseline - 20 weeks)												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	36	37	-	SMD **1.28 higher** (0.78 higher to 1.79 higher)	⨁⨁⨁◯ Moderate	NOT IMPORTANT (P<00001)
Functional Status: Performance Measure - TUG												
2	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	60	60	-	SMD **0.25 lower** (0.73 lower to 0.24 higher)	⨁◯◯◯ Very low	CRITICAL
Functional Status:Performance Measure - Endurance minutes												
2	randomised trials	serious[a]	not serious	serious[c]	very serious[a]	none	60	60	-	SMD **0.19 higher** (0.52 lower to 0.89 higher)	⨁◯◯◯ Very low	NOT IMPORTANT
Functional Status: Performance Measure - Sit-to-stand												
2	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	60	60	-	SMD **0.16 higher** (0.2 lower to 0.52 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of LifeRA: Physical Function (Baseline - 3 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic v active comparator >12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	85	41	-	SMD **0.02 higher** (0.35 lower to 0.4 higher)	⨁⨁◯◯ Low	CRITICAL

Quality of LifeRA: Emotional/psychological Function (Baseline - 3 months)

1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	85	41	-	SMD **0.02 higher** (0.35 lower to 0.39 higher)	⨁⨁◯◯ Low	IMPORTANT

Quality of LifeRA: Social Function (Baseline - 3 months)

1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	85	41	-	SMD **0.03 higher** (0.34 lower to 0.4 higher)	⨁⨁◯◯ Low	IMPORTANT

Quality of LifeRA: Self-recognized health status (Baseline - 3 months)

1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	85	41	-	SMD **0.07 higher** (0.3 lower to 0.44 higher)	⨁⨁◯◯ Low	IMPORTANT

Quality of lifeRA: Overall (Baseline - 3 months)

1	randomised trials	serious[d]	not serious	not serious	serious[b]	none	85	41	-	SMD **0.16 higher** (0.21 lower to 0.54 higher)	⨁⨁◯◯ Low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic v active comparator >12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
Fatigue (4 years)												
1	randomised trials	serious[d]	not serious	serious[g]	serious[b]	none	24	23	-	SMD **0.35 lower** (0.92 lower to 0.23 higher)	⨁◯◯◯ Very low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Unable to blind participants in trials of aerobic exercise

b. crosses line of no effect

c. artificial measure of functional status per instruction

d. unable to blind participants and assessor unblinded

e. imprecision of direction, cross line of no difference

f. studies produce estimates in different directions

g. surrogate measure of functional status per instruction

Data table from non-RCTs aerobic exercise versus inactive comparators >12 week outcomes

Author(s):

Question: NonRCT aerobic ex compared to inactive comparator >12 weeks for health problem or population

Setting:

Bibliography: . PICO4 6019 7169 1064 1778. Cochrane Database of Systematic Reviews [Year], Issue [Issue].

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
AIMS-Pain(0-10)												
1	observational studies	very serious[a,b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.12 lower** (0.88 lower to 0.65 higher)	⨁◯◯◯ Very low	CRITICAL
AIMS-Physical activity (0-10)												
1	observational studies	very serious[a,b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.11 higher** (0.66 lower to 0.87 higher)	⨁◯◯◯ Very low	IMPORTANT
AIMS-Mobility (0-10)												
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.07 higher** (0.7 lower to 0.83 higher)	⨁◯◯◯ Very low	IMPORTANT
AIMS-ADL (0-10)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.18 lower** (0.94 lower to 0.59 higher)	⨁◯◯◯ Very low	IMPORTANT

AIMS-Household Act(0-10)

1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.47 higher** (0.31 lower to 1.24 higher)	⨁◯◯◯ Very low	IMPORTANT

swollen joints(count)

1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.12 lower** (0.89 lower to 0.64 higher)	⨁◯◯◯ Very low	IMPORTANT

Swollen joints(number)

1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.03 higher** (0.73 lower to 0.8 higher)	⨁◯◯◯ Very low	IMPORTANT

Painful joints-number

1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.3 higher** (0.47 lower to 1.07 higher)	⨁◯◯◯ Very low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
Painful joints-count												
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.23 higher** (0.54 lower to 0.99 higher)	⨁◯◯◯ Very low	IMPORTANT
AIMS-social activities												
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.16 higher** (0.61 lower to 0.93 higher)	⨁◯◯◯ Very low	IMPORTANT
AIMS-depression												
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.27 lower** (1.04 lower to 0.5 higher)	⨁◯◯◯ Very low	IMPORTANT
AIMS-anxiety												
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.52 lower** (1.3 lower to 0.26 higher)	⨁◯◯◯ Very low	IMPORTANT
POMS												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	19	10	-	SMD **0.36 lower** (1.14 lower to 0.41 higher)	⨁◯◯◯ Very low	IMPORTANT

50ft walk time

1	observational studies	very serious[b]	not serious	serious[d]	serious[c]	none	19	10	-	SMD **0.06 lower** (0.82 lower to 0.71 higher)	⨁◯◯◯ Very low	IMPORTANT

intra-articular injections

1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	2/19 (10.5%)	0/10 (0.0%)	**OR 3.00** (0.13 to 68.71)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	IMPORTANT

Cortisone

1	observational studies	very serious[b]	not serious	not serious	serious[c]	none	1/19 (5.3%)	0/10 (0.0%)	**OR 1.70** (0.06 to 45.66)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	IMPORTANT

HAQ-6m

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	18	18	-	MD **0.98 lower** (1.67 lower to 0.28 lower)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
DAS28-6m												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	18	18	-	SMD **0.61 lower** (1.28 lower to 0.06 higher)	⨁◯◯◯ Very low	IMPORTANT
Lansbury Index												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23	23	-	MD **26 lower** (36.31 lower to 15.69 lower)	⨁◯◯◯ Very low	
% on Sick-Leave or Sick-Pension												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	1/100 (1.0%)	29/100 (29.0%)	**OR 0.02** (0.00 to 0.19)	**282 fewer per 1,000** (from 218 fewer to --)	⨁◯◯◯ Very low	IMPORTANT
Orthopedic Surgeries												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	10/23 (43.5%)	12/23 (52.2%)	**OR 0.71** (0.22 to 2.25)	**85 fewer per 1,000** (from 328 fewer to 189 more)	⨁◯◯◯ Very low	IMPORTANT
Change in Xray index												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23	23	-	MD **2.8 lower** (5.32 lower to 0.28 lower)	⨁◯◯◯ Very low	IMPORTANT

Stair Test (seconds)

1	observational studies	serious[a]	not serious	serious[d]	serious[c]	none	21	15	-	MD **6.7 lower** (15.69 lower to 2.29 higher)	⨁◯◯◯ Very low	IMPORTANT

Step Test (cm)

1	observational studies	serious[a]	not serious	serious[d]	serious[c]	none	21	17	-	MD **5.4 higher** (1.02 lower to 11.82 higher)	⨁◯◯◯ Very low	IMPORTANT

Walk Test (minutes)

1	observational studies	serious[a]	not serious	serious[d]	serious[c]	none	19	9	-	MD **0.95 higher** (0.98 lower to 2.88 higher)	⨁◯◯◯ Very low	IMPORTANT

Wash Hair

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	22/23 (95.7%)	19/23 (82.6%)	**OR 4.63** (0.48 to 45.09)	**130 more per 1,000** (from 131 fewer to 169 more)	⨁◯◯◯ Very low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
Wash Face												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23/23 (100.0%)	22/23 (95.7%)	**OR 3.13** (0.12 to 81.00)	**29 more per 1,000** (from 231 fewer to 43 more)	⨁◯◯◯ Very low	IMPORTANT
Intimate hygiene												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23/23 (100.0%)	21/23 (91.3%)	**OR 5.47** (0.25 to 120.37)	**70 more per 1,000** (from 189 fewer to 86 more)	⨁◯◯◯ Very low	IMPORTANT
Wash feet												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	21/23 (91.3%)	17/23 (73.9%)	**OR 3.71** (0.66 to 20.76)	**174 more per 1,000** (from 88 fewer to 244 more)	⨁◯◯◯ Very low	IMPORTANT
Toilet												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23/23 (100.0%)	22/23 (95.7%)	**OR 3.13** (0.12 to 81.00)	**29 more per 1,000** (from 231 fewer to 43 more)	⨁◯◯◯ Very low	IMPORTANT
Socks, on-off												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23/23 (100.0%)	20/23 (87.0%)	**OR 8.02** (0.39 to 164.73)	**112 more per 1,000** (from 147 fewer to 130 more)	⨁◯◯◯ Very low	IMPORTANT

Shirt, on-off

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	21/23 (91.3%)	16/19 (84.2%)	**OR 1.97** (0.29 to 13.21)	**71 more per 1,000** (from 235 fewer to 144 more)	⨁◯◯◯ Very low	IMPORTANT

Fasten buttons

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	21/23 (91.3%)	18/23 (78.3%)	**OR 2.92** (0.50 to 16.89)	**131 more per 1,000** (from 140 fewer to 201 more)	⨁◯◯◯ Very low	IMPORTANT

Rise from lying to standing

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	22/23 (95.7%)	19/22 (86.4%)	**OR 3.47** (0.33 to 36.24)	**93 more per 1,000** (from 187 fewer to 132 more)	⨁◯◯◯ Very low	IMPORTANT

Walk on level ground

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	22/23 (95.7%)	21/23 (91.3%)	**OR 2.10** (0.18 to 24.87)	**44 more per 1,000** (from 259 fewer to 83 more)	⨁◯◯◯ Very low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
Walk upstairs and downstairs												
1	observational studies	serious[a]	not serious	not serious	not serious	none	22/23 (95.7%)	13/21 (61.9%)	**OR 13.54** (1.52 to 120.85)	**337 more per 1,000** (from 93 more to 376 more)	⨁◯◯◯ Very low	IMPORTANT
Eat with knife and fork												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	22/23 (95.7%)	22/23 (95.7%)	**OR 1.00** (0.06 to 17.02)	**0 fewer per 1,000** (from 388 fewer to 41 more)	⨁◯◯◯ Very low	IMPORTANT
Cook												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	21/22 (95.5%)	20/23 (87.0%)	**OR 3.15** (0.30 to 32.85)	**85 more per 1,000** (from 203 fewer to 126 more)	⨁◯◯◯ Very low	IMPORTANT
Wash dishes												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	22/23 (95.7%)	20/23 (87.0%)	**OR 3.30** (0.32 to 34.35)	**87 more per 1,000** (from 189 fewer to 126 more)	⨁◯◯◯ Very low	IMPORTANT
Go shopping												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	17/20 (85.0%)	13/22 (59.1%)	**OR 3.92** (0.88 to 17.46)	**259 more per 1,000** (from 31 fewer to 371 more)	⨁◯◯◯ Very low	IMPORTANT

Clean up the house

1	observational studies	serious[a]	not serious	not serious	not serious	none	18/21 (85.7%)	9/22 (40.9%)	**OR 8.67** (1.96 to 38.40)	**448 more per 1,000** (from 167 more to 555 more)	⨁◯◯◯ Very low	IMPORTANT

Wash the laundry

1	observational studies	serious[a]	not serious	not serious	not serious	none	19/21 (90.5%)	12/22 (54.5%)	**OR 7.92** (1.47 to 42.54)	**359 more per 1,000** (from 93 more to 435 more)	⨁◯◯◯ Very low	IMPORTANT

Make the bed

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	18/21 (85.7%)	18/23 (78.3%)	**OR 1.67** (0.35 to 8.04)	**75 more per 1,000** (from 225 fewer to 184 more)	⨁◯◯◯ Very low	IMPORTANT

Use scissors

1	observational studies	serious[a]	not serious	not serious	serious[c]	none	21/23 (91.3%)	20/23 (87.0%)	**OR 1.57** (0.24 to 10.44)	**43 more per 1,000** (from 254 fewer to 116 more)	⨁◯◯◯ Very low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
Use public transport												
1	observational studies	serious[a]	not serious	not serious	not serious	none	17/22 (77.3%)	11/23 (47.8%)	**OR 3.71** (1.02 to 13.47)	**295 more per 1,000** (from 5 more to 447 more)	⨁◯◯◯ Very low	IMPORTANT
Pick up object from the floor												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	22/23 (95.7%)	21/23 (91.3%)	**OR 2.10** (0.18 to 24.87)	**44 more per 1,000** (from 259 fewer to 83 more)	⨁◯◯◯ Very low	IMPORTANT
Take object from shelf above shoulder level												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	23/23 (100.0%)	16/22 (72.7%)	**OR 18.52** (0.97 to 351.82)	**253 more per 1,000** (from 6 fewer to 272 more)	⨁◯◯◯ Very low	IMPORTANT
Write a letter												
1	observational studies	serious[a]	not serious	not serious	serious[c]	none	21/22 (95.5%)	22/23 (95.7%)	**OR 0.95** (0.06 to 16.27)	**2 fewer per 1,000** (from 388 fewer to 41 more)	⨁◯◯◯ Very low	IMPORTANT
Trunk Flexibility at 12-month												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious[a,b,e]	not serious	serious[d]	serious[c]	none	17	15	-	MD **5.4 lower** (12.17 lower to 1.37 higher)	⨁◯◯◯ Very low	IMPORTANT

Grip Strength (mm Hg) at 12 month

1	observational studies	very serious[a,b,e]	not serious	serious[d]	serious[c]	none	15	17	-	MD **15 higher** (23.27 lower to 53.27 higher)	⨁◯◯◯ Very low	IMPORTANT

Work Capacity Evaluation - Hands at 12 months

1	observational studies	very serious[a,b,e]	not serious	serious[d]	serious[c]	none	15	17	-	MD **0.1 higher** (0.73 lower to 0.93 higher)	⨁◯◯◯ Very low	IMPORTANT

Work Capacity Evaluation - Lift at 12 months

1	observational studies	very serious[a,b,e]	not serious	serious[d]	serious[c]	none	15	17	-	MD **0.3 higher** (0.05 lower to 0.65 higher)	⨁◯◯◯ Very low	IMPORTANT

Work Capacity Evaluation - Legs at 12 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious[a,b,e]	not serious	serious[d]	not serious	none	15	17	-	MD **0.4 higher** (0.01 higher to 0.79 higher)	⨁◯◯◯ Very low	IMPORTANT

Work Capacity Evaluation - Dictionary of Occupational Titles at 12 months

1	observational studies	very serious[a,b,e]	not serious	serious[d]	serious[c]	none	15	17	-	MD **0.5 higher** (0.13 lower to 1.13 higher)	⨁◯◯◯ Very low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. case-control

b. no matching

c. crosses line of no effect

d. artificial measure of functional status per instruction

e. assignment determined by proximity to facility

Data table from RCTs aerobic ex compared to inactive comparator < 12wks outcomes

Author(s):

Question: RCT aerobic ex compared to inactive comparator < 12wks for health problem or population

Setting:

Bibliography: . PICO4 6019 7169 1064 1778. Cochrane Database of Systematic Reviews [Year], Issue [Issue].

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator < 12wks	Relative (95% CI)	Absolute (95% CI)		
Adverse Events												
1	randomised trials	serious[a]	not serious	not serious	not serious	publication bias strongly suspected[b]	0/11 (0.0%)	0/22 (0.0%)	not estimable		⨁⨁◯◯ Low	CRITICAL
Functional status: HAQ, 6 weeks												
2	randomised trials	serious[a]	not serious	not serious	serious[d]	none	28	43	-	MD **0.08 higher** (0.06 lower to 0.22 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of Life: EuroQoL, 6 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	11	22	-	MD **5.1 lower** (8.78 lower to 1.42 lower)	⨁⨁◯◯ Low	IMPORTANT (P=0.007)
Self-efficacy for symptoms: ASES, 6 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	11	22	-	MD **26.1 higher** (8.49 higher to 43.71 higher)	⨁⨁◯◯ Low	IMPORTANT (P=0.004)
Depression (CES-D 6 wks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic ex	inactive comparator < 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	not serious	none	17	21	-	MD **6.74 lower** (12.02 lower to 1.46 lower)	⨁⨁⨁◯ Moderate	IMPORTANT P<0.05)

Fatigue (BRAF total)

1	randomised trials	serious[a]	not serious[e]	serious[e]	not serious	none	17	21	-	MD **11.08 lower** (16.15 lower to 6.01 lower)	⨁⨁◯◯ Low	CRITICAL P<0.05)

Fatigue (VAS 0-100)

1	randomised trials	serious[a]	not serious[e]	serious[e]	not serious	none	17	21	-	MD **16.05 lower** (24.36 lower to 7.74 lower)	⨁⨁◯◯ Low	CRITICAL P<0.05)

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio

Explanations

a. unblinded to intervention, self report nature of event reporting

b. adverse events not recorded by many clinical trials

c. feasibility study, not powered

d. crosses line of no effect

e. fatigue is surrogate measure for function

Data table from RCTs aerobic exercise versus active comparators <12 week outcomes

Author(s):

Question: RCT aerobic v active comparator <12wks compared to placebo for health problem or population

Setting:

Bibliography: . PICO4 6019 7169 1064 1778. Cochrane Database of Systematic Reviews [Year], Issue [Issue].

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
Number of swollen joints (8 weeks)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	15	15	-	MD **2.27 higher** (3.25 lower to 7.79 higher)	⨁⨁◯◯ Low	IMPORTANT
Number of tender joints (8 weeks)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	15	15	-	MD **3.2 higher** (6.17 lower to 12.57 higher)	⨁⨁◯◯ Low	IMPORTANT
Pain severity (8 weeks)												
1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	15	15	-	MD **0.91 higher** (0.72 lower to 2.54 higher)	⨁⨁◯◯ Low	CRITICAL
Pain in ADLs (8 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	15	15	-	MD **0.12 lower** (0.6 lower to 0.36 higher)	⨁⨁◯◯ Low	CRITICAL

50ft walking time (8 weeks)

1	randomised trials	serious[a]	not serious	serious[d]	serious[b]	none	15	15	-	MD **0.68 higher** (0.77 lower to 2.13 higher)	⨁◯◯◯ Very low	IMPORTANT

Global self-assessment (8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	15	15	-	MD **1.2 higher** (0.19 lower to 2.59 higher)	⨁⨁◯◯ Low	IMPORTANT

Global physician assessment (8 weeks)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	15	15	-	MD **0.76 higher** (0.19 higher to 1.33 higher)	⨁⨁◯◯ Low	IMPORTANT

HAQ

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials	serious[a,c]	serious[e]	not serious	serious[b]	none	46	44	-	MD **0.05 higher** (0.08 lower to 0.18 higher)	⨁◯◯◯ Very low	CRITICAL

AIMS-depression (8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	15	15	-	MD **0.13 higher** (0.66 lower to 0.92 higher)	⨁⨁◯◯ Low	IMPORTANT

AIMS-anxiety (8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	15	15	-	MD **0.14 higher** (0.62 lower to 0.9 higher)	⨁⨁◯◯ Low	IMPORTANT

ESR (8 weeks)

2	randomised trials	serious[a]	not serious	not serious	serious[b]	none	35	35	-	MD **1.33 higher** (6.43 lower to 9.09 higher)	⨁⨁◯◯ Low	IMPORTANT

CRP (8 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	15	15	-	MD **0.2 higher** (1.45 lower to 1.85 higher)	⨁⨁◯◯ Low	IMPORTANT

Rheumatoid factor (8 weeks)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	15	15	-	MD **83 higher** (402.97 lower to 568.97 higher)	⨁⨁◯◯ Low	IMPORTANT

Stifness (min), 15th day

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **10.65 lower** (29.95 lower to 8.65 higher)	⨁⨁◯◯ Low	IMPORTANT

Pain (VAS), 15th day

1	randomised trials	serious[a,c]	not serious	serious	serious[b]	none	20	20	-	MD **0.38 lower** (1.52 lower to 0.76 higher)	⨁◯◯◯ Very low	CRITICAL

Ritchie articular index (RAI)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	20	20	-	MD **0.25 lower** (5.02 lower to 4.52 higher)	⊕⊕◯◯ Low	IMPORTANT

Fatigue Severity Scale (change 0-8 weeks)

1	randomised trials	serious[a,c]	not serious	serious[f]	serious[b]	none	20	20	-	MD **1.53 lower** (2.81 lower to 0.26 lower)	⊕◯◯◯ Very low	IMPORTANT

Beck Depression Inventory (change: 0-8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **0.1 lower** (2.06 lower to 1.86 higher)	⊕⊕◯◯ Low	IMPORTANT

6 minute walk test (change: 0-8 weeks)

1	randomised trials	serious[a,c]	not serious	serious[d]	very serious[b,g]	none	20	20	-	MD **2.15 lower** (11.95 lower to 7.65 higher)	⊕◯◯◯ Very low	IMPORTANT

McGill Pain Questionnaire Short Form words subscale (change: 0-8 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **0.06 higher** (1.41 lower to 1.53 higher)	⨁⨁◯◯ Low	CRITICAL

McGill Pain Questionnaire Short Form visual analog scale subscale (change: 0-8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **0.16 higher** (0.33 lower to 0.64 higher)	⨁⨁◯◯ Low	CRITICAL

McGill Pain Questionnaire Short Form Likert subscale (change: 0-8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **0.25 higher** (0.15 higher to 0.35 higher)	⨁⨁◯◯ Low	CRITICAL

Pittsburg Sleep Quality Index (change: 0-8 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **0.15 higher** (0.54 lower to 0.84 higher)	⨁⨁◯◯ Low	IMPORTANT

Rheumatoid Arthritis Quality of Life (change: 0-8 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic v active comparator <12wks	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	20	20	-	MD **0.53 lower** (1.81 lower to 0.74 higher)	⨁⨁◯◯ Low	IMPORTANT

Active Joint Count (11 weeks)

1	randomised trials	serious[a,c]	not serious	not serious	serious[b]	none	11	9	-	SMD **0.07 higher** (0.81 lower to 0.95 higher)	⨁⨁◯◯ Low	IMPORTANT

Performance test: Grip strength (11 weeks)

1	randomised trials	serious[a,c]	not serious	serious[d]	serious[b]	none	11	9	-	SMD **0.58 lower** (1.48 lower to 0.33 higher)	⨁◯◯◯ Very low	IMPORTANT

Performance test: Walking on treadmill (11 weeks)

1	randomised trials	serious[a,c]	not serious	serious[d]	serious[b]	none	11	9	-	SMD **0.31 lower** (1.2 lower to 0.57 higher)	⨁◯◯◯ Very low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. participant unblinded

b. crosses line of no effect

c. assessor unblinded

d. artificial measure of functional status per instruction

e. magnitude

f. surrogate measure of functional status, downgrade per instruction

g. direction of effect

1. Andersson SEM, Lange E, Kucharski D, Svedlund S, Önnheim K, Bergquist M, et al. Moderate- to high intensity aerobic and resistance exercise reduces peripheral blood regulatory cell populations in older adults with rheumatoid arthritis. Immun Ageing. 2020;17:12.

2. Baxter SV, Hale LA, Stebbings S, Gray AR, Smith CM, Treharne GJ. Walking is a Feasible Physical Activity for People with Rheumatoid Arthritis: A Feasibility Randomized Controlled Trial. Musculoskeletal Care. 2016;14(1):47-56.

3. Breedland I, van Scheppingen C, Leijsma M, Verheij-Jansen NP, van Weert E. Effects of a group-based exercise and educational program on physical performance and disease self-management in rheumatoid arthritis: a randomized controlled study. Phys Ther. 2011;91(6):879-93.

4. da Silva KN, Teixeira LE, Imoto AM, Atallah AN, Peccin MS, Trevisani VF. Effectiveness of sensorimotor training in patients with rheumatoid arthritis: a randomized controlled trial. Rheumatol Int. 2013;33(9):2269-75.

5. Daltroy LH, Robb-Nicholson C, Iversen MD, Wright EA, Liang MH. Effectiveness of minimally supervised home aerobic training in patients with systemic rheumatic disease. Br J Rheumatol. 1995;34(11):1064-9.

6. de Jong Z, Munneke M, Zwinderman AH, Kroon HM, Jansen A, Ronday KH, et al. Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis? Results of a randomized controlled trial. Arthritis Rheum. 2003;48(9):2415-24.

7. de Jong Z, Munneke M, Zwinderman AH, Kroon HM, Ronday KH, Lems WF, et al. Long term high intensity exercise and damage of small joints in rheumatoid arthritis. Ann Rheum Dis. 2004;63(11):1399-405.

8. Durcan L, Wilson F, Cunnane G. The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. J Rheumatol. 2014;41(10):1966-73.

9. García-Morales JM, Lozada-Mellado M, Hinojosa-Azaola A, Llorente L, Ogata-Medel M, Pineda-Juárez JA, et al. Effect of a Dynamic Exercise Program in Combination With Mediterranean Diet on Quality of Life in Women With Rheumatoid Arthritis. J Clin Rheumatol. 2020;26(7S Suppl 2):S116-S22.

10. Hansen TM, Hansen G, Langgaard AM, Rasmussen JO. Longterm physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. Scand J Rheumatol. 1993;22(3):107-12.

11. Hsieh LF, Chen SC, Chuang CC, Chai HM, Chen WS, He YC. Supervised aerobic exercise is more effective than home aerobic exercise in female chinese patients with rheumatoid arthritis. J Rehabil Med. 2009;41(5):332-7.

12. Hurkmans EJ, van den Berg MH, Ronday KH, Peeters AJ, le Cessie S, Vlieland TP. Maintenance of physical activity after Internet-based physical activity interventions in patients with rheumatoid arthritis. Rheumatology (Oxford). 2010;49(1):167-72.

13. Jahanbin I, Hoseini Moghadam M, Nazarinia MA, Ghodsbin F, Bagheri Z, Ashraf AR. The effect of conditioning exercise on the health status and pain in patients with rheumatoid arthritis: a randomized controlled clinical trial. Int J Community Based Nurs Midwifery. 2014;2(3):169-76.

14. Katz P, Margaretten M, Gregorich S, Trupin L. Physical Activity to Reduce Fatigue in Rheumatoid Arthritis: A Randomized Controlled Trial. Arthritis Care Res (Hoboken). 2018;70(1):1-10.

15. Kucharski D, Lange E, Ross AB, Svedlund S, Feldthusen C, Önnheim K, et al. Moderate-to-high intensity exercise with person-centered guidance influences fatigue in older adults with rheumatoid arthritis. Rheumatol Int. 2019;39(9):1585-94.

16. Lange E, Kucharski D, Svedlund S, Svensson K, Bertholds G, Gjertsson I, et al. Effects of Aerobic and Resistance Exercise in Older Adults With Rheumatoid Arthritis: A Randomized Controlled Trial. Arthritis Care Res (Hoboken). 2019;71(1):61-70.

17. Lange E, Gjertsson I, Mannerkorpi K. Long-time follow up of physical activity level among older adults with rheumatoid arthritis. Eur Rev Aging Phys Act. 2020;17:10.

18. Melikoglu MA, Karatay S, Senel K, Akcay F. Association between dynamic exercise therapy and IGF-1 and IGFBP-3 concentrations in the patients with rheumatoid arthritis. Rheumatol Int. 2006;26(4):309-13.

19. Neuberger GB, Aaronson LS, Gajewski B, Embretson SE, Cagle PE, Loudon JK, et al. Predictors of exercise and effects of exercise on symptoms, function, aerobic fitness, and disease outcomes of rheumatoid arthritis. Arthritis Rheum. 2007;57(6):943-52.

20. Rahnama N, Mazloum V. Effects of strengthening and aerobic exercises on pain severity and function in patients with knee rheumatoid arthritis. Int J Prev Med. 2012;3(7):493-8.

21. Sanford Smith S, MacKay-Lyons M, Nunes-Clement s. Therapeutic Benefit of Aquaerobics for Individuals with Rheumatoid Arthritis. Physiotherapy Canada. 1998;50:7.

22. van den Berg MH, Ronday HK, Peeters AJ, le Cessie S, van der Giesen FJ, Breedveld FC, et al. Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. Arthritis Rheum. 2006;55(6):935-45.

23. van den Ende CH, Hazes JM, le Cessie S, Mulder WJ, Belfor DG, Breedveld FC, et al. Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomised clinical trial. Ann Rheum Dis. 1996;55(11):798-805.

24. Westby MD, Wade JP, Rangno KK, Berkowitz J. A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. J Rheumatol. 2000;27(7):1674-80.

25. Yang D-j, Xu F-y, Gan J-h. Assessment of curative effect of aerobic exercise with quality of life questionnaire for patients with rheumatoid arthritis. Chinese Journal of Clinical Rehabilitation. 2005;9.

26. Yentür SB, Ataş N, Öztürk MA, Oskay D. Comparison of the effectiveness of pilates exercises, aerobic exercises, and pilates with aerobic exercises in patients with rheumatoid arthritis. Ir J Med Sci. 2021;190(3):1027-34.

27. Azeez M, Clancy C, O'Dwyer T, Lahiff C, Wilson F, Cunnane G. Benefits of exercise in patients with rheumatoid arthritis: a randomized controlled trial of a patient-specific exercise programme. Clin Rheumatol. 2020;39(6):1783-92.

28. Brodin N, Eurenius E, Jensen I, Nisell R, Opava CH, Group PS. Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. Arthritis Rheum. 2008;59(3):325-31.

29. Loeppenthin K, Esbensen BA, Klausen JM, Østergaard M, Christensen JF, Tolver A, et al. Efficacy and Acceptability of Intermittent Aerobic Exercise on Polysomnography-Measured Sleep in People With Rheumatoid Arthritis With Self-Reported Sleep Disturbance: A Randomized Controlled Trial. ACR Open Rheumatol. 2022;4(5):395-405.

30. Minor MA, Hewett JE. Physical fitness and work capacity in women with rheumatoid arthritis. Arthritis Care Res. 1995;8(3):146-54.

31. Nordemar R, Ekblom B, Zachrisson L, Lundqvist K. Physical training in rheumatoid arthritis: a controlled long-term study. I. Scand J Rheumatol. 1981;10(1):17-23.

32. Nordemar R. Physical training in rheumatoid arthritis: A controlled long-term study. II. Functional capacity and general attitudes. Scand J Rheumatol. 1981;10(1):25-30.

33. Noreau L, Martineau H, Roy L, Belzile M. Effects of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. Am J Phys Med Rehabil. 1995;74(1):19-27.

34. Stavropoulos-Kalinoglou A, Metsios GS, Veldhuijzen van Zanten JJ, Nightingale P, Kitas GD, Koutedakis Y. Individualised aerobic and resistance exercise training improves cardiorespiratory fitness and reduces cardiovascular risk in patients with rheumatoid arthritis. Ann Rheum Dis. 2013;72(11):1819-25.

35. Abyad A, Boyer JT. Arthritis and aging. Curr Opin Rheumatol. 1992;4(2):153-9.

PICO 5: Should patients with RA engage in an aquatic exercise program?

Summary: Literature searches identified 8 randomized control trial studies [1,2,3,4,5,6,7,8] addressing this question. The eight studies made three types of comparisons:

- Aquatic exercises compared to no exercise (Table 1) [2,8]
- Aquatic exercises compared to land exercise (Table 2) [1,3,4,6,7,9]
- Aquatic exercise compared to warm-water immersion (Table 3) [5]

We deemed the intervention primarily aquatic in all eight trials, and aquatic regimens included both resistance [1, 5, 6, 7, 9] and aerobic [1, 3, 6, 8, 9] components performed anywhere between 1 [8], 2 [3, 5, 6], 3 [7, 8], or 5 [1]x/week with intervention lengths from 6 [4] weeks to 4 [9] years, with most around 12 weeks [1,2,3,5,6,7,8] of intervention. Intensities were generally ratings of perceived exertion (RPE) 13-15/20 on the Berg scale. Control groups maintained their typical levels of physical activity (self-report not objectively measured) and received no intervention, while land exercise control groups [1, 4, 6, 7] received similar interventions to the aquatic groups (i.e., weight training against gravity compared to foam weights and fins against water), except for a few who simply provided a home exercise program [3, 9]. Uniquely in one study, [5] a warm-water immersion group sat in equal temperature water to the aquatic group for the same amount of time the aquatic exercise took place.

Many exercise regimens included not only aquatic exercise, but also other forms of exercise such as aerobic or resistance exercise. Due to this multicomponent nature, we separately provide a summary of all multicomponent exercise studies (PICO 4-5-6), and 1 study discussed here is also discussed in that document.

Aquatic exercises ***compared to no exercise controls*** (two studies) had statistically significantly better functional status, specifically functional task proxies (surrogates) that signify the ability to perform the task. Additionally, pain also statistically significantly decreased compared to no exercise controls.

Aquatic exercises ***compared to land exercise*** (six studies) were not statistically significantly different in most outcomes related to functional status, however it is worth noting the decrease in patient global assessment in the aquatic exercise group compared to the land exercise group. This could potentially be due to the influence gravity has on land exercise in promoting and improving balance during exercise that may not be present to the same extent in aquatic exercise. Additionally, no significant differences in pain were shown but a slight lean toward land exercise having lower pain. Therefore, aquatic exercise and land exercise should be recommended equally.

Aquatic exercise ***compared to warm-water immersion*** (one study) was not statistically significantly different in most outcomes related to functional status and pain. However, wrist range of motion was improved, and grip strength dramatically improved compared to the warm-water immersion. Therefore, regarding pain aquatic exercise and warm-water immersion will yield similar results. When concerned about functional status aquatic exercise is recommended based on the proxy functional status measures of wrist range of motion and grip strength, two specific areas people with rheumatoid arthritis typical have difficulty in.

The RCTs were rated very low to moderate quality due to reliance on single studies, the use of proxy measures for functional status, and risk of bias assessment.

Table 4 provides data from any studies from which effect sizes were not computable (e.g., unreported dispersion).

Quality of evidence across all critical outcomes: Low. This is based only on the comparison of aquatic exercise vs no exercise, in which several direct functional status measures were rated Low, and a surrogate functional status measure was graded as Very Low; consequently, we chose to ignore the latter.

Tables 1-3. Data from randomized controlled trials

Table 1 Aquatic Exercise vs No exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Control (Nothing)	Relative (95% CI)	Absolute (95% CI)		
Pain: Pain > 12 weeks (3 months to 4 years)												
2	randomised trials	serious[c]	not serious	not serious	serious[b]	none	37	36	-	MD 10.25 lower (22.62 lower to 2.12 higher)	⨁⨁◯◯ Low	CRITICAL Not significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Control (Nothing)	Relative (95% CI)	Absolute (95% CI)		
Pain: Pain during testing > 12 weeks (3 months to 4 years)												
1	randomised trials	serious[c]	not serious	not serious	serious[b]	none	17	13	-	MD 1.44 lower (4.04 lower to 1.16 higher)	⨁⨁◯◯ Low	CRITICAL Not significant
Functional status: HAQ > 12 weeks (3 months to 4 years)												
2	randomised trials	serious[a]	not serious	not serious	serious[e]	none	53	57	-	MD 0.49 lower (1.17 lower to 0.19 higher)	⨁⨁◯◯ Low	CRITICAL Not significant
Functional status: Patient global assessment > 12 weeks (3 months to 4 years)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	33	34	-	MD 3.2 lower (4.48 lower to 1.92 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of aquatic exercise**
Function as inferred from Muscle Strength > 12 weeks (3 months to 4 years)												
1	randomised trials	serious[c]	not serious	serious[d]	serious[b]	none	17	13	-	MD 5.3 higher (0.48 higher to 10.12 higher)	⨁◯◯◯ Very low	CRITICAL **Statistically significant in favor of aquatic exercise**
Function as inferred from Fatigue RPE > 12 weeks (3 months to 4 years)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Control (Nothing)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[c]	not serious	serious[d]	serious[b,e]	none	17	13	-	MD 0.2 higher (1.12 lower to 1.52 higher)	⊕◯◯◯ Very low	CRITICAL Not significant

Functional status: SF-36 physical functioning > 12 weeks (3 months to 4 years)

1	randomised trials	not serious	not serious	not serious	serious[b]	none	20	23	-	MD 0.46 higher (0.15 lower to 1.07 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant

Functional status: SF-36 physical component > 12 weeks (3 months to 4 years)

1	randomised trials	not serious	not serious	not serious	serious[b]	none	20	23	-	MD 0.44 higher (0.16 lower to 1.05 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant

Functional status: AIMS-2 Physical > 12 weeks (3 months to 4 years)

1	randomised trials	not serious	not serious	not serious	serious[b]	none	20	23	-	MD 0.46 lower (1.07 lower to 0.15 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant

Functional status: IMF > 12 weeks (3 months to 4 years)

1	randomised trials	not serious	not serious	not serious	serious[b]	none	20	23	-	MD 0.85 lower (1.48 lower to 0.23 lower)	⊕⊕⊕◯ Moderate	CRITICAL **Statistically significant in favor of aquatic exercise**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Control (Nothing)	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from Chair Test > 12 weeks (3 months to 4 years)												
1	randomised trials	not serious	not serious	serious[d]	serious[b]	none	20	23	-	MD 0.95 higher (0.31 higher to 1.58 higher)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of aquatic exercise**
Disease activity: DAS-28 > 12 weeks (3 months to 4 years)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	33	34	-	MD 1.1 lower (1.56 lower to 0.64 lower)	⨁⨁◯◯ Low	IMPORTANT **Statistically significant in favor of aquatic exercise**
Disease activity: Joint mobility (lower score is better) > 12 weeks (3 months to 4 years)												
1	randomised trials	serious[c]	not serious	serious[d]	serious[c,e]	none	17	13	-	MD 1.8 lower (6.41 lower to 2.81 higher)	⨁◯◯◯ Very low	IMPORTANT Not significant
Mental health status: Mental Health > 12 weeks (3 months to 4 years)												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	20	23	-	MD 0.44 higher (0.17 lower to 1.05 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant
Mental health status: SF-36 mental component > 12 weeks (3 months to 4 years)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Control (Nothing)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[b,e]	none	20	23	-	MD 0.27 higher (0.33 lower to 0.88 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant

CI: confidence interval; **MD:** mean difference

Explanations

a. Loss to follow-up >25%

b,e. wide CI crosses 0

c. Assessor was not blinded to treatment group

d. Proxy measure

Table 2 Aquatic exercise vs Land exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Land exercise	Relative (95% CI)	Absolute (95% CI)		
Functional status: HAQ > 12 weeks (3 months to 6 months)												
2	randomised trials	serious[b]	not serious	not serious	serious[b]	none	56	54	-	MD 0.16 lower (0.65 lower to 0.33 higher)	⨁⨁◯◯ Low	CRITICAL Not significant
Functional status: Patient global assessment > 12 weeks (3 months to 6 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Land exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[b]	not serious	not serious	Not serious	none	33	33	-	MD 2 lower (3.38 lower to 0.62 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant in favor of land exercise**

Functional status: AIMS2-SF > 12 weeks (3 months to 6 months)

2	randomised trials	not serious	not serious	not serious	serious[d]	none	58	55	-	MD 0.22 higher (0.41 lower to 0.85 higher)	⨁⨁⨁◯ Moderate	CRITICAL Not significant

Function as inferred from Knee range of motion > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	serious[e]	serious[d]	none	35	34	-	MD 3.4 higher (8.23 lower to 15.03 higher)	⨁⨁◯◯ Low	CRITICAL Not significant

Function as inferred from Wrist range of motion > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	serious[e]	serious[d]	none	35	34	-	MD 15.5 higher (11.37 lower to 42.37 higher)	⨁⨁◯◯ Low	CRITICAL Not significant

Function as inferred from Grip strength > 12 weeks (3 months to 6 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Land exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[e]	serious[d]	none	35	34	-	MD 14.3 higher (24.43 lower to 53.03 higher)	⨁⨁◯◯ Low	CRITICAL Not significant

Functional status: AIMS2: Physical Capacity > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	not serious	serious[d]	none	35	34	-	MD 0.4 lower (1.32 lower to 0.52 higher)	⨁⨁⨁◯ Moderate	CRITICAL Not significant

Function as inferred from Grip strength < 12 weeks (6 weeks to 12 weeks)

1	randomised trials	not serious	not serious	serious[e]	serious[d]	none	11	9	-	MD 0.58 lower (1.48 lower to 0.33 higher)	⨁⨁◯◯ Low	CRITICAL Not significant

Function as inferred from Walking on treadmill < 12 weeks (6 weeks to 12 weeks)

1	randomised trials	serious[b]	not serious	serious[e]	serious[d]	none	11	9	-	MD 0.31 lower (1.2 lower to 0.57 higher)	⨁◯◯◯ Very low	CRITICAL Not significant

Functional status: HAQ < 12 weeks (6 weeks to 12 weeks)

1	randomised trials	serious[b]	not serious	not serious	serious[d]	none	11	9	-	MD 0.76 higher (0.16 lower to 1.68 higher)	⨁⨁◯◯ Low	CRITICAL Not significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Land exercise	Relative (95% CI)	Absolute (95% CI)		
Pain: McGill Pain Questionnaire_Sensory Pain > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	35	34	-	MD 0.04 lower (0.39 lower to 0.31 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant
Pain: McGill Pain Questionnaire_Affective Pain > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	35	34	-	MD 0.46 higher (0.32 lower to 1.24 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant
Pain: AIMS2: Pain > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	35	34	-	MD 1 higher (0.05 lower to 2.05 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant
Disease activity: DAS-28 > 12 weeks (3 months to 6 months)												
2	randomised trials	serious[a,b]	not serious	not serious	serious[c]	none	56	54	-	MD 0.01 lower (1.17 lower to 1.14 higher)	⊕⊕◯◯ Low	IMPORTANT Not significant
Disease activity: Ritchie articular index > 12 weeks (3 months to 6 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Land exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	Serious[d]	serious[c,d]	none	35	34	-	MD 3.5 lower (8.85 lower to 1.85 higher)	⨁⨁◯◯ Low	IMPORTANT Not significant

Disease activity: Duruoz Hand Index > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	Serious[d]	serious[c,d]	none	23	21	-	MD 5.9 higher (2.75 lower to 14.55 higher)	⨁⨁◯◯ Low	IMPORTANT Not significant

Disease activity: Morning stiffness minutes > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	serious[e]	serious[c,d]	none	35	34	-	MD 11.2 higher (9.29 lower to 31.69 higher)	⨁⨁◯◯ Low	IMPORTANT Not significant

Disease activity: Active Joint Count > 12 weeks (3 months to 6 months)

1	randomised trials	serious[b]	not serious	serious[e]	serious[c,d]	none	11	9	-	MD 0.07 higher (0.81 lower to 0.95 higher)	⨁◯◯◯ Very low	IMPORTANT Not significant

Quality-of-life: NHP > 12 weeks (3 months to 6 months)

1	randomised trials	serious[b]	not serious	not serious	serious[d]	none	23	21	-	MD 5.7 higher (72.43 lower to 83.83 higher)	⨁⨁◯◯ Low	IMPORTANT Not significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic exercise	Land exercise	Relative (95% CI)	Absolute (95% CI)		
Mental health status: AIMS2: Affect > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	35	34	-	MD 0.3 higher (4.43 lower to 5.03 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant
Work Status: AIMS2: Work > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	35	34	-	MD 0.5 lower (1.73 lower to 0.73 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant
Work status: SODA > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	23	21	-	MD 0.2 higher (5.12 lower to 5.52 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant

CI: confidence interval; **MD:** mean difference

Explanations

a. Assessor not blinded

b. Loss to follow up >25%

c,d. wide CI crossing 0

e. Proxy measure

Table 3 Aquatic Exercise vs Warm-Water Immersion

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic	Warm water immersion	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from Knee range of motion > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	35	-	MD 2.7 lower (14.24 lower to 8.84 higher)	⊕⊕◯◯ Low	CRITICAL Not significant
Function as inferred from Wrist range of motion > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	35	-	MD 10.1 higher (17.72 lower to 37.92 higher)	⊕⊕◯◯ Low	CRITICAL Not significant
Function as inferred from Grip strength > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	35	-	MD 25.6 higher (3.09 lower to 54.29 higher)	⊕⊕◯◯ Low	CRITICAL Not significant
Functional status: AIMS2: Physical Capacity > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0.3 higher (0.54 lower to 1.14 higher)	⊕⊕⊕◯ Moderate	CRITICAL Not significant
Pain: McGill Pain Questionnaire_Sensory Pain > 12 weeks (3 months to 6 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic	Warm water immersion	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0.01 higher (0.36 lower to 0.38 higher)	⨁⨁⨁◯ Moderate	CRITICAL Not significant

Pain:McGill Pain Questionnaire_Affective Pain > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0 (0.82 lower to 0.82 higher)	⨁⨁⨁◯ Moderate	CRITICAL Not significant

Pain: AIMS2: Pain > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0.3 higher (0.85 lower to 1.45 higher)	⨁⨁⨁◯ Moderate	CRITICAL Not significant

Disease activity: Ritchie articular index > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	35	-	MD 0.3 lower (5.54 lower to 4.94 higher)	⨁⨁◯◯ Low	IMPORTANT Not significant

Disease activity: Morning stiffness minutes > 12 weeks (3 months to 6 months)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	35	35	-	MD 4.3 higher (15.52 lower to 24.12 higher)	⨁⨁◯◯ Low	IMPORTANT Not significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Aquatic	Warm water immersion	Relative (95% CI)	Absolute (95% CI)		
Mental health status: AIMS2: Affect > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0 (0.64 lower to 0.64 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant
Mental health status: AIMS2: Social > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0.1 higher (0.49 lower to 0.69 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant
Work status: AIMS2: Work > 12 weeks (3 months to 6 months)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	35	35	-	MD 0.3 higher (0.78 lower to 1.38 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Not significant

CI: confidence interval; **MD:** mean difference

Explanations

a. wide CI crossing 0

b. Proxy measure

Table 4. Additional Data from RCT and Observational Studies

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1665 Baillet 2009	RCT	12 months	50 participants with RA Dynamic exercise program: (mean age = 51.6 years, mean disease duration = 10.5, 84% female) Conventional joint rehab: (mean age = 56.3 years, mean disease duration = 11.7, 78% female)	Interventions were 4 weeks -Dynamic exercise program (DEP) (n=25): individualized multicomponent intervention consisting of occupational therapy program including dexterity exercises and splinting and physical therapy program including aquatic exercises (60 min/day) and cycling/running/resisting pulley cord (45 min/day) (full desription in Table 2); sessions were 5 hours per day that were led by PT, OT, or rheumatologist. -Conventional joint rehab (n=23): 3-day multidisciplinary program (~20 hours) that focused on education on disease pathogenesis, RA managment, and joint protection; exercises were also perfomed; conducted in groups of 4-5 participants with individual	Radiographic progression: Simple Erosions Narrowing Score (SENS) (higher score=worse progression) at 12 months The authors state that there were no differences between the groups at 12 months for SENS and they did not observe significant worsening in SENS score in DEP compared to control No adverse effects were observed in either group

				discussion at end of each day	
1889 Siqueira 2017	RCT	16 weeks	100 participants with RA (100% women) Water-based: mean age = 55 years, mean disease duration = 9.2 years Land-based: mean age = 54 years, mean disease duration = 7.7 years Control: mean age = 53.2 years, mean disease duration = 8.5 years	Interventions were 16 weeks -Water-based (n=33): 11 lower extremity body weight exercises (mostly seated) in water; flotation noodles used for stabilization; performed 3 times per week for 16 weeks (15-30 min sessions) supervised by physical education professional -Land-based (n=33): 11 lower extremity body weight exercises (mostly seated) on land; performed 3 times per week for 16 weeks (15-30 min sessions) supervised by physical education professional -Control (n=34): did not participate in any physical activities and continued normal routines	Outcomes after 16 weeks Treatment-related harms: adverse events (total is included here; there are subtypes based on type in Table 7. Pain or joint swelling was the most common subtype) • Aquatic (n=33): 3 (9.1%) • Land (n=33): 14 (42.4%) • Control (n=34): 33 (97.1%) Long-term outcomes: mortality • Aquatic (n=33): 0 • Land (n=33): 1 (3%) • Control (n=34): 0
3276 Strenstrom 1991	Non randomized control trial	4 years	60 RA patients	Training group 4 years of 1x/wk aquatics exercise in a group of 5 with summer and holiday breaks; included range of motion, dynamic and static strength, muscle endurance, coordination,	No difference in: Ritchies articular index, larsens radiological index, soft tissue swelling, pain, outdoor walking 480m, indoor walking 12m, lifting, buttoning, leaning for distance, standing from chair. Training group R grip strength improved and comparison group decreased $p<0.01$. not significant for L grip strength

				balance, and relaxation exercises. Intense tempo >170% resting HR.	
8030, Everden, 2007	Randomized controlled trial	3 months	115 patients with RA, stable on DMARDs, no prior PT in 6 months	Intervention: hydrotherapy (aquatic exercise in a heated pool) for 6 weeks Control: land exercise for 6 weeks	No significant differences between groups in 10m walk time, HAQ score, quality of life, or pain score at 3-month followup. Of note, pain score increased significantly in both groups from baseline to 3 months post-treatment.

References:

1. Baillet, A., Payraud, E., Niderprim, V. A., Nissen, M. J., Allenet, B., Francois, P., ... & Gaudin, P. (2009). A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. *Rheumatology*, *48*(4), 410-415.

2. Bergman, K. D. (1999). Therapeutic Benefit of Aquaerobics for Individuals with Rheumatoid Arthritis. *The Journal of Aquatic Physical Therapy*, *7*(1), 26-27.

3. Bilberg, A., Ahlmen, M., & Mannerkorpi, K. (2005). Moderately intensive exercise in a temperate pool for patients with rheumatoid arthritis: a randomized controlled study. *Rheumatology*, *44*(4), 502-508.

4. Eversden, L., Maggs, F., Nightingale, P., & Jobanputra, P. (2007). A pragmatic randomised controlled trial of hydrotherapy and land exercises on overall well being and quality of life in rheumatoid arthritis. *BMC musculoskeletal disorders*, *8*(1), 1-7.

5. Hall, J., Skevington, S. M., Maddison, P. J., & Chapman, K. (1996). A randomized and controlled trial of hydrotherapy in rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, *9*(3), 206-215.

6. Rintala, P., Kettunen, H., & McCubbin, J. A. (1996). Effects of a water exercise program for individuals with rheumatoid arthritis. *Research in Sports Medicine: An International Journal*, *7*(1), 31-38.

7. Siqueira, U. S., Valente, L. G. O., de Mello, M. T., Szejnfeld, V. L., & Pinheiro, M. M. (2017). Effectiveness of aquatic exercises in women with rheumatoid arthritis: a randomized, controlled, 16-week intervention—the HydRA trial. *American journal of physical medicine & rehabilitation*, *96*(3), 167-175.

8. Smith, S. S. (1998). Therapeutic benefit of aquaerobics for individual with rheumatoid arthritis. *Physiotherapy Canada*, *50*, 40-46.

9. Stenström, C. H., Lindell, B., Swanberg, E., Swanberg, P., Harms-Ringdahl, K., & Nordemar, R. (1991). Intensive dynamic training in water for rheumatoid arthritis functional class II-a long-term study of effects. *Scandinavian journal of rheumatology*, *20*(5), 358-365.

PICO 6: Should patients with RA consistently engage in a resistance training exercise program?

Summary: We included 20 studies for this question; 18 were randomized controlled trials (RCTs), and two (Sul et al. (2020)[17], Joo et al. (2022)[7]) were prospective interventional controlled trials. The most common outcomes included in the studies were pain, functional status, disease activity, and quality of life, and only 1-2 studies were included for most individual outcomes. For this PICO, the only critical outcomes were pain and functional status.

Many exercise regimens included not only resistance exercise, but also other forms of exercise such as aerobic exercise. The interventions discussed in this summary were judged to *primarily* involve resistance exercise. However, due to the multicomponent nature, we separately provide a summary of all multicomponent exercise studies (PICO 4-5-6), and many studies discussed here are also discussed in that document.

Below, we summarize the evidence in 6 sections:
1) RCTs: Resistance exercise vs. no exercise (Table 1)
2) Nonrandomized study: Resistance exercise vs. no exercise (Table 2)
3) RCTs: Resistance exercise vs. aquatic exercise (Table 3)
4) RCTs: Resistance exercise vs. conservative exercise (Table 4)
5) RCTs: Resistance exercise (pilates) and aerobic exercise vs. aerobic exercise alone (Table 5)
6) RCTs: Resistance exercise (pilates) vs. aerobic exercise (Table 6)

The first comparison (resistance exercise vs. no exercise, see Table 1) had by far the largest number of outcomes (47 unique outcomes). Subgroups were established uniquely for this comparison as several studies included range of motion/stretching exercises rather than a completely inactive control. We did not differentiate (by comparison or subgroup) according to the type or amount of resistance exercise that was included in the interventions. The outcomes in which the resistance group was favored (over the control group) include pain at 6 weeks and ≥12 weeks, walking speed at 24 months, sit to stand at ≥12 weeks, grip strength at 5 years, 30-second arm curl test at 24 weeks, steps per day at 24 weeks, DAS-28 at ≥12 weeks, number of painful joints from 0-12 weeks, and morning stiffness at ≥12 weeks. Only joint pain (adverse effect) favored the control group (inactive control). All other outcomes had no significant difference between groups (including functional status, disease activity, quality of life, work status, and radiographic progression outcomes). The evidence for this comparison was very low or low certainty for most outcomes.

We next discuss the subgroup analyses (studies with a true no-exercise control group, separately from studies with a ROM/stretching control group). For many outcomes, results were similar between these two subgroups of studies (HAQ at ≥12 weeks, Walk test at ≥12 weeks, pain at ≥12 weeks, DAS-28 at ≥12 weeks). However, for both the sit to stand test at ≥12 weeks and morning stiffness at ≥12 weeks, resistance exercise was statistically significantly better than ROM/stretching, but resistance exercise was not statistically significantly better than no exercise. This

counterintuitive finding may be due to subtleties between interventions between the subgroups could contribute to these differences (e.g., perhaps studies using a ROM/stretching subgroup used more intensive forms of resistance exercise).

One non-randomized study (Sul et al., 2020)[17] was included comparing resistance exercise to no exercise, and it examined 9 outcomes. Two of the outcomes (SF-36 physical function and left lower extremity strength from 0-12 weeks) significantly favored resistance and the remaining outcomes had no significant differences between groups (very low certainty evidence). Another non-randomized study (Joo et al. (2022))[7] was also included comparing resistance exercise to no exercise, and it examined 5 outcomes (which were all combined with those from Sul et al. None of the outcomes favored a group (very low certainty evidence). (The addition of this study does not change any conclusions about the outcomes or the certainty of evidence). Due to the presence of many RCTs of this same comparison (see comparison #1 above), the results of this study are largely irrelevant, but we include its data for reference purposes.

For the third comparison (resistance vs. aquatic), one study (Siqueira et al., 2017)[15] was included which examined three outcomes. Two of the outcomes (HAQ and pain at 16 weeks) favored the aquatic group and the third outcome (DAS-28 at 16 weeks) had no significant differences between groups (very low to low certainty evidence).

For the fourth comparison (resistance vs. conservative), one study (van der Ende et al., 2000)[18] was included which examined 7 outcomes. The conservative exercise intervention included range of motion and isometric exercises. No statistically significant differences were found between groups in any of the outcomes (primarily moderate certainty evidence).

For the fifth and sixth comparisons, one study (Yentur et al. 2021)[20] was included, which was an RCT of 8 weeks. This study contained two comparisons: Pilates vs. Aerobic AND Pilates (5^{th} comparison) and Pilates vs. Aerobic (6^{th} comparison). In this study, the Pilates group received education about principles of Pilates and related stretches and strengthening exercises; 3 times per week for about 45 minutes per session. The Aerobic group walked on a treadmill (60-80% of max heart rate); 3 times per week for 30 minutes per session. The Combined group (Pilates and Aerobic) received both interventions (aerobic first and then Pilates after 15-minute rest); 3 times per week.

Pilates and Aerobic Exercise vs. Aerobic Exercise Only (5^{th} comparion)

Out of 8 outcomes, 7 outcomes significantly (6 Minute Walk Test, McGill Pain Questionnaire [Words subscale], Fatigue Severity Scale, RA Quality of Life, Pittsburgh Sleep Quality Index, and Beck Depression Inventory) or slightly (McGill Pain Questionnaire [VAS score]) favored the Combined Pilates and Aerobic group compared with Aerobic only [9]. Only the McGill Pain Questionnaire (Likert subscale) significantly favored the Aerobic only group [9]. From these findings, we can conclude that Pilates and Aerobic combined is more effective than Aerobic only.

Pilates vs. Aerobic Exercise (6^{th} comparison)

Out of 8 outcomes, the only outcome that favored the Pilates group was the Fatigue Severity Scale (statistically significant) [9]. For all other outcomes, the control group either slightly favored (6 Minute Walk Test, RA Quality of Life) or significantly favored (McGill Pain

Questionnaire Short Form [Words Subscale, VAS, and Likert Subscale], Pittsburgh Sleep Quality Index, and Beck Depression Inventory) the Aerobic group [9].

Overall, resistance exercise was favored in several outcomes related to pain, functional status, and disease activity, although significant differences were inconsistent across studies and quality of evidence was very low or low for many outcomes.

Overall Quality of evidence comparing resistance exercise to no exercise: Very low

Overall Quality of evidence comparing resistance exercise to aquatic exercise: Very Low

Overall Quality of evidence comparing resistance exercise to conservative exercise: Low

Overall Quality of evidence comparing resistance exercise (pilates) to aerobic exercise: Very Low

Table 1: Resistance exercise versus no exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
Functional Status – HAQ (12 weeks-24 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
7	randomised trials	serious[e]	not serious	not serious	serious[f]	none	161	175	-	SMD **0.13 lower** (0.35 lower to 0.08 higher) This corresponds to MD **0.084 lower** (0.23 lower to 0.052 higher) on a 0-3 scale	⨁⨁◯◯ Low	Critical No significant difference

Functional Status – HAQ (12 weeks-24 months) - Inactive control

4	randomised trials	very serious[g]	not serious	not serious	serious[f]	none	85	96	-	SMD **0.27 lower** (0.56 lower to 0.03 higher) This corresponds to MD **0.16 lower** (0.32 lower to 0.017 higher) on a 0-3 scale	⨁◯◯◯ Very low	Critical No significant difference

Functional Status – HAQ (24 weeks-24 months) - ROM/stretch control

3	randomised trials	serious[h]	not serious	not serious	serious[f]	none	76	79	-	SMD **0.02 higher** (0.3 lower to 0.34 higher) This corresponds to MD **0.063 higher** (0.95 lower to 1.07 higher) on a 0-3 scale	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS dexterity (1-10) (change: 0-12 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	16	19	-	MD **0.8 lower** (2.59 lower to 0.99 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS Mobility (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	17	20	-	MD **0.7 lower** (2.09 lower to 0.69 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS Physical activity (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	15	19	-	MD **0** (1.3 lower to 1.3 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS household activity (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	16	17	-	MD **0.1 lower** (0.59 lower to 0.39 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS social activity (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	15	20	-	MD **0.2 lower** (1.26 lower to 0.86 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS ADL (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	15	19	-	MD **0.2 lower** (0.74 lower to 0.34 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS Pain (1-10) (change: 0-12 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	16	19	-	MD **0.6 higher** (0.57 lower to 1.77 higher)	⨁⨁◯◯ Low	Critical No significant difference

Functional status: AIMS Arthritis impact (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	16	19	-	MD **0.6 lower** (2.17 lower to 0.97 higher)	⨁⨁◯◯ Low	Critical No significant difference

Function as inferred from walking speed (m/s)- 24 months

1	randomised trials	very serious[a]	not serious	serious[b]	serious[c]	none	31	31	-	MD **0.3 higher** (0.03 higher to 0.57 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from walking speed (m/s)- 24 months - ROM/stretch control

1	randomised trials	very serious[a]	not serious	serious[b]	serious[c]	none	31	31	-	MD **0.3 higher** (0.03 higher to 0.57 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from sit to stand (seconds)-12 weeks

1	randomised trials	very serious[d]	not serious	serious[b]	serious[c]	none	17	23	-	MD **7 lower** (12.89 lower to 1.11 lower)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from Grip strength (Kgf) (R) (24 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[f]	none	27	33	-	MD **1.43 higher** (3.13 lower to 5.99 higher)	⨁⨁◯◯ Low	Critical No significant difference

Function as inferred from Grip strength (Kgf) (L) (24 weeks)

1	randomised trials	not serious	not serious	serious[b]	serious[f]	none	27	33	-	MD **0.18 higher** (3.88 lower to 4.24 higher)	⨁⨁◯◯ Low	Critical No significant difference

Function as inferred from Grip-5yr

1	randomised trials	very serious[m]	not serious	serious[b]	serious[c]	none	29	30	-	MD **11.7 higher** (9.1 higher to 14.3 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from Grip-5yr - ROM/stretch control

1	randomised trials	very serious[m]	not serious	serious[b]	serious[c]	none	29	30	-	MD **11.7 higher** (9.1 higher to 14.3 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from 30-s arm curl test (# of reps) (24 weeks)

1	randomised trials	serious[k]	not serious	serious[b]	serious[c]	none	13	15	-	MD **4.3 higher** (1.02 higher to 7.58 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from 30-s arm curl test (# of reps) (24 weeks) - ROM/stretch control

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[k]	not serious	serious[b]	serious[c]	none	13	15	-	MD **4.3 higher** (1.02 higher to 7.58 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from 50ft Walk test (sec) (12 weeks-24 weeks)

2	randomised trials	very serious[j]	not serious	serious[b]	serious[f]	none	31	38	-	MD **1.54 lower** (3.14 lower to 0.05 higher)	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from 50ft Walk test (sec) (24 weeks) - ROM/stretch control

1	randomised trials	serious[k]	not serious	serious[b]	serious[f]	none	13	15	-	MD **2.12 lower** (4.42 lower to 0.18 higher)	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from 50ft Walk test (sec) (change: 0-12 weeks) - Inactive control

1	randomised trials	serious[k]	not serious	serious[b]	serious[f]	none	18	23	-	MD **1.00 lower** (3.22 lower to 1.22 higher)	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from Sit to stand test (12 weeks-24 weeks) (# of stands in 30 sec)

2	randomised trials	very serious[j]	serious[i]	serious[b]	serious[f]	none	29	31	-	MD **4.21 higher** (1.27 higher to 7.16 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from Sit to stand test (24 weeks) - ROM/stretch control (# of stands in 30 sec)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[k]	not serious	serious[b]	serious[c]	none	13	15	-	MD **4.40 higher** (1.36 higher to 7.44 higher)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Function as inferred from Sit to stand test (12 weeks) - Inactive control (# of stands in 30 sec)

1	randomised trials	serious[k]	not serious	serious[b]	serious[f]	none	16	16	-	MD **1.50 higher** (10.12 lower to 13.12 higher)	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from TUG (sec) (12 weeks)

1	randomised trials	serious[k]	not serious	serious[b]	serious[f]	none	16	16	-	MD **0.8 lower** (6.2 lower to 4.6 higher)	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from Fatigue: Hours before fatigue (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	serious[p]	serious[f]	none	19	21	-	MD **0.8 higher** (1.43 lower to 3.03 higher)	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from Fatigue VAS (0-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	serious[p]	serious[f]	none	17	18	-	MD **0.5 higher** (1.27 lower to 2.27 higher)	⨁◯◯◯ Very low	Critical No significant difference

Pain (VAS 0-10) (12 weeks-24 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
6	randomised trials	very serious[i]	not serious	not serious	not serious	none	175	187	-	SMD **0.27 lower** (0.48 lower to 0.06 lower) This corresponds to MD **0.74 lower** (1.32 lower to 0.17 lower) on a 0-10 scale	⨁⨁◯◯ Low	Critical Statistically Significant Favors Resistance

Pain (VAS 0-10) (12 months-24 months) - ROM/stretch control

2	randomised trials	very serious[i]	not serious	not serious	serious[f]	none	63	64	-	MD **0.86 lower** (1.67 lower to 0.05 lower)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Pain (VAS 0-10) (12 weeks-24 weeks) - Inactive control

4	randomised trials	very serious[i]	not serious	not serious	serious[f]	none	112	123	-	MD **0.67 lower** (1.33 lower to 0.01 lower)	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

Functional status: HAQ (0-3) 6 weeks

1	randomised trials	not serious	not serious	not serious	serious[f]	none	17	18	-	MD **0.1 lower** (0.42 lower to 0.22 higher)	⨁⨁⨁◯ Moderate	Critical No significant difference

Function as inferred from TUG (sec) (6 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[c]	none	17	18	-	MD **1.8 lower** (2.85 lower to 0.75 lower)	⨁⨁◯◯ Low	Critical No significant difference

Pain: VAS 6 weeks

1	randomised trials	not serious	not serious	not serious	serious[c]	none	17	18	-	MD **1.5 lower** (2.99 lower to 0.01 lower)	⨁⨁⨁◯ Moderate	Critical Statistically Significant Favors Resistance

Disease Activity - DAS28 (0-10) (16 weeks-24 moths)

5	randomised trials	serious[e]	not serious	not serious	not serious	none	136	146	-	MD **0.38 lower** (0.67 lower to 0.09 lower)	⨁⨁⨁◯ Moderate	Important Statistically Significant Favors Resistance

Disease Activity - DAS28 (0-10) (16 weeks-24 weeks) - Inactive control

2	randomised trials	very serious[a]	serious[i]	not serious	serious[f]	none	60	67	-	MD **0.32 lower** (0.88 lower to 0.24 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease Activity - DAS28 (0-10) (24 weeks-24 months) - ROM/stretch control

3	randomised trials	very serious[i]	not serious	not serious	serious[f]	none	76	79	-	MD **0.47 lower** (0.95 lower to 0.00 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from ESR (mm/hr) (24 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[m]	not serious	serious[o]	serious[f]	none	31	31	-	MD **4.5 lower** (9.82 lower to 0.82 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from ESR (mm/hr) (24 months) - ROM/stretch control

1	randomised trials	very serious[m]	not serious	serious[o]	serious[f]	none	31	31	-	MD **4.5 lower** (9.82 lower to 0.82 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from Ritchie Index (24 months0

1	randomised trials	very serious[m]	not serious	serious[o]	serious[f]	none	31	31	-	MD **0.8 lower** (2.78 lower to 1.18 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from Ritchie Index (24 months) - ROM/stretch control

1	randomised trials	very serious[m]	not serious	serious[o]	serious[f]	none	31	31	-	MD **0.8 lower** (2.78 lower to 1.18 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from Self-reported joint count (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	serious[o]	serious[f]	none	18	17	-	MD **5.4 lower** (12.52 lower to 1.72 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from Number of painful joints (change: 0-12 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[k]	not serious	serious[o]	serious[c]	none	19	22	-	MD **1.4 lower** (2.6 lower to 0.2 lower)	⨁◯◯◯ Very low	Important Statistically Significant Favors Resistance

Disease activity as inferred from Physician's joint count (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	serious[o]	serious[f]	none	19	22	-	MD **2.7 lower** (8.53 lower to 3.13 higher)	⨁◯◯◯ Very low	Important No significant difference

Disease activity as inferred from Morning stiffness (duration in minutes) (12 weeks-24 months)

2	randomised trials	very serious[i]	not serious	serious[o]	serious[c]	none	50	53	-	MD **19.24 units lower** (34.29 lower to 4.19 lower)	⨁◯◯◯ Very low	IMPORTANT Statistically Significant Favors Resistance

Disease activity as inferred from Morning stiffness (duration in minutes) (24 months) - ROM/stretch control

1	randomised trials	very serious[i]	not serious	serious[o]	serious[c]	none	31	31	-	MD **21.40 lower** (38.54 lower to 4.26 lower)	⨁◯◯◯ Very low	IMPORTANT Statistically Significant Favors Resistance

Disease activity as inferred from Morning stiffness (duration in minutes) (change: 0-12 weeks) - Inactive control

1	randomised trials	serious[i]	not serious	serious[o]	serious[e]	none	19	22	-	MD **12.00 lower** (43.39 lower to 19.39 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
Adverse effect: Joint pain (presence of pain) 12 weeks												
1	randomised trials	serious[k]	not serious	not serious	very serious[l]	none	9/16 (56.3%)	0/16 (0.0%)	**RR 19.00** (1.20 to 301.16)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	Important Statistically Significant Favors no exercise
QoL: SF-36 Functional capacity (0-100) (12 weeks-24 weeks)												
2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **10.50 higher** (2.34 lower to 23.34 higher)	⨁⨁◯◯ Low	Important No significant difference
QoL: SF-36: Social aspects (0-100) (12 weeks-24 weeks)												
2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	49	43	-	MD **3.01 higher** (9.47 lower to 15.49 higher)	⨁⨁◯◯ Low	Important No significant difference
QoL: SF-36: Vitality (0-100) (12 weeks-24 weeks)												
2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **0.41 higher** (8.73 lower to 9.55 higher)	⨁⨁◯◯ Low	Important No significant difference
QoL: SF-36: Emotional aspects (0-100) (12 weeks-24 weeks)												
2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **0.80 higher** (18.41 lower to 20.01 higher)	⨁⨁◯◯ Low	Important No significant difference
QoL: SF-36: Physical aspects limitation (0-100) (12 weeks-24 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **4.02 higher** (15.69 lower to 23.74 higher)	⨁⨁◯◯ Low	Important No significant difference

QoL: SF-36: General Health (0-100) (12 weeks-24 weeks)

2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **3.90 higher** (6.03 lower to 13.83 higher)	⨁⨁◯◯ Low	Important No significant difference

QoL: SF-36: Pain (0-100) (12 weeks-24 weeks)

2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **5.60 higher** (4.59 lower to 15.79 higher)	⨁⨁◯◯ Low	Important No significant difference

QoL: SF-36: Mental Health (0-100) (12 weeks-24 weeks)

2	randomised trials	serious[e]	not serious	not serious	serious[f]	none	43	49	-	MD **0.40 higher** (9.57 lower to 10.37 higher)	⨁⨁◯◯ Low	Important No significant difference

Work status: Physical Loading of Work (1-7) (24 months)

1	randomised trials	very serious[m]	not serious	not serious	serious[f]	none	31	31	-	MD **0.5 higher** (0.28 lower to 1.28 higher)	⨁◯◯◯ Very low	Important No significant difference

Work status: Physical Loading of Work (1-7) (24 months) - ROM/stretch control

1	randomised trials	very serious[m]	not serious	not serious	serious[f]	none	31	31	-	MD **0.5 higher** (0.28 lower to 1.28 higher)	⨁◯◯◯ Very low	Important No significant difference

Radiographic progression: Larsen Score (0-100) (24 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[m]	not serious	not serious	serious[f]	none	31	31	-	MD **1.6 lower** (3.32 lower to 0.12 higher)	⨁◯◯◯ Very low	Important No significant difference

Radiographic progression: Larsen Score (0-100) (24 months) - ROM/stretch control

1	randomised trials	very serious[m]	not serious	not serious	serious[f]	none	31	31	-	MD **1.6 lower** (3.32 lower to 0.12 higher)	⨁◯◯◯ Very low	Important No significant difference

Depression: AIMS Depression (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	16	20	-	MD **0.5 lower** (1.32 lower to 0.32 higher)	⨁⨁◯◯ Low	Important No significant difference

Anxiety AIMS Anxiety (1-10) (change: 0-12 weeks)

1	randomised trials	serious[k]	not serious	not serious	serious[f]	none	15	19	-	MD **0.8 lower** (1.78 lower to 0.18 higher)	⨁⨁◯◯ Low	Important No significant difference

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Several categories at high risk of bias

b. performance surrogate for functional status

c. Wide CI and low N (<200)

d. one category had high risk of bias

e. a few studies had a category with unclear or high risk of bias

f. CI overlaps with no effect line

g. most studies had one category with high risk of bias

h. several categories with unclear risk of bias

i. I-squared =50-70%

j. studies had several categories with high or unclear risk of bias

k. one category with high risk of bias and at least one category with unclear risk of bias

l. very wide CI

m. most categories had high risk of bias

n. several categories had unclear risk of bias

o. Not a direct measure of disease activity

p. Fatigue is an indirect measure of functional status

Table 2: Nonrandomized: Resistance exercise versus Inactive control

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nonrand: Resistance	inactive control	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from 6MWT (m) (12 weeks)												
2	observational studies	very serious[a]	not serious	serious[d]	serious[b]	none	42	30	-	MD **9.3 higher** (6.51 lower to 25.11 higher)	⨁◯◯◯ Very low	Critical No significant difference
Function as inferred from Sit to Stand (# of stands)-- Mean Change (0-12 weeks)												
1	observational studies	very serious[a]	not serious	serious[d]	serious[b]	none	18	17	-	MD **2.3 lower** (6.82 lower to 2.22 higher)	⨁◯◯◯ Very low	Critical No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nonrand: Resistance	inactive control	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from Borg Scale (0-10)-- Mean Change (0-12 weeks)												
1	observational studies	very serious[a]	not serious	serious[d]	serious[b]	none	18	17	-	MD **0.1 lower** (0.76 lower to 0.56 higher)	⨁◯◯◯ Very low	Critical No significant difference
Function as inferred from UE Strength (Rt) (12 weeks)												
2	observational studies	very serious[a]	serious[e]	serious[d]	serious[b]	none	42	30	-	SMD **0.25 lower** (1.06 lower to 0.55 higher) This corresponds to MD **2.7 lower** (11.45 lower to 5.94 higher) on a lb scale	⨁◯◯◯ Very low	Critical No significant difference
Function as inferred from UE Strength (Lt) (12 weeks)												
2	observational studies	very serious[a]	not serious	serious[d]	serious[b]	none	42	30	-	SMD **0.12 higher** (0.35 lower to 0.6 higher) This corresponds to MD **1.0 higher** (2.99 lower to 4.99 higher) on a lb scale	⨁◯◯◯ Very low	Critical No significant difference
Function as inferred from LE Strength (Rt) (12 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nonrand: Resistance	inactive control	Relative (95% CI)	Absolute (95% CI)		
2	observational studies	very serious[a]	not serious	serious[d]	serious[b]	none	42	30	-	SMD **0.38 higher** (0.1 lower to 0.85 higher) This corresponds to MD **2.67 higher** (0.70 lower to 5.98 higher) on a lb scale	⨁◯◯◯ Very low	Critical No significant difference

Function as inferred from LE Strength (Lt) (12 weeks)

2	observational studies	very serious[a]	not serious	serious[d]	serious[b]	none	42	30	-	SMD **0.66 higher** (0.17 higher to 1.14 higher) This corresponds to MD **4.80 higher** (1.24 higher to 8.30 higher) on a lb scale	⨁◯◯◯ Very low	Critical Statistically Significant Favors Resistance

QoL: SF36 Mental Health (0-100)- Mean Change (0-12 weeks)

1	observational studies	very serious[a]	not serious	not serious	serious[b]	none	18	17	-	MD **5.1 higher** (3.56 lower to 13.76 higher)	⨁◯◯◯ Very low	Important No significant difference

QoL: SF36 Physical function (0-100)- Mean Change (0-12 weeks)

1	observational studies	very serious[a]	not serious	not serious	serious[c]	none	18	17	-	MD **9.9 higher** (2.17 higher to 17.63 higher)	⨁◯◯◯ Very low	Important Statistically Significant Favors Resistance

CI: confidence interval; **MD:** mean difference

Explanations

a. several categories had high risk of bias

b. CI crosses no effect line

c. Wide CI and low N (<200)

d. performance measure is an indirect measure of functional status

e. i2 = 50-70%

Table 3: Resistance exercise versus Aquatic exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance exercise	Aquatic exercise	Relative (95% CI)	Absolute (95% CI)		
Functional status: HAQ (0-3) 16 weeks												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	33	33	-	MD **0.4 higher** (0.15 higher to 0.65 higher)	⊕◯◯◯ Very low	Critical Statistically Significant Favors Aquatic
Disease activity: DAS-28 (0-10) 16 weeks												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	33	33	-	MD **0.5 higher** (0.03 lower to 1.03 higher)	⊕◯◯◯ Very low	Important No significant difference
Pain: Patient global assessment (0-10) 16 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance exercise	Aquatic exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	serious[d]	serious[c]	none	33	33	-	MD **2 higher** (0.62 higher to 3.38 higher)	⨁◯◯◯ Very low	Important Statistically Significant Favors Aquatic

CI: confidence interval; **MD:** mean difference

Explanations

a. several categories in study were at high risk of bias

b. CI crosses the no effect line

c. Wide CI and low N (<200)

d. not a direct measurement of pain

Table 4: Resistance exercise versus Conservative exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance exercise	conservative exercise	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS (0-10), 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	30	24	-	MD **0.4 higher** (0.92 lower to 1.72 higher)	⨁⨁⨁◯ Moderate	Critical No significant difference
Functional status: HAQ (0-3), 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	20	-	MD **0.2 lower** (0.67 lower to 0.27 higher)	⨁⨁⨁◯ Moderate	Critical No significant difference
Function as inferred from 50ft walk test (sec), 12 weeks												
1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	28	23	-	MD **1.4 lower** (6.35 lower to 3.55 higher)	⨁⨁◯◯ Low	Critical No significant difference
Treatment-related harms, psychological strain, study period (# of dropouts) (~30 days)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	2/34 (5.9%)	0/30 (0.0%)	**RR 4.43** (0.22 to 88.74)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁⨁⨁◯ Moderate	Important No significant difference
Treatment-related harms, pain, study period (# of dropouts) (~30 days)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	2/34 (5.9%)	0/30 (0.0%)	**RR 4.43** (0.22 to 88.74)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁⨁⨁◯ Moderate	Important No significant difference
Disease activity: DAS (0-10), 12 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	24	20	-	MD **0.5 lower** (1.18 lower to 0.18 higher)	⨁⨁⨁◯ Moderate	Important No significant difference
Disease activity as inferred from Number of swollen joints, 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance exercise	conservative exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	28	23	-	MD **1 lower** (4.43 lower to 2.43 higher)	⨁⨁◯◯ Low	Important No significant difference

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

Explanations

a. CI crosses no effect line

b. not a direct measure of disease activity

c. performance measure is an indirect measure of functional status

Table 5: RCTs: Combined (Pilates & Aerobic Exercise) compared to Aerobic Exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Combined (Pilates & Aerobic)	Aerobic	Relative (95% CI)	Absolute (95% CI)		
Critical Outcomes <12 Weeks (8 weeks)												
FUNCTIONAL STATUS: 6 Min Walk Test (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	10	10	-	MD **7.1 higher** (4.4 higher to 9.8 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Combined**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Combined (Pilates & Aerobic)	Aerobic	Relative (95% CI)	Absolute (95% CI)		
FUNCTIONAL STATUS: inferred from Fatigue Severity Scale (9-63) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	serious[d]	serious[b]	none	10	10	-	MD **2.2 lower** (3.12 lower to 1.28 lower)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Combined**
PAIN: McGill Pain Questionnaire Short Form (Words Subscale) (0-45) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	10	10	-	MD **0.7 lower** (1.2 lower to 0.2 lower)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Combined**
PAIN: McGill Pain Questionnaire Short Form (VAS) (0-10) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	10	10	-	MD **0.1 lower** (0.31 lower to 0.11 higher)	⨁◯◯◯ Very low	CRITICAL No Significant Difference
PAIN: McGill Pain Questionnaire Short Form (Likert Subscale) (0-5) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	10	10	-	MD **0.2 higher** (0.06 higher to 0.34 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Aerobic Only**

Important Outcomes <12 Weeks (8 weeks)

QOL: RA Quality of Life (0-30) (Mean Change Scores) (8 weeks)

1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	10	10	-	MD **1.2 lower** (1.72 lower to 0.68 lower)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Combined**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Combined (Pilates & Aerobic)	Aerobic	Relative (95% CI)	Absolute (95% CI)		
SLEEP: Pittsburgh Sleep Quality Index (0-21) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	10	10	-	MD **0.2 lower** (0.38 lower to 0.02 lower)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Combined**
MENTAL HEALTH: Beck Depression Inventory (0-63) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	10	10	-	MD **1.1 lower** (1.46 lower to 0.74 lower)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Combined**

CI: confidence interval; **MD:** mean difference

Explanations

a. 1064 Revman Bias Table: 3H, 2L, 1U. No blinding and some selective reporting.

b. Single study.

c. Single study, and confidence interval spans across the null value.

d. Outcome is a surrogate measure.

Table 6: RCTs: Pilates compared to Aerobic Exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Pilates	Aerobic	Relative (95% CI)	Absolute (95% CI)		

Critical Outcomes <12 Weeks (8 weeks)

FUNCTIONAL STATUS: 6 minute walk test (Mean Change Scores) (8 weeks) [Meters]

1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	10	10	-	MD **2.9 lower** (6.27 lower to 0.47 higher)	⨁◯◯◯ Very low	CRITICAL No Significant Difference

FUNCTIONAL STATUS: inferred from Fatigue Severity Scale (9-63) (Mean Change Scores) (8 weeks)

1	randomised trials	very serious[a]	not serious	serious[d]	serious[c]	none	10	10	-	MD **0.9 lower** (1.72 lower to 0.08 lower)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Pilates**

PAIN: McGill Pain Questionnaire Short Form (words subscale) (0-45) (Mean Change Scores) (8 weeks)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	10	10	-	MD **0.8 higher** (0.46 higher to 1.14 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Aerobic**

PAIN: McGill Pain Questionnaire Short Form (VAS) (0-10) (Mean Change Scores) (8 weeks)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	10	10	-	MD **0.4 higher** (0.24 higher to 0.56 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Aerobic**

PAIN: McGill Pain Questionnaire Short Form (Likert subscale) (0-5) (Mean Change Scores) (8 weeks)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	10	10	-	MD **0.3 higher** (0.16 higher to 0.44 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Aerobic**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Pilates	Aerobic	Relative (95% CI)	Absolute (95% CI)		
Important Outcomes <12 Weeks (8 weeks)												
QOL: RA Quality of Life (0-30) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	10	10	-	MD **0.1 higher** (0.24 lower to 0.44 higher)	⨁◯◯◯ Very low	IMPORTANT No Significant Difference
SLEEP: Pittsburg Sleep Quality Index (0-21) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	10	10	-	MD **0.5 higher** (0.33 higher to 0.67 higher)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Aerobic**
MENTAL HEALTH: Beck Depression Inventory (0-63) (Mean Change Scores) (8 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	10	10	-	MD **0.9 higher** (0.6 higher to 1.2 higher)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Aerobic**

CI: confidence interval; **MD:** mean difference

Explanations

a. 1064 Revman Bias Table: 3H, 2L, 1U. No blinding and some selective reporting.

b. Single study, and confidence interval spans across the null value.

c. Single study.

d. Outcome is a surrogate measure.

Table 7: Additional Data (not in Revman)

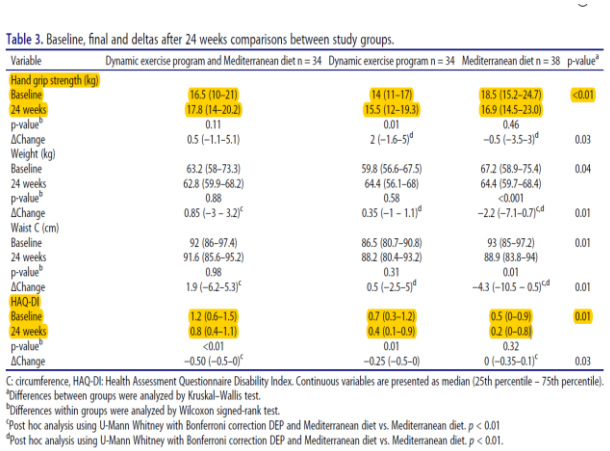
Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
803 Pineda-Juarez 2020	RCT, single blind	24 weeks	106 participants with active RA, all female and over 18 years old	MD Group - prescribed a Mediterranean diet, received general physical activity recommendations DEP/MD Group - received both interventions (a 3rd group received only DEP, which was included for data in PICO 1 for the comparison DEP vs MD+DEP)	

Table 3. Baseline, final and deltas after 24 weeks comparisons between study groups.

Variable	Dynamic exercise program and Mediterranean diet n = 34	Dynamic exercise program n = 34	Mediterranean diet n = 38	p-value[a]
Hand grip strength (kg)				
Baseline	16.5 (10–21)	14 (11–17)	18.5 (15.2–24.7)	<0.01
24 weeks	17.8 (14–20.2)	15.5 (12–19.3)	16.9 (14.5–23.0)	
p-value[b]	0.11	0.01	0.46	
ΔChange	0.5 (−1.1–5.1)	2 (−1.6–5)[d]	−0.5 (−3.5–3)[d]	0.03
Weight (kg)				
Baseline	63.2 (58–73.3)	59.8 (56.6–67.5)	67.2 (58.9–75.4)	0.04
24 weeks	62.8 (59.9–68.2)	64.4 (56.1–68)	64.4 (59.7–68.4)	
p-value[b]	0.88	0.58	<0.001	
ΔChange	0.85 (−3 – 3.2)[c]	0.35 (−1 – 1.1)[d]	−2.2 (−7.1–0.7)[c,d]	0.01
Waist C (cm)				
Baseline	92 (86–97.4)	86.5 (80.7–90.8)	93 (85–97.2)	0.01
24 weeks	91.6 (85.6–95.2)	88.2 (80.4–93.2)	88.9 (83.8–94)	
p-value[b]	0.98	0.31	0.01	
ΔChange	1.9 (−6.2–5.3)[c]	0.5 (−2.5–5)[d]	−4.3 (−10.5 – 0.5)[c,d]	0.01
HAQ-DI				
Baseline	1.2 (0.6–1.5)	0.7 (0.3–1.2)	0.5 (0–0.9)	0.01
24 weeks	0.8 (0.4–1.1)	0.4 (0.1–0.9)	0.2 (0–0.8)	
p-value[b]	<0.01	0.01	0.32	
ΔChange	−0.50 (−0.5–0)[c]	−0.25 (−0.5–0)	0 (−0.35–0.1)[c]	0.03

C: circumference, HAQ-DI: Health Assessment Questionnaire Disability Index. Continuous variables are presented as median (25th percentile – 75th percentile).
[a]Differences between groups were analyzed by Kruskal–Wallis test.
[b]Differences within groups were analyzed by Wilcoxon signed-rank test.
[c]Post hoc analysis using U-Mann Whitney with Bonferroni correction DEP and Mediterranean diet vs. Mediterranean diet. $p < 0.01$
[d]Post hoc analysis using U-Mann Whitney with Bonferroni correction DEP and Mediterranean diet vs. Mediterranean diet. $p < 0.01$.

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1022 Seneca 2015	RCT	12 weeks	51 participants with early RA (≤ 5 years) RA -Partly supervised exercises (n=25): (median age=61 years, median	Interventions were 12 weeks -Partly supervised (PS) exercises: 6 weeks of supervised training: 30-min bike (15-16 RPE), 30-min muscle strength	All results are median change scores (range) from baseline to 12 weeks Pain (NPRS): • PS group (n=15): -2.0 (-6.0 to 3.0) • SA group (n=21): 0.0 (-4.0 to 4.0)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
			disease duration=1 year, 68% female) -Self-administered exercises (n=26): (median age=54 years, median disease duration=1.5 years, 69% female)	training (legs, shoulders, trunk extensors/flexors) at 12 rep max; sessions were twice per week in groups of 2-4 patients supervised by same physical therapist; 6 weeks of self-administered exercise (see below) -Self-administered (SA) exercises: 12 weeks of self-administered exercises in local community (participants recommended to exercise at same intensity level as supervised group)	• Between groups p value: 0.263 Disease activity (DAS28-CRP) • PS group (n=15): -0.58 (-2.46 to 0.88) • SA group (n=21): 0.06 (-1.62 to 1.77) • Between groups p value: 0.006 Functional status (HAQ-DI) • PS group (n=21): 0.0 (-0.63 to 0.5) • SA group (n=24): 0.0 (-0.63 to 0.3) • Between groups p value: 0.972 Functional status (SF-36 physical component score) • PS group (n=21): 1.3 (-10.3 to 13.6) • SA group (n=24): 0.9 (-5.1 to 20.9) • Between groups p value: 0.802 Functional status (SF-36 mental component score) • PS group (n=21): 2.8 (-7.36 to 17.9) • SA group (n=24): -1.2 (-20.9 to 20.8) • Between groups p value: 0.089
1674 Ekdahl 1990	RCT	18 weeks	67 participants with RA (mean age=53 years; mean disease duration=10.6 years; 64% female)	Interventions were 6 weeks; all programs occurred at primary health center for one hour supervised by a physical therapist. All participants were encouraged to continue home program for another 3 months. The dynamic and static groups were collapsed in the results. -Dynamic (12 visits): multicomponent with	The following outcomes are change scores from baseline to 18 weeks (3 months after intervention) between the dynamic groups and the static groups The authors report that there were no significant differences between groups with the same type of program (dynamic 12 and 4 visits; static 12 and 4 visits) so the groups were combined. Pain (Pain intensity after muscle tests): • Dynamic groups: 0.0 • Static groups: 0.4 • Between groups P value >0.05 Pain (Pain intensity after bicycle ergometer) • Dynamic groups: -0.4 • Static groups: -0.2

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
				twice per week for 6 weeks; included cycling, body weight exercises for lower extremities, pulley exercises for hips; home program based on above -Dynamic (4 visits): similar exercises as above; 4 visits total -Static (12 visits): joint protection discussion, lower extremity mobility and body weight exercises; home program based on above -Static 4 visits): similar exercises as above; 4 visits total	• Between groups P value >0.05 Functional status (MF-index strength): • Dynamic groups: -1.2 • Static groups: -0.1 • Between groups P value <0.01 Functional status (MF-index endurance): • Dynamic groups: -1.8 • Static groups: 0.5 • Between groups P value <0.001 Functional status (MF-index balance/coordination): • Dynamic groups: 1.2 • Static groups: 0.9 • Between groups P value >0.05 Functional status (60-m walk test) • Dynamic groups: -1.9 • Static groups: 0.1 • Between groups P value >0.05 Functional status (Walking up/down 8 steps) • Dynamic groups: -2.7 • Static groups: -1.2 • Between groups P value <0.05 The authors report (giving no specific results of p values) that no significant differences between the dynamic and static groups were found in changes in pain intensity (during previous week), index of joint mobility, Ritchie total index, or indices of ADL, lower extremity ADL, ESR, and CRP from 0-18 weeks. There were significant differences between dynamic and static groups in 0-18 week change scores for Ritchie lower extremity index (p=0.01), Hb (p=0.01), and morning stiffness (p=0.002), indicating that the dynamic group reported fewer painful joints, had higher Hb values, and

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					reported fewer minutes of morning stiffness than the static group.
1889 Siqueira 2017	RCT	16 weeks	100 participants with RA (100% women) Water-based: mean age = 55 years, mean disease duration = 9.2 years Land-based: mean age = 54 years, mean disease duration = 7.7 years Control: mean age = 53.2 years, mean disease duration = 8.5 years	Interventions were 16 weeks -Water-based (n=33): 11 lower extremity body weight exercises (mostly seated) in water; flotation noodles used for stabilization; performed 3 times per week for 16 weeks (15-30 min sessions) supervised by physical education professional -Land-based (n=33): 11 lower extremity body weight exercises (mostly seated) on land; performed 3 times per week for 16 weeks (15-30 min sessions) supervised by physical education professional	Outcomes after 16 weeks Treatment-related harms: adverse events (total is included here; there are subtypes based on type in Table 7. Pain or joint swelling was the most common subtype) • Aquatic (n=33): 3 (9.1%) • Land (n=33): 14 (42.4%) • Control (n=34): 33 (97.1%) Long-term outcomes: mortality • Aquatic (n=33): 0 • Land (n=33): 1 (3%) • Control (n=34): 0

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
				-Control (n=34): did not participate in any physical activities and continued normal routines	

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
4678 Hakkinen 1998	RCT	6 month strength training, final assessment at 6 months and 42 months. 42 was last comparison end point	Articles says there were 38 RA patients and 5 PsA total but does not describe how the 5 PsA were distributed across groups Experimental (n = 21, mean age = 41.4 yrs, 47.6% female, mean symptom duration 10.5 months) Control (n = 18, mean age 45.6 years, 55.6% female, mean symptom duration 18.5 months) Healthy control (n = 18, mean age =41.8 years, 50% female)	Experimental – strength training 2-3 times/wk for all major upper and lower extremity muscle groups using rubber bands for resistance (theraband). Load increased progressively at 2 month intervals. During last 2 months, at 70-80% of repetition maximum load. After 6months to next 36 months, not supervised and did habitual activities Control and healthy controls maintained habitual physical activities (walking, biking, swimming)	Experimental group at 42 mo (median and inter-quartile range) Ritchies articular index 6 (0-10) Larsen's index 7.5 (1-16) HAQ 0 (0-6) Control group at 42 mo (median and inter-quartile range) Ritchies articular index 4(2-10) Larsen's index 13 (4-20) HAQ 2.5 (0-8)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
6071, Hakkinen, 2004 Affiliated studies: 1990 2891 5147	RCT	5 year fu	62 RA patients	Dynamic resistance training (RT) versus control	The respective median (IQR) Larsen scores in the RT and CG were 0 (0–1), 0 (0–2), 0 (0–3), and 0 (0–3), 2 (0–5), 2 (0–4) at baseline, and at the 2 and 5 year check ups, respectively.
2059, Strasser, 2011	RCT	6 month	40 patients	Strength and endurance versus control	After the 6-month training period, no significant changes inthe maximum strength (1RM) and maximum workload(Wmax) were observed between the groups.
7895, Wessel, 1984	3 arm RCT	7 weeks	32	Isometric ex versus isokinetic ex versus control	"The pain experienced by the isometric group before and after treatment sessions was significantly higher than that experienced by the isokinetic group." **Pre- and post-test mean values of the pain-rating index (score derived from the ranked values of word descriptors of pain)** **Group / Pretest / Post-test** Control 22.3 19.1 Isometric 27.3 24.8 Isokinetic 23.6 18.1* *significantly different from pre-test

References:

1. Ekdahl, C., et al. (1990). "Dynamic versus static training in patients with rheumatoid arthritis." Scandinavian Journal of Rheumatology **19**(1): 17-26.

2. Häkkinen, A., et al. (2004). "A home-based two-year strength training period in early rheumatoid arthritis led to good long-term compliance: a five-year followup." Arthritis Care & Research **51**(1): 56-62.

3. Häkkinen, A., et al. (2004). "Sustained maintenance of exercise induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5 year follow up." Annals of the Rheumatic Diseases **63**(8): 910-916.

4. Häkkinen, A., et al. (2001). "A randomized two-year study of the effects of dynamic strength training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis." Arthritis & Rheumatism **44**(3): 515-522.

5. Häkkinen, A., et al. (1999). "Dynamic strength training in patients with early rheumatoid arthritis increases muscle strength but not bone mineral density." The Journal of Rheumatology **26**(6): 1257-1263.

6. Häkkinen, A., et al. (2003). "Effects of dynamic strength training on physical function, Valpar 9 work sample test, and working capacity in patients with recent-onset rheumatoid arthritis." Arthritis Care & Research: Official Journal of the American College of Rheumatology **49**(1): 71-77.

7. Joo, Y. B., Lee, K. B., Sul, B., Lee, H.-S., Lim, S. H., & Park, Y.-J. (2022). Effect of resistance exercise on serum leptin levels in a prospective longitudinal study of women patients with rheumatoid arthritis. Arthritis Research & Therapy, 24(1), 1-9.

8. Komatireddy, G., et al. (1997). "Efficacy of low load resistive muscle training in patients with rheumatoid arthritis functional class II and III." The Journal of Rheumatology **24**(8): 1531-1539.

9. Lemmey, A. B., et al. (2009). "Effects of high-intensity resistance training in patients with rheumatoid arthritis: A randomized controlled trial." Arthritis Care & Research **61**(12): 1726-1734.

10. Lourenzi, F. M., et al. (2017). "Effectiveness of an overall progressive resistance strength program for improving the functional capacity of patients with rheumatoid arthritis: a randomized controlled trial." Clinical Rehabilitation **31**(11): 1482-1491.

11. McMeeken, J., et al. (1999). "The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis." Physiotherapy Research International **4**(1): 55-67.

12. Pineda-Juárez, J. A., et al. (2022). "Changes in hand grip strength and body weight after a dynamic exercise program and Mediterranean diet in women with rheumatoid arthritis: A randomized clinical trial." Physiotherapy Theory and Practice **38**(4): 504-512.

13. Rodrigues, R., et al. (2020). "Low-Load Resistance Training With Blood-Flow Restriction in Relation to Muscle Function, Mass, and Functionality in Women With Rheumatoid Arthritis." Arthritis Care & Research **72**(6): 787-797.

14. Seneca, T., et al. (2015). "Comparable effect of partly supervised and self-administered exercise programme in early rheumatoid arthritis--a randomised, controlled trial." Dan Med J **62**(8): A5127.

15. Siqueira, U. S., et al. (2017). "Effectiveness of aquatic exercises in women with rheumatoid arthritis: a randomized, controlled, 16-week intervention—the HydRA trial." American Journal of Physical Medicine & Rehabilitation **96**(3): 167-175.

16. Strasser, B., et al. (2011). "The effects of strength and endurance training in patients with rheumatoid arthritis." Clinical Rheumatology **30**(5): 623-632.

17. Sul, B., et al. (2020). "Twelve weeks of strengthening exercise for patients with rheumatoid arthritis: A prospective intervention study." Journal of Clinical Medicine **9**(9): 2792.

18. Van den Ende, C., et al. (2000). "Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial." Annals of the Rheumatic Diseases **59**(8): 615-621.

19. Wessel, J. and H. Quinney (1984). "Pain experienced by persons with rheumatoid arthritis during isometric and isokinetic exercise." Physiotherapy Canada **36**(3): 131-134.

20. Yentür SB, Ataş N, Öztürk MA, Oskay D. Comparison of the effectiveness of pilates exercises, aerobic exercises, and pilates with aerobic exercises in patients with rheumatoid arthritis. Irish Journal of Medical Science. 2021;190:1027-1034.

PICO 4-5-6: Should patients with RA consistently engage in a combined exercise program?

Evidence Summary: A review of the literature revealed 22 studies that evaluated exercise regimens that were a combination of aerobic exercise, resistance exercises, and/or water-based exercises. One of these studies was non-randomized, while the other 21 were randomized controlled trials that compared an active group to either no exercise (Table 1 RCTs and Table 2 non-RCT), or compared an "intense" exercise regimen to a standard regimen (Table 3), or compared combination exercise to an active control such as basic range of motion exercises or isometric exercises for the large joints (Table 4). The critical outcomes for this PICO were pain and function.

- For RCTs comparing combination exercise to no exercise (Table 1), of the 13 statistical tests involving the critical outcomes of pain or function, six were statistically significant in favor of combination exercise, and the other seven were statistically non-significant (Table 1).
- For the non-RCT of this comparison (Table 2), the one critical outcome (function as measured by HAQ) was statistically significant in favor of combination exercise.
- When combination exercise was compared with conservative exercise (i.e., less intense), neither critical outcome (pain, function) was statistically significant (Table 3)
- When combination exercise was compared with active control (Table 4), only 1 of 10 statistical tests of critical outcomes were statistically significant.

Data for studies that did not report sufficient information for the calculation of effect sizes and standard errors) appear at the end of the document.

Many studies suffered from similar methodological problems, including inability to blind the participants, personnel, and assessors, largely due to the nature of the interventions.

Quality of evidence across critical outcomes: Low

Table 1: Combination exercise vs no exercise (RCTs)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
HAQ >12 weeks												
4	randomised trials	serious[a]	not serious	not serious	serious[b]	none	308	311	-	MD **0.06 lower** (0.13 lower to 0)	⨁⨁◯◯ Low	CRITICAL Not statistically significant
HAQ >12 weeks - low accountability/contact												
3	randomised trials	serious[a]	not serious	not serious	serious[b]	none	172	166	-	MD **0.07 lower** (0.19 lower to 0.04 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant
HAQ >12 weeks - High accountability/contact												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	136	145	-	MD **0.07 lower** (0.15 lower to 0.01 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant
SF-36 physical function >12 weeks												
3	randomised trials	serious[a]	not serious	not serious	serious[c]	none	168	156	-	MD **0.53 higher** (0.04 higher to 1.01 higher)	⨁⨁◯◯ Low	CRITICAL Statistically significant improvement in outcome, favoring combination exercise
SF-36 physical function - low accountabiilty/contact >12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious[a]	not serious	not serious	serious[c]	none	132	129	-	MD **1.84 higher** (2.49 lower to 6.18 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

Pain, 12 weeks (VAS 0-100)

1	randomised trials	serious[a]	not serious	not serious	not serious	none	40	38	-	MD **18.8 lower** (29.66 lower to 7.94 lower)	⨁⨁⨁◯ Moderate	CRITICAL Statistically significant improvement in outcome, favoring combination exercise

Functional status: MACTAR, 12-24 months (-38 = maximal deterioration, +38 = maximal improvement)

2	randomised trials	serious[d]	not serious	not serious	not serious	none	213	220	-	MD **2.43 higher** (0.68 higher to 4.19 higher)	⨁⨁⨁◯ Moderate	CRITICAL Statistically significant improvement in outcome, favoring combination exercise

ASES Function, 22 weeks (0-5; greater scores indicate better self-efficacy)

1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	17	15	-	MD **0.19 higher** (0.14 lower to 0.52 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

ASES pain and other symptoms, 22 weeks (0-5; greater scores indicated better self-efficacy)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	17	15	-	MD **0.14 higher** (0.41 lower to 0.69 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

Dutch-AIMS2 physical health, 22 weeks (0-10, with lower scores indicating better health)

1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	17	15	-	MD **0.54 lower** (1.08 lower to 0)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

SF-36 physical function - high accountability/contact >12 weeks

1	randomised trials	serious[a]	not serious	not serious	not serious[c]	none	36	27	-	MD **0.51 higher** (0.03 higher to 0.99 higher)	⨁⨁⨁◯ Moderate	CRITICAL Statistically significant improvement in outcome, favoring combination exercise

Function as inferred from Fatigue Severity Scale (FSS) 12 months (1-7, with higher scores indicating more fatigue)

1	randomised trials	serious[a]	not serious	serious	not serious	none	40	38	-	MD **9.2 lower** (17.1 lower to 1.3 lower)	⨁⨁◯◯ Low	CRITICAL Statistically significant improvement in outcome, favoring combination exercise

Disease activity as inferred from Stiffness (VAS 0-100) (12 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious	not serious	none	40	38	-	MD **18.4 lower** (31.05 lower to 5.75 lower)	⨁⨁◯◯ Low	IMPORTANT Statistically significant improvement in outcome, favoring combination exercise

SF-36 mental health >12 weeks

3	randomised trials	serious[a]	not serious	not serious	serious[c]	none	168	156	-	MD **0.59 higher** (0.13 higher to 1.05 higher)	⨁⨁◯◯ Low	IMPORTANT Statistically significant improvement in outcome, favoring combination exercise

SF-36 mental health - low accountability/contact >12 weeks

2	randomised trials	serious[a]	not serious	not serious	serious[c]	none	132	129	-	MD **0.29 higher** (4.2 lower to 4.78 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

SF-36 mental health - high accountability/contact >12 weeks

1	randomised trials	serious[a]	not serious	not serious	not serious	none	36	27	-	MD **0.59 higher** (0.13 higher to 1.05 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Statistically significant improvement in outcome, favoring combination exercise

Disease activity, DAS28, 6-12 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious[d]	not serious	not serious	serious[e]	none	94	90	-	MD **0.3 higher** (0.22 lower to 0.82 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

Dutch-AIMS2 pyschological health, 22 weeks (0-10; lower scores indicated better health)

1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	17	15	-	MD **0.42 lower** (1.29 lower to 0.45 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

Dutch-AIMS2 social interaction, 22 weeks (0-10; lower scores indicate better social interaction)

1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	17	15	-	MD **0.4 higher** (0.97 lower to 1.77 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

Pittsburgh Sleep Quality Index, 12 months (0-21, with higher scores indicating worse sleep quality)

1	randomised trials	serious[a]	not serious	not serious	serious[f]	none	40	38	-	MD **0.8 higher** (0.82 lower to 2.42 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

Radiographic damage: Larsen score for large joints, 24 months (0-60, with higher scores representing increased joint damage)

1	randomised trials	serious[d]	not serious	not serious	serious[g]	none	136	145	-	MD **0** (0.23 lower to 0.23 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
Radiographic progression: Number with relevant progression, 24 months												
1	randomised trials	serious[d]	not serious	not serious	not serious	none	20/136 (14.7%)	15/145 (10.3%)	**OR 1.49** (0.73 to 3.05)	**43 more per 1,000** (from 26 fewer to 157 more)	⨁⨁⨁◯ Moderate	IMPORTANT Not statistically significant
Mental health: HADS, 24 months (0-42, with higher scores indicating increased anxiety and/or depression)												
1	randomised trials	serious[d]	not serious	not serious	not serious	none	136	145	-	MD **1.3 lower** (2.25 lower to 0.35 lower)	⨁⨁⨁◯ Moderate	IMPORTANT Statistically significant improvement in outcome, favoring combination exercise
Disease activity: DAS4 (Ritchie index + number swollen joints), 24 months												
1	randomised trials	serious[d]	not serious	serious[h]	serious[e]	none	136	145	-	MD **0.2 lower** (0.47 lower to 0.07 higher)	⨁◯◯◯ Very low	IMPORTANT Not statistically significant
Radiographic damage: Feet only, 24 months (Larsen scoring)												
1	randomised trials	serious[d]	not serious	not serious	serious[g]	none	136	145	-	MD **0.8 lower** (1.6 lower to 0)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant
Radiographic damage: Hands only, 24 months (Larsen scoring)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT aerobic/resistance ex	inactive comparator > 12wks	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[d]	not serious	not serious	serious[g]	none	136	145	-	MD **1.3 lower** (3.1 lower to 0.5 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant
RAQol score, 24 months (scores 0-30; lower scores indicate better quality of life)												
2	randomised trials	serious[d]	not serious	not serious	serious[c]	none	132	128	-	MD **0.94 lower** (2.01 lower to 0.13 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant
SF-36 global health												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	36	27	-	MD **0.72 higher** (0.23 higher to 1.21 higher)	⨁⨁⨁◯ Moderate	IMPORTANT Statistically significant improvement in outcome, favoring combination exercise

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio

Explanations

a. participants and assessors not blinded. Also selective reporting of endpoints in this study

b. Confidence intervals cross 0, indicating no change in mean difference of HAQ

c. Confidence intervals cross 0, indicating no change in functional assessment

d. Participants and personnel not blinded

e. Confidence intervals cross 0, indicating no change in disease activity

f. confidence interval crosses 0, indicating no change in sleep quality

g. Confidence interval crosses 0, indicating no change in radiographic assessment

h. Outcomes measure is not one of the preferred disease activity measures

Table 2: Combination exercise vs no exercise (non-RCT)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRCT aerobic/resistance ex	inactive comparator >12 weeks	Relative (95% CI)	Absolute (95% CI)		
Function, **HAQ, 6 months**												
1	observational studies	serious[a]	not serious	not serious	not serious	none	18	18	-	MD **0.98 SD lower** (1.67 lower to 0.28 lower)	⨁◯◯◯ Very low	CRITICAL Statistically significant change, favoring combination exercise
Disease activity, **DAS28, 6 months**												
1	observational studies	serious[a]	not serious	not serious	not serious	none	18	18	-	MD **0.61 SD lower** (1.28 lower to 0.06 higher)	⨁◯◯◯ Very low	IMPORTANT Not statistically significant

CI: confidence interval; **MD:** mean difference

Explanations

a. Poor random sequence generation and allocation concealment. No blinding of patients and personnel

Table 3: Combination intensive exercise vs conservative exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intense exercise program	conservative exercise	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS, 12 weeks (0-10cm)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	24	-	MD **0.4 higher** (0.92 lower to 1.72 higher)	⨁⨁◯◯ Low	CRITICAL Findings not statistically significant
Functional status: HAQ, 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	25	20	-	MD **0.2 lower** (0.67 lower to 0.27 higher)	⨁⨁◯◯ Low	CRITICAL Findings not statistically significant
Functional performance, inferred from 50ft walk test (sec), 12 weeks												
1	randomised trials	serious[a]	not serious	serious[e]	serious[f]	none	28	23	-	MD **1.4 lower** (6.35 lower to 3.55 higher)	⨁◯◯◯ Very low	CRITICAL Findings not statistically significant
Treatment-related harms, pain, study period (~30 days)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intense exercise program	conservative exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	2/34 (5.9%)	0/30 (0.0%)	**RR 4.43** (0.22 to 88.74)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL Findings not statistically significant

Disease activity: DAS, 12 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	24	20	-	MD **0.5 lower** (1.18 lower to 0.18 higher)	⨁⨁◯◯ Low	IMPORTANT Findings not statistically significant

Disease activity, inferred from number of swollen joints, 12 weeks

1	randomised trials	serious[a]	not serious	serious[d]	serious[b]	none	28	23	-	MD **1 lower** (4.43 lower to 2.43 higher)	⨁◯◯◯ Very low	IMPORTANT Findings not statistically significant

Treatment-related harms, psychological strain, study period (~30 days)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	2/34 (5.9%)	0/30 (0.0%)	**RR 4.43** (0.22 to 88.74)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	IMPORTANT Findings not statistically significant

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

Explanations

a. Study participants and personnel were not blinded to the intervention

b. Confidence intervals cross 0, indicating no change in disease activity

c. Very few adverse events and CI crosses 1

d. The outcome measure is not a complete measurement of disease activity, and is used as a subdomain in most established outcomes measures for disease activity

e. The outcome measure is not a full assessment of function

f. Confidence interval crosses 0, indicating no change in functional performance measure

Table 4: RCT- aerobic/resistance v active comparator >12wks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT- aerobic/resistance >12wks	active comparator	Relative (95% CI)	Absolute (95% CI)		
Pain, VAS 0-100mm, 12 weeks												
2	randomised trials	serious[a]	serious[b]	not serious	serious[c]	none	96	92	-	MD **3.72 higher** (6.7 lower to 14.14 higher)	⨁◯◯◯ Very low	CRITICAL Not statistically significant
HAQ, >12 weeks												
4	randomised trials	serious[a]	not serious	not serious	serious[c]	none	155	150	-	MD **0.02 lower** (0.1 lower to 0.06 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant
Function, AIMS2-SF, 12 months (0-60)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT-aerobic/resistance >12wks	active comparator	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	23	21	-	MD **1.1 higher** (3.07 lower to 5.27 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

Function inferred from SODA, 12 months (0-108; higher scores indicated better hand function)

1	randomised trials	serious[a]	not serious	serious	serious[c]	none	23	21	-	MD **0.2 higher** (5.12 lower to 5.52 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

Functional Status, inferred from Performance Measure - VO2/kg/min, ml (Baseline - 20 weeks)

1	randomised trials	serious[a]	not serious	serious[g]	not serious	none	36	37	-	MD **1.28 SD higher** (0.78 higher to 1.79 higher)	⨁⨁◯◯ Low	CRITICAL Statistically significant difference in favor of combination exercise

Functional Status: as inferred from TUG (timed up-and-go, seconds) >20 weeks

2	randomised trials	serious[a]	not serious	serious	serious[c]	none	60	60	-	MD **0.25 SD lower** (0.73 lower to 0.24 higher)	⨁◯◯◯ Very Low	CRITICAL Not statistically significant

Function inferred from Duruoz Hand Index, 12 months (0-90, with higher scores indicating decreased function)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT-aerobic/resistance >12wks	active comparator	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious	serious[f]	none	23	21	-	MD **5.9 higher** (2.75 lower to 14.55 higher)	⨁◯◯◯ Very Low	CRITICAL Not statistically significant

Function as inferred from 50ft walk test, 24 weeks (seconds)

1	randomised trials	serious[a]	not serious	serious	serious[c]	none	72	69	-	MD **0.57 lower** (1.2 lower to 0.06 higher)	⨁◯◯◯ Very Low	CRITICAL Not statistically significant

Functional Status:Performance Measure - Endurance, minutes, 12 months (lower times indicate better enduranc)

2	randomised trials	serious[a]	serious[b]	not serious	serious[c]	none	60	60	-	MD **0.19 SD higher** (0.52 lower to 0.89 higher)	⨁◯◯◯ Very low	CRITICAL Not statistically significant

Function as inferred from fatigue (4 years)

1	randomised trials	serious[a]	not serious	serious	serious[c]	none	24	23	-	MD **0.35 SD lower** (0.92 lower to 0.23 higher)	⨁◯◯◯ Very Low	CRITICAL Not statistically significant

Functional Status: Performance Measure - Sit-to-stand (higher number indicates better function)

2	randomised trials	serious[a]	not serious	not serious	serious[c]	none	60	60	-	MD **0.16 SD higher** (0.2 lower to 0.52 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT-aerobic/resistance >12wks	active comparator	Relative (95% CI)	Absolute (95% CI)		
Disease activity, inferred from swollen joint count, 24 weeks												
1	randomised trials	serious[a]	not serious	serious[d]	not serious	none	72	69	-	MD **1.49 lower** (2.37 lower to 0.6 lower)	⨁⨁◯◯ Low	IMPORTANT Statistically significant difference in favor of combination exercise
Disease activity, inferred from Ritchie articular index, 24 weeks												
1	randomised trials	serious[a]	not serious	serious[e]	serious[c]	none	72	69	-	MD **0.24 lower** (1.91 lower to 1.44 higher)	⨁◯◯◯ Very low	IMPORTANT Not statistically significant
Disease activity, inferred from global assessment of disease activity, 24 weeks (VAS 0-10 cm)												
1	randomised trials	serious[a]	not serious	serious[e]	serious[c]	none	72	69	-	MD **0.73 higher** (0.32 lower to 1.78 higher)	⨁◯◯◯ Very low	IMPORTANT Not statistically significant
Disease Activity - DAS28, >12 weeks												
3	randomised trials	serious[a]	serious[b]	not serious	serious[c]	none	71	68	-	MD **0.13 lower** (0.81 lower to 0.55 higher)	⨁◯◯◯ Very low	IMPORTANT Not statistically significant
Disease Activity - CDAI (20 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT-aerobic/resistance >12wks	active comparator	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[f]	none	24	24	-	MD **1.6 lower** (7.83 lower to 4.63 higher)	⨁⨁◯◯ Low	IMPORTANT Not statistically significant

CI: confidence interval; **MD:** mean difference

Explanations

a. Poor blinding of participants and study assessors

b. Large I^2 and heterogeneous effect direction

c. Effect confidence intervals cross 0

d. Outcome is a component of more accepted measurements of disease activity

e. Outcome measure not preferred for measurement of disease activity

f. large confidence interval, which crosses 0

g. Obscure functional measure. Unlikely to be pertinent

PICO question: 4-5-6. Studies of combination exercise for which effect sizes were not computable

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
3176, Stavropoulos-Kalinoglou,	Case-control	6 month	RA fulfilling 1987 revised American College of	Individualized exercise program versus control	At 6 months: Median and range

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
2013	matched design		Rheumatology criteria, sedentary lifestyle (no participation in structured exercise for the preceding 6 months), and stable disease (no changes in disease-modifying antirheumatic drugs (DMARDs) — including biologics — or oral steroids and no parenteral steroid administration in the last 3 months).		Exercise group CRP = 4.0 (3.0–8.0), Control CRP = 7.0 (3.0–15.0)
Ref ID, Author, year	**Study type**	**Duration**	**Population Description**	**Treatment given to relevant population**	**Results**
3563, Hansen, 1993	Randomized controlled trial	24 months	75 patients with RA, age 20-60, not exercising more than 2x per week at baseline	5 groups: Self-training, training with a PT, training as a group, training in a group and in a pool, versus no training (control)	No statistically significant difference in number of swollen joints, pain score, HAQ, x-ray progression, functional score, muscle strength, or aerobic fitness at 24 months between any of the groups.

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
3886, Kucharski, 2019	Sub-study of an RCT	52 weeks	74 patients with RA	Intervention: 20 weeks of gym-based aerobic and resistance exercise Control: light home-based exercise for 20 weeks	At 20 weeks, there was a statistically significant difference in physical fatigue (p = 0.002), mental fatigue (p = 0.048) and depression (p = 0.039) however at 52 weeks these differences were no longer statistically significant. There were no differences in general fatigue, reduced motivation, reduced activity, VAS-fatigue, or anxiety at either timepoint.
2059, Strasser, 2011	RCT	6 month	40 patients	Strength and endurance versus control	After the 6-month training period, no significant changes inthe maximum strength (1RM) and maximum workload(Wmax) were observed between the groups.
803 Pineda-Juarez 2020	RCT, single blind	24 weeks	106 participants with active RA, all female and over 18 years old	MD Group - prescribed a Mediterranean diet, received general physical activity recommendations DEP/MD Group - received both interventions (a 3rd group received only DEP, which was included for data in PICO 1 for the comparison DEP vs MD+DEP)	Table 3. Baseline, final and deltas after 24 weeks comparisons between study groups. Variable | Dynamic exercise program and Mediterranean diet n = 34 | Dynamic exercise program n = 34 | Mediterranean diet n = 38 | p-value[a] Hand grip strength (kg) Baseline | 16.5 (10–21) | 14 (11–17) | 18.5 (15.2–24.7) | <0.01 24 weeks | 17.8 (14–20.2) | 15.5 (12–19.3) | 16.9 (14.5–23.0) p-value[b] | 0.11 | 0.01 | 0.46 ΔChange | 0.5 (−1.1–5.1) | 2 (−1.6–5)[d] | −0.5 (−3.5–3)[d] | 0.03 Weight (kg) Baseline | 63.2 (58–73.3) | 59.8 (56.6–67.5) | 67.2 (58.9–75.4) | 0.04 24 weeks | 62.8 (59.9–68.2) | 64.4 (56.1–68) | 64.4 (59.7–68.4) p-value[b] | 0.88 | 0.58 | <0.001 ΔChange | 0.85 (−3 – 3.2)[c] | 0.35 (−1 – 1.1)[d] | −2.2 (−7.1–0.7)[c,d] | 0.01 Waist C (cm) Baseline | 92 (86–97.4) | 86.5 (80.7–90.8) | 93 (85–97.2) | 0.01 24 weeks | 91.6 (85.6–95.2) | 88.2 (80.4–93.2) | 88.9 (83.8–94) p-value[b] | 0.98 | 0.31 | 0.01 ΔChange | 1.9 (−6.2–5.3)[c] | 0.5 (−2.5–5)[d] | −4.3 (−10.5 – 0.5)[c,d] | 0.01 HAQ-DI Baseline | 1.2 (0.6–1.5) | 0.7 (0.3–1.2) | 0.5 (0–0.9) | 0.01 24 weeks | 0.8 (0.4–1.1) | 0.4 (0.1–0.9) | 0.2 (0–0.8) p-value[b] | <0.01 | 0.01 | 0.32 ΔChange | −0.50 (−0.5–0)[c] | −0.25 (−0.5–0) | 0 (−0.35–0.1)[c] | 0.03 C: circumference, HAQ-DI: Health Assessment Questionnaire Disability Index. Continuous variables are presented as median (25th percentile – 75th percentile). [a]Differences between groups were analyzed by Kruskal–Wallis test. [b]Differences within groups were analyzed by Wilcoxon signed-rank test. [c]Post hoc analysis using U-Mann Whitney with Bonferroni correction DEP and Mediterranean diet vs. Mediterranean diet. $p < 0.01$ [d]Post hoc analysis using U-Mann Whitney with Bonferroni correction DEP and Mediterranean diet vs. Mediterranean diet. $p < 0.01$.
1674 Ekdahl 1990	RCT	18 weeks	67 participants with RA (mean age=53 years; mean disease duration=10.6 years; 64% female)	Interventions were 6 weeks; all programs occurred at primary health center for one hour supervised by a physical therapist. All	The following outcomes are change scores from baseline to 18 weeks (3 months after intervention) between the dynamic groups and the static groups The authors report that there were no significant differences between groups with the same type of program (dynamic 12

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
				participants were encouraged to continue home program for another 3 months. The dynamic and static groups were collapsed in the results. -Dynamic (12 visits): multicomponent with twice per week for 6 weeks; included cycling, body weight exercises for lower extremities, pulley exercises for hips; home program based on above -Dynamic (4 visits): similar exercises as above; 4 visits total -Static (12 visits): joint protection discussion, lower extremity mobility and body weight exercises; home program based on above -Static 4 visits): similar exercises as above; 4 visits total	and 4 visits; static 12 and 4 visits) so the groups were combined. Pain (Pain intensity after muscle tests): • Dynamic groups: 0.0 • Static groups: 0.4 • Between groups P value >0.05 Pain (Pain intensity after bicycle ergometer) • Dynamic groups: -0.4 • Static groups: -0.2 • Between groups P value >0.05 Functional status (MF-index strength): • Dynamic groups: -1.2 • Static groups: -0.1 • Between groups P value <0.01 Functional status (MF-index endurance): • Dynamic groups: -1.8 • Static groups: 0.5 • Between groups P value <0.001 Functional status (MF-index balance/coordination): • Dynamic groups: 1.2 • Static groups: 0.9 • Between groups P value >0.05 Functional status (60-m walk test) • Dynamic groups: -1.9 • Static groups: 0.1 • Between groups P value >0.05 Functional status (Walking up/down 8 steps) • Dynamic groups: -2.7 • Static groups: -1.2 • Between groups P value <0.05 The authors report (giving no specific results of p values) that no significant differences between the dynamic and static

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					groups were found in changes in pain intensity (during previous week), index of joint mobility, Ritchie total index, or indices of ADL, lower extremity ADL, ESR, and CRP from 0-18 weeks. There were significant differences between dynamic and static groups in 0-18 week change scores for Ritchie lower extremity index (p=0.01), Hb (p=0.01), and morning stiffness (p=0.002), indicating that the dynamic group reported fewer painful joints, had higher Hb values, and reported fewer minutes of morning stiffness than the static group.
1022 Seneca 2015	RCT	12 weeks	51 participants with early RA (≤ 5 years) RA -Partly supervised exercises (n=25): (median age=61 years, median disease duration=1 year, 68% female) -Self-administered exercises (n=26): (median age=54 years, median disease duration=1.5 years, 69% female)	Interventions were 12 weeks -Partly supervised (PS) exercises: 6 weeks of supervised training: 30-min bike (15-16 RPE), 30-min muscle strength training (legs, shoulders, trunk extensors/flexors) at 12 rep max; sessions were twice per week in groups of 2-4 patients supervised by same physical therapist; 6 weeks of self-administered exercise (see below)	All results are median change scores (range) from baseline to 12 weeks Pain (NPRS): • PS group (n=15): -2.0 (-6.0 to 3.0) • SA group (n=21): 0.0 (-4.0 to 4.0) • Between groups p value: 0.263 Disease activity (DAS28-CRP) • PS group (n=15): -0.58 (-2.46 to 0.88) • SA group (n=21): 0.06 (-1.62 to 1.77) • Between groups p value: 0.006 Functional status (HAQ-DI) • PS group (n=21): 0.0 (-0.63 to 0.5) • SA group (n=24): 0.0 (-0.63 to 0.3) • Between groups p value: 0.972 Functional status (SF-36 physical component score) • PS group (n=21): 1.3 (-10.3 to 13.6) • SA group (n=24): 0.9 (-5.1 to 20.9)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
				-Self-administered (SA) exercises: 12 weeks of self-administered exercises in local community (participants recommended to exercise at same intensity level as supervised group)	• Between groups p value: 0.802 Functional status (SF-36 mental component score) • PS group (n=21): 2.8 (-7.36 to 17.9) • SA group (n=24): -1.2 (-20.9 to 20.8) • Between groups p value: 0.089

References:

1. Seneca et al. Comparable effect of partly supervised and self-administered exercise programme in early rheumatoid arthritis – a randomised, controlled trial. Dan Med J 2015;62(8):A5127.
2. Bailett et al. A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. Rheumatology 2009;48:410–415.
3. Ekdahl et al. Dynamic versus Static Training in Patients with Rheumatoid Arthritis. Scand J Rheumatology 19: 17-26, 1990.
4. Garcia-Morales et al. Effect of a Dynamic Exercise Program in Combination With Mediterranean Diet on Quality of Life in Women With Rheumatoid Arthritis. J Clin Rheumatol 2020;26: S116–S122.
5. Durcan et al. The Effect of Exercise on Sleep and Fatigue in Rheumatoid Arthritis: A Randomized Controlled Study. J Rheumatol 2014;41:1966–73.
6. Breedland et al. Effects of a Group-Based Exercise and Educational Program on Physical Performance and Disease Self-Management in Rheumatoid Arthritis: A Randomized Controlled Study. Phys Ther.2011;91:879 – 893
7. Lange et al. Effects of Aerobic and Resistance Exercise in Older Adults With Rheumatoid Arthritis: A Randomized Controlled Trial. Arthritis Care & Research Vol. 71, No. 1, January 2019, pp 61–70.
8. Strasser et al. The effects of strength and endurance training in patients with rheumatoid arthritis. Clin Rheumatol (2011) 30:623–632.
9. Rahnama et al. Effects of Strengthening and Aerobic Exercises on Pain Severity and Function in Patients with Knee Rheumatoid Arthritis. International Journal of Preventive Medicine, Vol 3, No 7, July 2012

10. Stavropoulos-Kalinoglou et al. Individualised aerobic and resistance exercise training improves cardiorespiratory fitness and reduces cardiovascular risk in patients with rheumatoid arthritis. Ann Rheum Dis 2013;72:1819–1825.
11. De Jong et al. Is a Long-Term High-Intensity Exercise Program Effective andSafe in Patients With Rheumatoid Arthritis? ARTHRITIS & RHEUMATISM Vol. 48, No. 9, September 2003, pp 2415–2424.
12. De Jong et al. Long term high intensity exercise and damage of small joints in rheumatoid arthritis. Ann Rheum Dis 2004;63:1399–1405.
13. Hansen et al. Longterm Physical Training in Rheumatoid Arthritis. A Randomized Trial with Different Training Programs and Blinded Observers. candinavian Journal of Rheumatology, 22:3, 107-112.
14. Andersson et al. Moderate- to high intensity aerobic and resistance exercise reduces peripheral blood regulatory cell populations in older adults with rheumatoid arthritis. Immunity & Ageing (2020) 17:12.
15. Kucharski et al. Moderate-to-high intensity exercise with person-centered guidance influences fatigue in older adults with rheumatoid arthritis. Rheumatology International (2019) 39:1585–1594.
16. Van den Ende et al. Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomized clinical trial. Ann Rheum Dis 1996;55:798-805.
17. Jahanbin et al. The Effect of Conditioning Exercise on the Health Status and Pain in Patients with Rheumatoid Arthritis: A Randomized Controlled Clinical Trial. JCBNM. 2014;2(3):169-176.
18. Van der Ende et al. Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial. Ann Rheum Dis 2000;59:615–621.
19. Lange et al. Long-time follow up of physical activity level among older adults with rheumatoid arthritis. European Review of Aging and Physical Activity (2020) 17:10.
20. Hurkmans et al. Maintenance of physical activity after Internet-based physical activity interventions in patients with rheumatoid arthritis. Rheumatology 2010;49:167–172.
21. Pineda-Juarez et al. Changes in hand grip strength and bodyweight after a dynamic exercise program and Mediterranean diet in women with rheumatoid arthritis: a randomized clinical trial. Physiotherapy Theory and Practice, 38:4, 504-512.
22. Van Den Berg et al. Using Internet Technology to Deliver a Home-Based Physical Activity Intervention for Patients With Rheumatoid Arthritis: A Randomized Controlled Trial. Arthritis & Rheumatism (Arthritis Care & Research) Vol. 55, No. 6, December 15, 2006, pp 935–945.
23. Hansen et al. Longterm Physical Training in Rheumatoid Arthritis. A Randomized Trial with Different Training Programs and Blinded Observers, Scandinavian Journal of Rheumatology. 1993;22:3, 107-112.

PICO 7: Should patients with RA engage in a mind-body exercise program?

Summary: The literature search and inclusion criteria resulted in our inclusion of 6 RCTs [1,2,3,4,5,7], 1 case-control trial [6], and 2 non-randomized clinical trials [8,9]. They made 3 comparisons:

- Yoga vs usual care (5 RCTs [1,2,3,4,5] and 1 case-control trial [6])
- Yoga vs education (1 RCT [7] and 1 non-RCT [8])
- Tai chi vs education (1 non-RCT [9])

We discuss these comparisons in two sections below (Yoga, Tai Chi), and each received a different certainty of evidence rating (which appears at the beginning of each section).

Yoga

***Yoga**: Overall quality of evidence across critical outcomes: Very Low*

Yoga vs. Usual Care

Six studies were identified that examined Yoga versus Usual Care [1,2,3,4,5,6].

RCTs

<12 weeks

There were 4 RCTs [1,2,3,4] that assessed the effectiveness of Yoga versus Usual Care with interventions lasting less than 12 weeks. Gautam 2020 [1] and Gautam 2021 [2] administered an 8-week intervention of a yoga program comprised of yoga practices by qualified yoga instructors and counseling on stress management, nutrition, and personal lifestyle management; the program was 2 hours per day, 5x per week, 8 weeks, with no home regimen (they were different trials due to their different trial ID numbers). Evans [3] described a 6-week intervention that encompassed yoga group (Iyengar yoga with a variety of poses/postures); classes were 90 minutes and twice per week with a max of 7 participants. The usual care controls were a waitlist control group [3]. Finally, Ward [4] administered a 9-week intervention of Yoga, which was a 75-minute group class 1x/week with qualified instructor for 8 weeks; class and postures were progressive; there was also home practice 3x per week for 20 mins [4].

Disease Activity

For Disease Activity, Gautam 2020 [1] and 2021 [2] both found a significant effect favoring Yoga for DAS28-ESR; however, Evans [3] found no differences for DAS28. For the Clinical Disease Activity Index, Ward found a non-significant effect favoring the control group [4].

Functional Status

For Functional Status, there was several studies measured the Health Assessment Questionnaire. Gautam 2021 [2] and Ward [4] measured mean change scores, and both significantly favored Yoga; however, Ward [4] had a wide confidence interval. Evans [3] also measured HAQ Disability and HAQ Health subscales, and found effects that non-significantly favor Controls, and significantly favors Yoga, respectively.

Evans [3] also measured several other outcomes for Functional Status. For the SF36 subscales, there were no differences for the Bodily Pain subscale whereas the General Health, Vitality (significant), and Mental Health (significant) subscales all favored Yoga [3]. For the Brief Symptom Inventory subscales, the Somatization, Depression (significant), Anxiety, and Global Severity subscales all favored Yoga [3]. Lastly, the Global Improvement Scale significantly favored Yoga [3].

Pain

Only 2 measures of pain were reported, and for both there were very minimal differences between the Yoga and Control Groups: Pain Disability Index [3] and Pain VAS Scores [4].

QOL

Gautam 2020 [1] and Ward [4] both assessed Quality of Life using the Word Health Organization Quality of Life Questionnaire (WHOQOL-BREF) and European Quality of Life Scale (EuroQOL), respectively. For the WHOQOL-BREF, the Physical, Psychological, and Social subscales significantly favored the Yoga Group; the Environmental subscale non-significantly favored the Yoga Group [1]. For the EuroQOL, both the Total Score and VAS subscales found minimal differences between groups [4].

Fatigue

For Fatigue, Evans [3] found that the FACIT Fatigue Score effect significantly favored Yoga. Ward [4] had non-Revmannable Fatigue data for the Bristol Rheumatoid Arthritis Fatigue Numerical Rating Scales (BRAF-NRS): these included Level, Effect, and Coping subscales. For each of these subscales, neither the Yoga Group nor the Control Group had significant changes compared with baseline [4].

Treatment Harms

Only Ward [4] measured Treatment Harms. This study found that for the time periods of during the intervention and during follow-up, both the Yoga Group and Control Group experienced similar frequencies of adverse events [4].

Self-efficacy

Only Evans [3] measured Self-Efficacy, and they used the Arthritis Self-Efficacy Scale. There were minimal differences observed for the Function subscale, while the Pain subscale non-significantly favored the Yoga Group [3].

Sleep

Only Ward [4] measured Sleep. Ward employed the Insomnia Severity Scale, and found that the effect non-significantly favored Yoga [4].

Mental Health

Mental Health was assessed in several ways across 2 different studies. For the Hospital Anxiety and Depression Scale (HADS), the Depression subscale and Anxiety subscale both non-significantly favored Yoga [4]. For the Chronic Pain Acceptance Questionnaire, the effect non-significantly favored Yoga [3]. Lastly, for the 5-Facet Mindfulness Questionnaire (FFMQ), only the NonJudge subscale significantly favored the Yoga Group; the Observe, Describe, Awareness, and NonReact subscales found minimal differences between the Yoga and Control Groups [3].

12 weeks

Ganesan [5] administered a 12-week Yoga intervention that included yoga elements such as asanas, pranayamas, and meditation. Intervention participants were supervised by yoga instructors and the groups were held for 30 minutes, 3 times per week; the Intervention Group also received standard medical treatment [5]. The Control Group received standard medical treatment only. Ganesan assessed 4 outcomes, which were all measures of disease activity [5]. DAS28 AND Interleukin 1alpha significantly favored the Yoga Group, while Interleukin 6 and TNF Alpha both non-significantly favored the Yoga Group [5].

Non-Randomized

2 weeks

Dash conducted a 2-week case-control trial where the Intervention Group received Yoga for 60 minutes per session over the course of 2 weeks (unclear how many sessions were conducted), which included asanas or postures, pranayama or voluntary regulated breathing, meditation, and lectures about yoga philosophy, and joint loosening exercises [6]. The control group received usual care. For each of the 3 outcomes measured (Grip Strength Left Hand, Grip Strength Right Hand, and #NSAIDS/Day), all favored Yoga [6].

Table 1: RCTs: Yoga compared to Usual Care

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		

Critical Outcomes <12 Weeks (6 to 9 weeks)

FUNCTIONAL STATUS: Health Assessment Questionnaire (HAQ) (Mean Change Scores) (8-9 weeks)

2	randomised trials	serious[h]	not serious	not serious	serious[i]	none	48	47	-	MD **0.18 lower** (0.26 lower to 0.1 lower)	⨁⨁◯◯ Low	CRITICAL ***Significant** **Favors Yoga**

FUNCTIONAL STATUS: Health Assessment Questionnaire (Disability) (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **0.2 higher** (0.34 lower to 0.74 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Health Assessment Questionnaire (Health) (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	11	15	-	MD **24.4 lower** (47.59 lower to 1.21 lower)	⨁⨁◯◯ Low	CRITICAL ***Significant** **Favors Yoga**

Function as inferred from FACIT-fatigue (6 weeks)

1	randomised trials	serious[a]	not serious	serious[f]	serious[d]	none	11	15	-	MD **10.5 higher** (3.39 higher to 17.61 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Yoga**

PAIN: SF-36 Bodily pain (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **0.2 lower** (17.73 lower to 17.33 higher)	⨁◯◯◯ Very low	CRITICAL NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
PAIN: Pain Disability Index (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **1.9 lower** (14.15 lower to 10.35 higher)	⨁◯◯◯ Very low	CRITICAL NS
PAIN: Pain VAS (0-100) (9 weeks)												
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **3 lower** (30.2 lower to 24.2 higher)	⨁◯◯◯ Very low	CRITICAL NS

Important Outcomes ≥12 Weeks (12 Weeks)

DISEASE ACTIVITY: DAS 28 (12 weeks)												
1	randomised trials	serious[c]	not serious	not serious	serious[d]	none	68	75	-	MD **0.5 lower** (0.76 lower to 0.24 lower)	⨁⨁◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**
DISEASE ACTIVITY: inferred from Interleukin 1alpha (12 weeks)												
1	randomised trials	serious[c]	not serious	serious[f]	serious[d]	none	68	75	-	MD **3.37 lower** (6.02 lower to 0.72 lower)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Yoga**
DISEASE ACTIVITY: inferred from Interleukin 6 (12 weeks)												
1	randomised trials	serious[c]	not serious	serious[f]	very serious[b]	none	68	75	-	MD **18.93 lower** (43.15 lower to 5.29 higher)	⨁◯◯◯ Very low	IMPORTANT NS
DISEASE ACTIVITY: inferred from TNF-alpha (12 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[c]	not serious	serious[f]	very serious[b]	none	68	75	-	MD **15.55 lower** (33.26 lower to 2.16 higher)	⊕◯◯◯ Very low	IMPORTANT NS

Important Outcomes <12 Weeks (6 to 9 weeks)

DISEASE ACTIVITY: DAS 28 (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **0** (1.08 lower to 1.08 higher)	⊕◯◯◯ Very low	IMPORTANT NS

DISEASE ACTIVITY: DAS28-ESR (Mean Change Scores) (8 weeks)

2	randomised trials	serious[e]	not serious	not serious	not serious	none	68	68	-	MD **0.7 lower** (0.89 lower to 0.52 lower)	⊕⊕⊕◯ Moderate	IMPORTANT ***Significant** **Favors Yoga**

DISEASE ACTIVITY: Clinical Disease Activity Index (Mean Change Score) (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **2.1 higher** (3.64 lower to 7.84 higher)	⊕◯◯◯ Very low	IMPORTANT NS

QOL: Global Improvement Scale (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	11	15	-	MD **1.2 higher** (0.45 higher to 1.95 higher)	⊕⊕◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**

QOL: SF-36 General health (6 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **13.1 higher** (2.11 lower to 28.31 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: SF-36 Vitality (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	11	15	-	MD **26.8 higher** (13.63 higher to 39.97 higher)	⨁⨁◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**

QOL: WHOQOL-BREF Physical Domain (D1) (Mean Change Scores) (8 weeks)

1	randomised trials	serious[i]	not serious	not serious	serious[d]	none	33	33	-	MD **14.5 higher** (10.88 higher to 18.12 higher)	⨁⨁◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**

QOL: WHOQOL-BREF Psychological Domain (D2) (Mean Change Scores) (8 weeks)

1	randomised trials	serious[i]	not serious	not serious	serious[d]	none	33	33	-	MD **15.8 higher** (12.37 higher to 19.23 higher)	⨁⨁◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**

QOL: WHOQOL-BREF Social Domain (D3) (Mean Change Scores) (8 weeks)

1	randomised trials	serious[i]	not serious	not serious	serious[d]	none	33	33	-	MD **7 higher** (4.6 higher to 9.4 higher)	⨁⨁◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**

QOL: WHOQOL-BREF Environmental Domain (D4) (Mean Change Scores) (8 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[i]	not serious	not serious	very serious[b]	none	33	33	-	MD **1.6 higher** (0.23 lower to 3.43 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: EuroQOL EQ-5D-3L (Mean Change Score) (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **0.02 higher** (0.11 lower to 0.15 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: EuroQOL EQ-5D-3L VAS (Mean Change Score) (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **3.9 lower** (17.82 lower to 10.02 higher)	⨁◯◯◯ Very low	IMPORTANT NS

TREATMENT HARMS: Treatment-related adverse events during intervention (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13/13 (100.0%)	9/12 (75.0%)	**OR 9.95** (0.46 to 215.84)	**218 more per 1,000** (from 170 fewer to 248 more)	⨁◯◯◯ Very low	IMPORTANT NS

TREATMENT HARMS: Treatment-related adverse events during follow-up (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	6/13 (46.2%)	7/12 (58.3%)	**OR 0.61** (0.13 to 2.98)	**123 fewer per 1,000** (from 429 fewer to 223 more)	⨁◯◯◯ Very low	IMPORTANT NS

SELF EFFICACY: Arthritis Self-efficacy Scale-function (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **0.19 lower** (1.96 lower to 1.58 higher)	⨁◯◯◯ Very low	IMPORTANT NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
SELF-EFFICACY: Arthritis Self-efficacy Scale-pain (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **1.84 higher** (0.85 lower to 4.53 higher)	⨁◯◯◯ Very low	IMPORTANT NS
SLEEP: Insomnia Severity Scale (Mean Change Score) (9 weeks)												
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **2.5 lower** (5.88 lower to 0.88 higher)	⨁◯◯◯ Very low	IMPORTANT NS
MENTAL HEALTH: Brief Symptom Inventory-Global severity (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **4.2 lower** (8.5 lower to 0.1 higher)	⨁◯◯◯ Very low	IMPORTANT NS
MENTAL HEALTH: Brief Symptom Inventory-Somatization (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **1.3 lower** (3.05 lower to 0.45 higher)	⨁◯◯◯ Very low	IMPORTANT NS
MENTAL HEALTH: Brief Symptom Inventory-Depression (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	11	15	-	MD **2.1 lower** (3.8 lower to 0.4 lower)	⨁⨁◯◯ Low	IMPORTANT ***Significant Favors Yoga**
MENTAL HEALTH: Brief Symptom Inventory-Anxiety (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **0.7 lower** (2.4 lower to 1 higher)	⨁◯◯◯ Very low	IMPORTANT NS
MENTAL HEALTH: SF-36 Mental health (6 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	11	15	-	MD **13.8 higher** (3.12 higher to 24.48 higher)	⨁⨁◯◯ Low	IMPORTANT *Significant Favors Yoga

MENTAL HEALTH: Hospital Anxiety and Depression Scale (HADS) Depression (Mean Change Score) (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **0.6 lower** (2.21 lower to 1.01 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: Hospital Anxiety and Depression Scale (HADS) Anxiety (Mean Change Score) (9 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	13	12	-	MD **2 lower** (4.2 lower to 0.2 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: inferred from Chronic Pain Acceptance Questionnaire-Total (6 weeks)

1	randomised trials	serious[a]	not serious	serious[f]	very serious[b]	none	11	15	-	MD **5 higher** (7.73 lower to 17.73 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: 5-Facet Mindfulness Q'aire (FFMQ)-Observe (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **2.2 higher** (2.65 lower to 7.05 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: 5-Facet Mindfulness Q'aire (FFMQ)-Describe (6 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **1.6 lower** (6.2 lower to 3 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: 5-Facet Mindfulness Q'aire (FFMQ)-Awareness (6 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **0.1 lower** (5.09 lower to 4.89 higher)	⨁◯◯◯ Very low	IMPORTANT NS
MENTAL HEALTH: 5-Facet Mindfulness Q'aire (FFMQ)-Nonjudge (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	11	15	-	MD **4.8 higher** (0.48 higher to 9.12 higher)	⨁⨁◯◯ Low	IMPORTANT ***Significant** **Favors Yoga**
MENTAL HEALTH: 5-Facet Mindfulness Q'aire (FFMQ)-Nonreact (6 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	11	15	-	MD **1 higher** (2.12 lower to 4.12 higher)	⨁◯◯◯ Very low	IMPORTANT NS

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio

Explanations

a. 3051 Revman Bias Table: 3U, 2L, 1H. Issues with randomization and blinding.

b. Single study, and confidence interval spans across the null value.

c. 1865 Revman Bias Table: 4L, 2H. No one was blinded.

d. Single study.

e. 6842 (4L, 1H, 1U) and 1736 (5L, 1H) each have a serious ROB classification.

f. Outcome is a surrogate measure.

g. 6840 Revman Bias Table: 4L, 2H. Participants not blinded, and some selective reporting.

h. Each study has serious bias classification.

i. 1 study has wide confidence interval.

j. 1736 Revman Bias Table: 5L, 1H. Participants not blinded.

Table 2: Additional data on yoga vs usual care

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td>6840
Ward
2018</td><td>Pilot RCT</td><td>8 weeks</td><td>26 adults with RA</td><td>Yoga vs. Usual Care</td><td>
<table>
<tr><th></th><th colspan="2">Yoga</th><th colspan="2">Control</th></tr>
<tr><th>Fatigue Measure</th><th>Median change</th><th>IQR</th><th>Median change</th><th>IQR</th></tr>
<tr><td>BRAF-NRS Level</td><td>-1</td><td>-4, 1</td><td>-1</td><td>-3, 1</td></tr>
<tr><td>BRAF-NRS Effect</td><td>-1</td><td>-4, 1</td><td>-1</td><td>-2, 2</td></tr>
<tr><td>BRAF-NRS Coping</td><td>0</td><td>-1, 3</td><td>-1</td><td>-3, 2</td></tr>
</table>
BRAF-NRS, Bristol Rheumatoid Arthritis Fatigue Numerical Rating Scales
</td></tr>
</table>

Table 3: Nonrandomized study: Yoga compared to Usual Care

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nonrand: Yoga	Usual Care	Relative (95% CI)	Absolute (95% CI)		
Critical Outcomes <12 Weeks (2 weeks)												
FUNCTIONAL STATUS: inferred from Left Grip strength (kg) (2 weeks)												
1	observational studies	very serious[a]	not serious	serious[b]	serious[c]	none	20 cases 20 controls		**RR 12.50** (7.87 to 17.13)	-	⨁◯◯◯ Very low	CRITICAL *Significant Favors Yoga
							-	0.0%		**0 fewer per 1,000** (from 0 fewer to 0 fewer)		
FUNCTIONAL STATUS: inferred from Right grip strength (kg) (2 weeks)												
1	observational studies	very serious[a]	not serious	serious[b]	serious[c]	none	20 cases 20 controls		**RR 12.80** (8.53 to 17.07)	-	⨁◯◯◯ Very low	CRITICAL *Significant Favors Yoga
							-	0.0%		**0 fewer per 1,000** (from 0 fewer to 0 fewer)		
PAIN: inferred from # NSAIDs per day (2 weeks)												
1	observational studies	very serious[a]	not serious	serious[b]	serious[c]	none	20 cases 20 controls		**RR -1.23** (-1.84 to -0.62)	-	⨁◯◯◯ Very low	CRITICAL NS
							-	0.0%		**0 fewer per 1,000** (from 0 fewer to 0 fewer)		

CI: confidence interval; **RR:** risk ratio

Explanations

a. 3115 Revman Bias Table: 4H, 2L. Issues with random sequence generation, allocation concealment, and blinding.

b. Outcome is a surrogate measure.

c. Single study.

Yoga vs. Education

RCTs

12 Weeks

The first study to examine Yoga versus education was Puksic [7], which was a 12-week RCT where the Intervention Group received Yoga, including relaxation, asanas, and breathing exercises; performed twice weekly for 90 minutes per session (6-8 participants per group). The Control Group received education through once weekly 60 minute lecture by rheumatologist on arthritis-related topics. For all 15 outcomes, Yoga was favored [7]. However, only 2 outcomes were significant, SF36 (Physical Function subscale), and FACIT Fatigue Score. The rest were not significant: CRP, DAS-28-CRP, SF36 (Role Physical, Bodily Pain, General Health, Vitality, Social Function, Role Emotional, and Mental Health Subscales), Pain VAS, Hospital Anxiety and Depression Scale (HADS) (Depression and Anxiety subscales), and Perceived Stress Scale. In the non-Revmannable data for Treatment Harms, there were some adverse events in both groups but all of them were determined to be unrelated to the study.

Non-Randomized

8 Weeks

Badsha also examined Yoga versus education using a non-randomized clinical trial design with a waitlist control [8]. The Experimental Group received a yoga program taught in groups of 10 by a licensed practitioner, the Vishwas-Raj yoga program; 1 hour class 2x per week for 6 weeks and exercises for home. The waitlist Control Group received usual care. The data was non-Revmannable due to reporting means without standard deviations or other measures of dispersion [8]. The Yoga group significantly improved their Health Assessment Questionnaire (HAQ), DAS28, Tender Joint Count, and Swollen Joint Count, while the Control group made no improvements for each of these 4 outcomes [8]. For the remaining outcomes (Patient Global Assessment, Fatigue VAS, SF36 [Physical Functioning, Role Limitations due to Physical Functioning, Pain, General Health, Energy/Fatigue, Social Role Limitations due to Emotional Problems, and Mental Health subscales]), neither the Yoga nor the Control group made any improvements [8].

Table 4: RCT: Yoga compared to Education

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Yoga	Education	Relative (95% CI)	Absolute (95% CI)		
Critical Outcomes ≥12 Weeks (24 weeks)												
FUNCTIONAL STATUS: SF-36 physical function (24 weeks)												
1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	23	23	-	MD **13.04 higher** (0.36 higher to 25.72 higher)	⨁⨁◯◯ Low	CRITICAL ***Significant** **Favors Yoga**
FUNCTIONAL STATUS: SF-36 role-physical (24 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **17.4 higher** (7.27 lower to 42.07 higher)	⨁◯◯◯ Very low	CRITICAL NS
FUNCTIONAL STATUS: inferred from FACIT-fatigue (24 weeks)												
1	randomised trials	serious[a]	not serious	serious[b]	very serious[c]	none	23	23	-	MD **5.74 higher** (0.52 higher to 10.96 higher)	⨁◯◯◯ Very low	CRITICAL ***Significant** **Favors Yoga**
PAIN: SF-36 Bodily Pain (24 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **5.43 higher** (5.69 lower to 16.55 higher)	⨁◯◯◯ Very low	CRITICAL NS
PAIN: Pain VAS (24 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **0.82 lower** (2.33 lower to 0.69 higher)	⨁◯◯◯ Very low	CRITICAL NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Yoga	Education	Relative (95% CI)	Absolute (95% CI)		

Important Outcomes ≥12 Weeks (24 weeks)

DISEASE ACTIVITY: inferred from C Reactive Protein (24 weeks)

1	randomised trials	serious[a]	not serious	serious[b]	very serious[c]	none	23	23	-	MD **0.97 lower** (3.21 lower to 1.27 higher)	⨁◯◯◯ Very low	IMPORTANT NS

DISEASE ACTIVITY: DAS 28 CRP (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **0.18 lower** (0.69 lower to 0.33 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: SF-36 General health (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **8.05 higher** (2.84 lower to 18.94 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: SF-36 Vitality (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **5 higher** (4.93 lower to 14.93 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: SF-36 Social function (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **5.43 higher** (5.69 lower to 16.55 higher)	⨁◯◯◯ Very low	IMPORTANT NS

QOL: SF-36 Role emotional (24 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Yoga	Education	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **8.7 higher** (16.8 lower to 34.2 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: SF-36 Mental health (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **2.09 higher** (7.11 lower to 11.29 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: Hospital Anxiety and Depression Scale-depression (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **1 lower** (2.51 lower to 0.51 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: Hospital Anxiety and Depression Scale-anxiety (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **1.51 lower** (3.51 lower to 0.49 higher)	⨁◯◯◯ Very low	IMPORTANT NS

MENTAL HEALTH: Perceived Stress Scale (24 weeks)

1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	23	23	-	MD **1.62 lower** (5.23 lower to 1.99 higher)	⨁◯◯◯ Very low	IMPORTANT NS

CI: confidence interval; **MD:** mean difference

Explanations

a. 2077 Revman Bias Table: 5L, 1H. Participants and personnel not blinded.

b. Outcome is a surrogate measure.

c. Single study, and confidence interval spans across the null value.

d. Single study.

Table 5: Additional data on yoga vs education

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td>2077
Puksic
2021</td><td>RCT</td><td>24 weeks</td><td>57 participants with RA

Yoga group: mean age = 52.9 years, mean disease duration = 7.4 years, 100% female

Control group: mean age = 57.9 years, mean disease duration = 8.7 years, 89% female</td><td>Interventions were 12 weeks

Yoga group (n=30): included relaxation, asanas, and breathing exercises; performed twice weekly for 90 minutes per session (6-8 participants per group)

Control group (n=27): received education through once weekly 60 minute lecture by rheumatologist on arthritis-related topics</td><td>Adverse events:
Yoga group: one participant had persistent positional vertigo. Another participant had acute diverticulitis in the follow-up period that was not considered intervention-related

Control group: one participant presented with acute thyroiditis and anemia before receiving the allocated intervention. Another participant had disease relapse and another experienced forearm fracture. None of these events were considered study-related.</td></tr>
<tr><td>588
Badsha
2009</td><td>Non-randomized Trial</td><td>8 weeks</td><td>47 RA patients</td><td>Yoga vs. Waitlist Control (Control group receives Education)</td><td>
<table>
<tr><th></th><th colspan="2">Yoga</th><th colspan="2">Control</th></tr>
<tr><th>Measure</th><th>Base</th><th>8-wk</th><th>Base</th><th>8-wk</th></tr>
<tr><td>HAQ</td><td>0.8</td><td>0.49**</td><td>0.78</td><td>0.75</td></tr>
<tr><td>DAS28</td><td>3.9</td><td>3.3**</td><td>3.8</td><td>3.9</td></tr>
<tr><td>Tender joint count</td><td>3.5</td><td>2.11**</td><td>5</td><td>5.3</td></tr>
<tr><td>Swollen joint count</td><td>3.2</td><td>1**</td><td>3.9</td><td>3.8</td></tr>
<tr><td>Patient global assessment</td><td>32</td><td>25</td><td>26</td><td>40</td></tr>
<tr><td>Fatigue VAS</td><td>34</td><td>26</td><td>32</td><td>44</td></tr>
<tr><td>SF-36 – Physical Functioning</td><td>65</td><td>66</td><td>63</td><td>65</td></tr>
<tr><td>SF-36 – Role limitation due to PF</td><td>61</td><td>64</td><td>59</td><td>48</td></tr>
<tr><td>SF-36 - Pain</td><td>43</td><td>33</td><td>39</td><td>39</td></tr>
<tr><td>SF-36 – General Health</td><td>52</td><td>53</td><td>51</td><td>53</td></tr>
<tr><td>SF-36 – Energy/ fatigue</td><td>52</td><td>55</td><td>51</td><td>55</td></tr>
<tr><td>SF-36 – Social</td><td>49</td><td>49</td><td>50</td><td>47</td></tr>
<tr><td>SF-36 – Role limitations due to emotional problems</td><td>73</td><td>85</td><td>69</td><td>68</td></tr>
<tr><td>SF-36 - Mental Health</td><td>62</td><td>64</td><td>64</td><td>63</td></tr>
</table>
**Significantly different than baseline
</td></tr>
</table>

Tai chi

***Tai chi**: Overall quality of evidence across critical outcomes: Very Low*

Tai Chi vs. Education

Non-Randomized

3 months (13 weeks)

The only study to examine Tai Chi as an intervention was Shin [9]. In this non-randomized clinical trial, the intervention group received a Tai Chi exercise program, "Twelve movement Tai Chi for arthritis," in a group setting, 1x per week for 60 minutes over 3 months [9]. The control group received information about lifestyle modification including smoking cessation, weight reduction, and exercise advice [9]. The only significant finding was for Tender Joint Count, which favored Tai Chi [9]. Among the non-significant findings, DAS28-ESR, RAPID3, Swollen Joint Count, and Health Assessment Questionnaire (HAQ) favored Tai Chi, while CRP and ESR favored Education (control group) [9].

Table 6: Non-randomized study: Tai Chi compared to Education

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRand: Tai Chi	Education	Relative (95% CI)	Absolute (95% CI)		
Critical Outcomes ≥12 Weeks (3 months)												
FUNCTIONAL STATUS: Health Assessment Questionnaire (HAQ) (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	not serious	very serious[c]	none	29	14	-	MD **0.13 lower** (0.28 lower to 0.02 higher)	⨁◯◯◯ Very low	CRITICAL NS
Important Outcomes ≥12 Weeks (3 months)												
DISEASE ACTIVITY: inferred from CRP (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	serious[b]	very serious[c]	none	29	14	-	MD **0.1 higher** (0.33 lower to 0.53 higher)	⨁◯◯◯ Very low	IMPORTANT NS
DISEASE ACTIVITY: inferred from ESR (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	serious[b]	very serious[c]	none	29	14	-	MD **3.1 higher** (6.13 lower to 12.33 higher)	⨁◯◯◯ Very low	IMPORTANT NS
DISEASE ACTIVITY: DAS28-ESR (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	not serious	very serious[c]	none	29	14	-	MD **0.4 lower** (1.1 lower to 0.3 higher)	⨁◯◯◯ Very low	IMPORTANT NS
DISEASE ACTIVITY: RAPID3 (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	not serious	very serious[c]	none	29	14	-	MD **1.2 lower** (3.86 lower to 1.46 higher)	⨁◯◯◯ Very low	IMPORTANT NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NonRand: Tai Chi	Education	Relative (95% CI)	Absolute (95% CI)		
DISEASE ACTIVITY: inferred from Tender Joint Count (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	serious[b]	very serious[c]	none	29	14	-	MD **2.6 lower** (4.87 lower to 0.33 lower)	⨁◯◯◯ Very low	IMPORTANT ***Significant** **Favors Tai Chi**
DISEASE ACTIVITY: inferred from Swollen Joint Count (Mean Change Score) (3 months)												
1	observational studies	very serious[a]	not serious	serious[b]	very serious[c]	none	29	14	-	MD **0.6 lower** (2.1 lower to 0.9 higher)	⨁◯◯◯ Very low	IMPORTANT NS

CI: confidence interval; **MD:** mean difference

Explanations

a. 579 Revman Bias Table: 5H, 1L. Issues with random sequence generation, allocation concealment, blinding, and attrition.

b. Outcome is a surrogate measure.

c. Single study, and confidence interval spans across the null value.

Summary of the any-exercise analyses

On 7-18-22, For the ACR integrative guideline, the lit review team leader combined the data for PICO 4-5-6-7.

These analyses were restricted to:

- Randomized trials only
- Comparisons to no exercise
- Pain and function outcomes (the only critical outcomes for PICO's 4-5-6-7)
- Direct measurements (no surrogates)
- >=12 weeks (our threshold for long-term)

Because different studies measured outcomes differently, we used the standardized mean difference (SMD). The core team decided *a priori* (on 7-15-22) that the threshold between a small and trivial effect was standardized mean difference (SMD)=0.15. Thus, if the CI for the random-effects meta-analyses was fully above 0.15 (or fully below -0.15), then there would be no downgrade for imprecision (assuming small quantitative heterogeneity as measured by I^2).

During the planning of these analyses, the lit review team leader made decisions about various aspects of these analyses, including which control groups could be considered inactive, which measure of pain to choose when a study reported two or more, which measure of function to choose when a study reported two or more, combining exercise groups when a study included 2+ exercise groups, and reversing effect sizes when some studies used positive scales (higher is better) instead of negative scales (lower is better). A full list of these decisions can be obtained from ACR upon request.

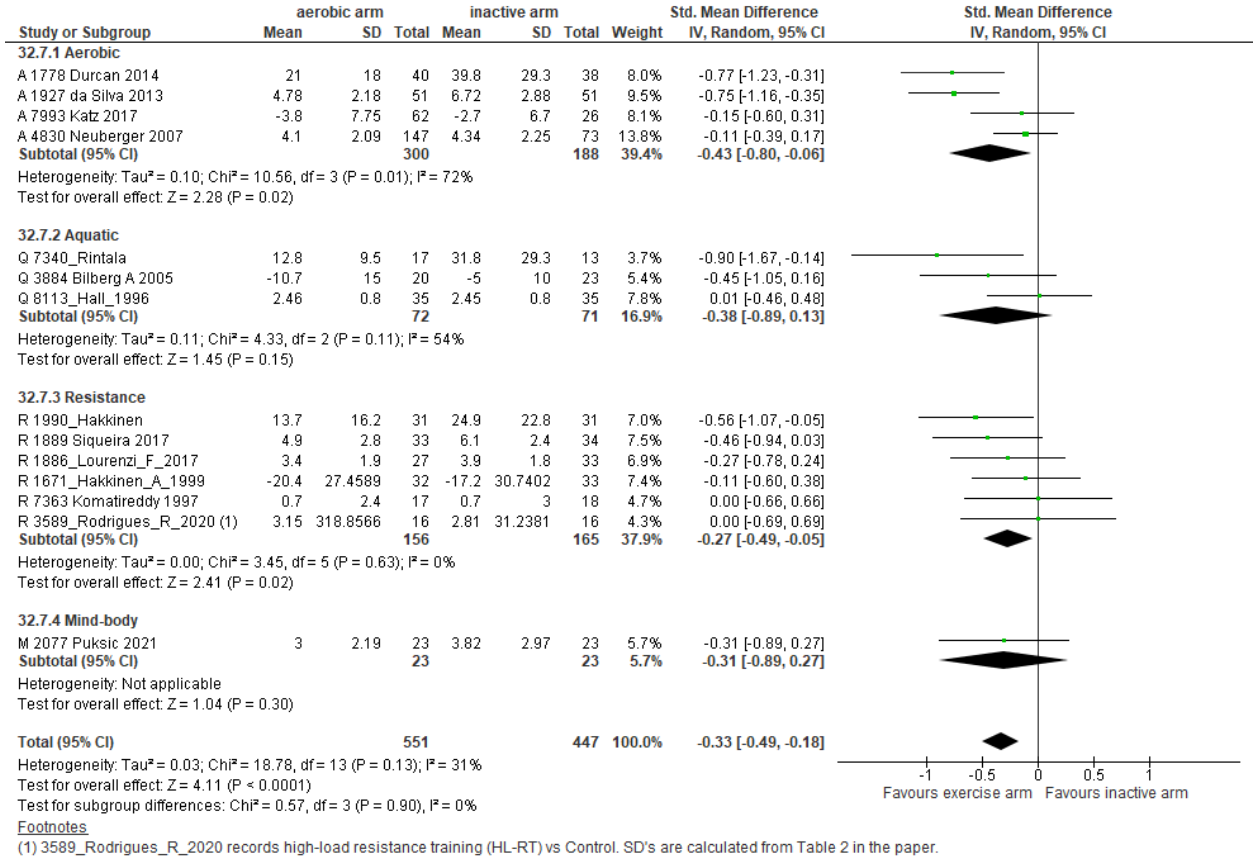
The GRADE for both outcomes was Moderate, which was based on a single downgrade for serious risk of bias (see Table 1 below).

Pain

This outcome included 14 studies with 998 participants (4 aerobic, 3 aquatic, 6 resistance, 1 mind-body). The summary effect size (combining all PICO's) was SMD = -0.33 (95% CI -0.49 to -0.18), indicating a statistically significant benefit of exercise, with I^2=31%, which is small. The GRADE for the pain evidence on any exercise for RA is Moderate (after a single downgrade for risk of bias).

Figure 1 below shows the forest plot (with four PICO subgroups). The four estimates were remarkably similar (ranging from -0.27 for resistance to -0.43 for aerobic), and the test comparing subgroups was p=0.90.

Figure 1. Forest plot of long-term pain after any exercise vs no exercise

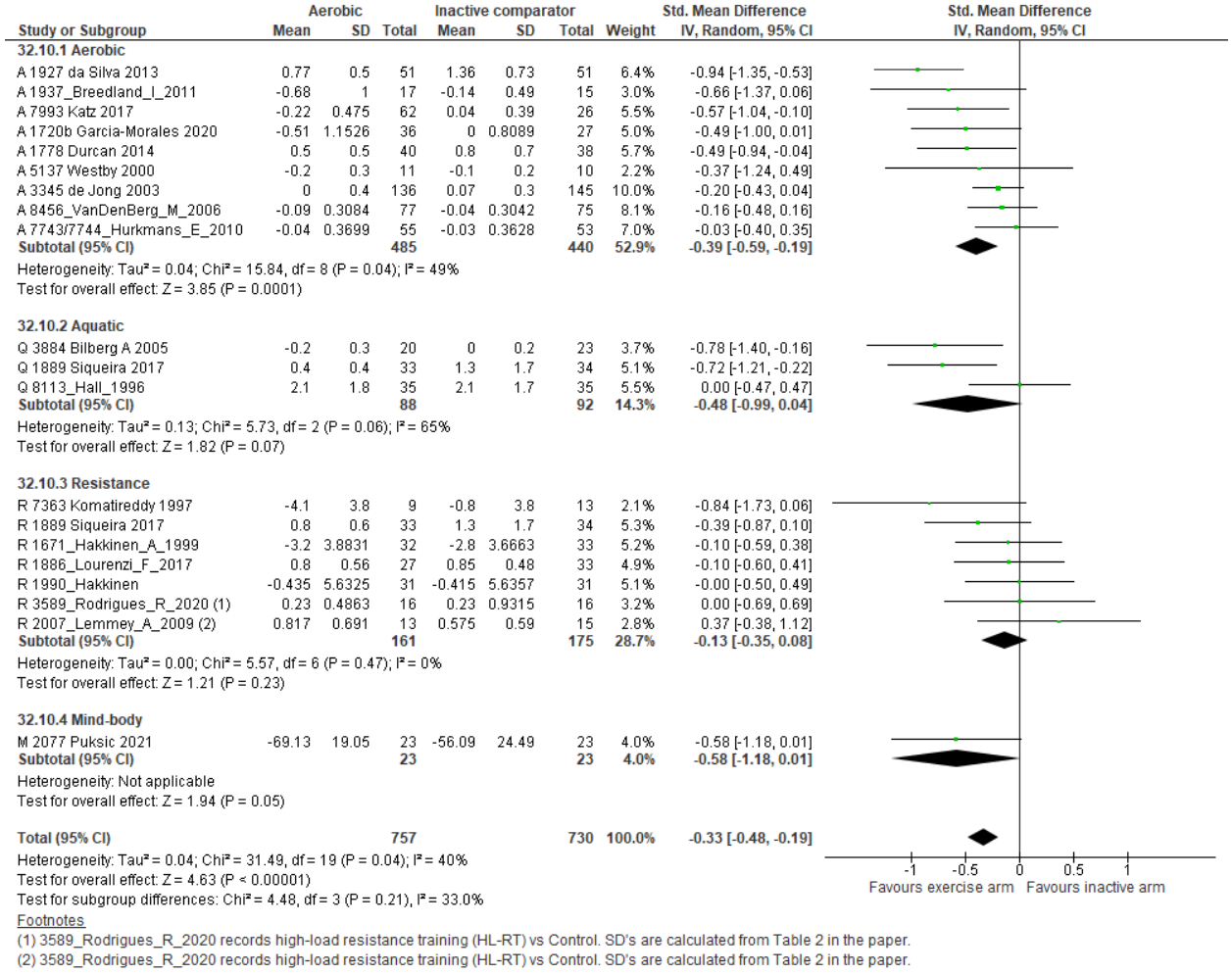


Function

This outcome included 20 studies with 1487 participants (9 aerobic, 3 aquatic, 7 resistance, 1 mind-body). The summary effect size (combining all PICO's) was SMD = -0.33 (95% CI-0.48 to -0.19), indicating a statistically significant benefit of exercise, with I^2=40%, which is small. The GRADE for the pain evidence on any exercise for RA is Moderate (after a single downgrade for risk of bias).

Figure 2 below shows the forest plot (with four PICO subgroups). The four estimates were a little more disparate for function (ranging from -0.13 for resistance to -0.58 for mind-body), and the test comparing subgroups was p=0.21. The aquatic and mind-body estimates were based on far less evidence, so likely the difference between resistance (-0.13) and aerobic (-0.39) is the cause of that somewhat-low p value.

Figure 2. Forest plot of long-term function after any exercise vs no exercise



Footnotes

(1) 3589_Rodrigues_R_2020 records high-load resistance training (HL-RT) vs Control. SD's are calculated from Table 2 in the paper.

(2) 3589_Rodrigues_R_2020 records high-load resistance training (HL-RT) vs Control. SD's are calculated from Table 2 in the paper.

Table 1: Any exercise versus no exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistance	No exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (>= 12 weeks)												
14	randomised trials	serious[e]	not serious	not serious	not serious	none	551	447	-	SMD = - 0.33 (95% CI - 0.49 to - 0.18)	⨁⨁⨁◯ Moderate	Critical Statistically significant in favor of any exercise
Functional Status (>= 12 weeks)												
20	randomised trials	serious[g]	not serious	not serious	not serious	none	757	730	-	SMD = - 0.33 (95% CI - 0.48 to - 0.19)	⨁⨁⨁◯ Moderate	Critical Statistically significant in favor of any exercise

References

1. Gautam S, Kumar M, Kumar U, Dada R. Effect of an 8-Week yoga-based lifestyle intervention on psycho-neuro-immune axis, disease activity, and perceived quality of life in rheumatoid arthritis patients: a randomized controlled trial. Frontiers in Psychology. 2020;11:1-17.
2. Gautam S, Kumar U, Kumar M, Rana D, Dada R. Yoga improves mitochondrial health and reduces severity of autoimmune inflammatory arthritis: A randomized controlled trial. Mitochondrion. 2021;58:147-159.
3. Evans S, Moieni M, Lung K et al. Impact of Iyengar Yoga on Quality of Life in Young Women With Rheumatoid Arthritis. Clin J Pain. 2013;29:988-997.
4. Ward L, Stebbings S, Athens J, Cherkin D, Baxter GD. Yoga for the management of pain and sleep in rheumatoid arthritis: a pilot randomized controlled trial. Musculoskeletal Care. 2018;16:39-47.
5. Ganesan S, Gaur GS, Negi VS, Sharma VK, Pal GK. Effect of yoga therapy on disease activity, inflammatory markers, and heart rate variability in patients with rheumatoid arthritis. The Journal of Alternative and Complementary Medicine. 2020;26(6):501-507.

6. Dash M, Telles S. Improvement in hand grip strength in normal volunteers and rheumatoid arthritis patients following yoga training. Indian J Physiol Pharmacol. 2001;45(3):355-360.
7. Puksic S, Mirovic J, Culo MI, et al. Effects of Yoga in Daily Life program in rheumatoid arthritis: A randomized controlled trial. Complementary Therapies in Medicine. 2021;57:102639.
8. Badsha H, Chhabra V, Leibman C, Mofti A, Kong KO. The benefits of yoga for rheumatoid arthritis: results of a preliminary, structured 8-week program. Rheumatol Int. 2009;29:1417-1421.
9. Shin JH, Lee Y, Kim SG, Choi BY, Lee HS, Bang SY. The beneficial effects of Tai Chi exercise on endothelial function and arterial stiffness in elderly women with rheumatoid arthritis. Arthritis Research and Therapy. 2015;17:380.

PICO 8: Should patients with RA and hand involvement perform resistive hand exercises?

Evidence summary**:** Seven randomized controlled trials and one non-randomized study investigated resistive hand exercises as an intervention for adults with rheumatoid arthritis. The RCTs made three types of comparisons:

- Resistive hand exercises vs. waitlist/usual care (Tables 1 and 2) (1-4)
- Resistive hand exercises vs. non-resistive hand exercises (Tables 3 and 4) (1, 5, 6)
- Resistive hand exercises vs. Education (Tables 5 and 6) (5, 7)

In each comparison, the first table contains the GRADEs for each outcome, and the second table contains additional data provided by the studies of that comparison for which effect sizes could not be reported due to insufficient reporting (e.g., no dispersion).

Heterogeneity in control groups, timepoints, and outcomes precluded meta-analysis and contributed to a low certainty of evidence across outcomes. Other contributors to a low certainty of evidence include serious risk of bias and imprecision on the estimates, likely due to small sample sizes.

Evidence from randomized controlled trials regarding resistive hand exercises to improve pain and functional status (critical outcomes) was mostly inconclusive in the short term (<12 weeks(1-3, 5-7)) and long term (6-12 months(4, 5, 7)). Some studies reported statistically significant positive effects (2, 4-6), while others reported no statistically significant effect (1, 4, 5, 7), or inconsistent effects (3, 5, 7) for multiple measures (e.g., two different questionnaires for functional status). Results varied for other important outcomes. For disease activity, most studies (3 out of 4 (2, 4, 7, 8)) reported no statistically significant or inconsistent effects. Hand resistance exercises may benefit long-term performance-based outcomes(4, 5), but no benefit was observed in the short-term (5, 7). One RCT reported positive effects of hand resistance exercises on work and self-efficacy, and no statistically significant effects on mental health or quality of life.(4)

The non-randomized trial(8) provided very low certainty evidence that resistive hand exercise resulted in greater reduction on an ultrasound measure of disease activity compared to a no-exercise control. See Table 7. We did not consider this non-randomized trial when rating the overall quality of evidence.

Overall Quality of evidence across critical outcomes: Low.

Table 1: Resistive hand exercise vs. Control (usual care/waitlist)(2, 4)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
Pain: Pain troublesomeness score (0-20, higher is better), 12 months												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **0.22 higher** (3.75 lower to 4.19 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference
Pain: MHQ Pain (0-100, lower is better), 12 months												
1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **2.25 lower** (5.98 lower to 1.48 higher)	⨁⨁⨁◯ Moderate	CRITICAL No significant difference
Functional Status: MHQ overall hand function (0-100, higher is better), 12 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **4.37 higher** (0.67 higher to 8.07 higher)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**

Functional status: MHQ ADL (both hands; 0-100, higher is better), 12 months

1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **3.62 higher** (0.43 higher to 6.81 higher)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**

Functional Status: SF 12 Physical Component Score (PCS; 0-100, higher is better), 12 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **1.16 higher** (0.21 lower to 2.53 higher)	⊕⊕⊕◯ Moderate	CRITICAL No significant difference

Function as inferred from Nine hole peg test (seconds; continuous, lower is better), 12 months

1	randomised trials	not serious	not serious	serious[e]	not serious	none	216	222	-	MD **1.24 lower** (2.22 lower to 0.26 lower)	⊕⊕⊕◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**

Pain: VAS (0-100, lower is better), 3 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	not serious	none	50	50	-	MD **37.6 lower** (44.28 lower to 30.92 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**

Functional status: ADL scale (Single question, 0-6, higher is better), 3 weeks

1	randomised trials	serious[a]	not serious	serious[c]	not serious	none	50	50	-	MD **1 higher** (0.52 higher to 1.48 higher)	⨁⨁◯◯ Low	CRITICAL **Statistically significant favoring resistive hand exercise**

Work: MHQ Work (0-100, higher is better), 12 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **5.01 higher** (1.04 higher to 8.98 higher)	⨁⨁⨁◯ Moderate	IMPORTANT **Statistically significant favoring resistive hand exercise**

Mental Health: SF 12 Mental Component Score (MCS; 0-100, higher is better), 12 months

1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **1.78 higher** (0.15 lower to 3.71 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No significant difference

Quality of Life: EQ-5D health state (0-1, higher is better), 12 months

1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **0.01 higher** (0.03 lower to 0.05 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No significant difference

Self-efficacy: Arthritis Self-efficacy Scale (0-100 version suspected; higher is better), 12 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[d]	none	216	222	-	MD **4.08 higher** (0.36 higher to 7.8 higher)	⨁⨁⨁◯ Moderate	IMPORTANT **Statistically significant favoring resistive hand exercise**

Disease activity: Swollen joint count (0-22, lower is better), 12 months

1	randomised trials	not serious	not serious	serious[f]	serious[d]	none	216	222	-	MD **0.11 lower** (0.99 lower to 0.77 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Disease activity: Tender joint count (0-22, lower is better), 12 months

1	randomised trials	not serious	not serious	serious[f]	serious[d]	none	216	222	-	MD **0.19 higher** (0.82 lower to 1.2 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Disease Activity: Hand component of the Ritchie Articular Index (range unclear, lower is better), 3 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Resistive hand exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[b]	not serious	none	50	50	-	MD **4.32 lower** (5.46 lower to 3.18 lower)	⨁⨁◯◯ Low	IMPORTANT **Statistically significant favoring resistive hand exercise**

CI: confidence interval; **MD:** mean difference

a. Unblinded participants, unclear if outcome assessors blinded, allocation concealment not mentioned

b. Outdated measure of disease activity

c. Vague description of the scale

d. Wide confidence interval

e. Surrogate measure of physical function

f. surrogate measure of disease activity

Table 2: Additional Data on Resistive hand exercise vs. Control (usual care/waitlist)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1728 Dellhag 1992(3)	Randomized controlled trial	4 weeks	52 Patients with RA, younger than age 70, with impairment of hand function randomized into 4 groups	**Group 1: Hot wax bath + Hand resistance exercises, 20 minutes, three times a week** **Group 2: Exercises alone**	

	Pain with resisted motion		Pain with non-resisted motion	
Measure	**Base**	**End**	**Base**	**End**
Group 1: Heat+Resistance	**1.4**	**0.8**	**29.3**	**22.1**

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Group 3: Wax only

Group 4: Nothing</td><td>
<table>
<tr><td>Group 2: Exercise only</td><td>1.1</td><td>1.3</td><td>28.8</td><td>17.0**</td></tr>
<tr><td>Group 3: Heat only</td><td>1.5</td><td>1.6</td><td>20.3</td><td>25.9</td></tr>
<tr><td>Group 4: Control</td><td>1.3</td><td>1.5</td><td>27.7</td><td>33.1</td></tr>
</table>
**Significantly better than baseline
No standard deviations reported.
</td></tr>
<tr><td>5135 Hoenig 1993(1)</td><td>Randomized controlled trial</td><td>3 months</td><td>57 participants, mean age 57 years old, 11.3 years since diagnosis, randomized into 4 groups</td><td>Group 1 (n=11): Range of Motion exercises
Group 2 (n=9): Resistance exercises
Group 3 (n=10): ROM + Resistance Exercise
Group 4 (n=11): Control (maintain active lifestyle)</td><td>
<table>
<tr><td></td><td></td><td>Number of painful joints</td><td></td><td>Dexterity (Nine hole peg test)</td><td></td></tr>
<tr><td>Measure</td><td></td><td>Base</td><td>3-mo</td><td>Base</td><td>3-mo</td></tr>
<tr><td rowspan="2">Group 1: ROM</td><td>L</td><td>2.3</td><td>2.7</td><td>23.9</td><td>23.6</td></tr>
<tr><td>R</td><td>2.6</td><td>2.2**</td><td>23.2</td><td>23.3</td></tr>
<tr><td rowspan="2">Group 2: Resistance</td><td>L</td><td>2.0</td><td>3.3</td><td>29.2</td><td>28.0</td></tr>
<tr><td>R</td><td>3.0</td><td>3.4</td><td>32.3</td><td>30.1</td></tr>
<tr><td rowspan="2">Group 3: ROM+ Resistance</td><td>L</td><td>2.5</td><td>2.4</td><td>29.5</td><td>24.4**</td></tr>
<tr><td>R</td><td>3.5</td><td>3.2</td><td>26.4</td><td>28.8</td></tr>
<tr><td rowspan="2">Group 4: Control</td><td>L</td><td>1.6</td><td>2.6</td><td>26.2</td><td>26.5</td></tr>
<tr><td>R</td><td>1.5</td><td>2.7</td><td>24.3</td><td>25.0</td></tr>
</table>
**Significant difference compared to control for change over time p<0.05
No standard deviations reported.
</td></tr>
</table>

Table 3: Resistive hand exercise vs. Other hand exercise (no resistance) (5)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	active hand exercise (no resistance)	Relative (95% CI)	Absolute (95% CI)		
Functional status: AIMS upper limb function (0-10, lower is better), 6 months												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	18	16	-	MD **1.18 lower** (2.08 lower to 0.28 lower)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**
Functional status: AIMS hand and finger function (0-10, lower is better), 6 months												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	18	16	-	MD **0.79 lower** (2.08 lower to 0.5 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
Pain: Brief Pain Inventory (0-10, lower is better), 12 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	active hand exercise (no resistance)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious	not serious	not serious	not serious	none	103	103	-	MD **1.59 lower** (2.17 lower to 1.02 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant favoring resistive hand exercise**

Functional status: Hand function (AIMS, SF-SACRAH; scaled to AIMS 0-10, lower is better)), 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	active hand exercise (no resistance)	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious	not serious	not serious	not serious	none	121	120	-	SMD **0.68 lower** (0.94 lower to 0.42 lower) *On the scale of AIMS-2 hand/finger function (0-10 scale where lower scores are better),* ***MD 1.42 lower (1.97 lower to 0.88 lower)***	⨁⨁◯◯ Low	CRITICAL **Statistically significant favoring resistive hand exercise**

Functional status: AIMS upper limb function (0-10 lower is better), 12 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	active hand exercise (no resistance)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	not serious	none	18	17	-	MD **1.19 lower** (2.2 lower to 0.18 lower)	⊕⊕⊕◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**

Disease activity: Patient global assessment (suspected 0-10, lower is better), 6 months

1	randomised trials	serious[a]	not serious	serious[c]	very serious[b]	none	21	24	-	MD **0.84 higher** (0.6 lower to 2.28 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Disease activity: Swollen joint count (# of joints unclear, lower is better), 6 months

1	randomised trials	serious[a]	not serious	serious[c]	very serious[b]	none	21	24	-	MD **0.89 higher** (0.77 lower to 2.55 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	active hand exercise (no resistance)	Relative (95% CI)	Absolute (95% CI)		
Disease activity: Tender joint count (# of joints unclear, lower is better, 6 months												
1	randomised trials	serious[a]	not serious	serious[c]	very serious[b]	none	21	24	-	MD **1.34 higher** (0.44 lower to 3.12 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

a. Unblinded participants, moderate attrition

b. wide confidence intervals

c. surrogate measure of the outcome

Table 4: Additional data on Resistive hand exercise vs. Other hand exercise (no resistance)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
5135 Hoenig 1993(1)	Randomized controlled trial	3 months	57 participants, mean age 57 years old, 11.3 years since diagnosis, randomized into 4 groups	**Group 1 (n=11): Range of Motion exercises** **Group 2 (n=9): Resistance exercises** **Group 3 (n=10): ROM + Resistance Exercise**	

		Number of painful joints		Dexterity (Nine hole peg test)	
Measure		**Base**	**3-mo**	**Base**	**3-mo**
Group 1: ROM	L	2.3	2.7	23.9	23.6
	R	2.6	2.2**	23.2	23.3
Group 2: Resistance	L	2.0	3.3	29.2	28.0
	R	3.0	3.4	32.3	30.1
	L	2.5	2.4	29.5	24.4**

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Group 4 (n=11): Control (maintain active lifestyle)</td><td>
<table>
<tr><td>Group 3: ROM+ Resistance</td><td>R</td><td>3.5</td><td>3.2</td><td>26.4</td><td>28.8</td></tr>
<tr><td rowspan="2">Group 4: Control</td><td>L</td><td>1.6</td><td>2.6</td><td>26.2</td><td>26.5</td></tr>
<tr><td>R</td><td>1.5</td><td>2.7</td><td>24.3</td><td>25.0</td></tr>
</table>
**Significant difference compared to control for change over time p<0.05
No standard deviations are reported.</td></tr>
<tr><td>1155 Obrien 2006(5)</td><td>RCT</td><td>6 months</td><td>67 participants with RA
-Hand strengthening (mean age=62.3 years, mean disease duration=17.7 years, 71% female)

-Hand stretching (mean age=57.3 years, mean disease duration=13.2 years, 63% female)

-Education control (mean age=59.5 years, mean disease duration=9.7 years, 73% female)</td><td>Interventions were performed at home for 6 months.
-Hand strengthening and mobilization home exercise (n=21)
-Hand stretching (Active control) (n=24)
-Education Control (n=22)</td><td>Performance-based functional status (Jebsen-Taylor hand function, Lower scores indicate quicker time in seconds)
<u>Change from 0-12 weeks</u> [median change scores (IQR)]
• Hand strengthening group (n=18): -7.62 (15.97)
• Hand stretching group (n=17): -5.47 (13.16)
• Joint protection information group (n=19): -4.75 (11.82)
<u>Change from 0-6 months [median change scores (IQR)]</u>
• Hand strengthening group (n=18): -7.92 (16.56)
• Hand stretching group (n=16): -3.38 (15.26)
• Joint protection information group (n=18): -3.46 (13.73)</td></tr>
</table>

Table 5. Resistive hand exercise compared to Education control (5, 7)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	Education control	Relative (95% CI)	Absolute (95% CI)		
Functional status: AIMS upper limb function (0-10, lower is better), 6 months												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	18	18	-	MD **1.3 lower** (2.05 lower to 0.55 lower)	⊕⊕⊕◯ Moderate	CRITICAL **Statistically significant favoring resistive hand exercise**
Functional status: AIMS hand and finger function (0-10, lower is better), 6 months												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	18	18	-	MD **0.59 lower** (1.7 lower to 0.52 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Functional status: AIMS upper limb function (0-10, lower is better), 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	18	19	-	MD **0.69 lower** (1.6 lower to 0.22 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	Education control	Relative (95% CI)	Absolute (95% CI)		
Functional status: AIMS hand and finger function (0-10, lower is better), 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	18	19	-	MD **0.13 lower** (1.18 lower to 0.92 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Pain: VAS Right (0-100, lower is better), 8 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **1.67 lower** (8.65 lower to 5.31 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Pain: VAS Left (0-100, lower is better), 8 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **4.48 lower** (16.93 lower to 7.97 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Functional status, HAQ (0-3, lower is better), 8 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	Education control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **0.1 lower** (0.22 lower to 0.02 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Functional Status AMPS ADL process ability (range unclear, measure expressed in logits, higher is better), 8 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **0.03 higher** (0.27 lower to 0.33 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Functional Status AMPS ADL motor ability (range unclear, measure expressed in logits, higher is better), 8 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **0.04 higher** (1.53 lower to 1.61 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Functional status: ADL-Questionnaire (ADL-Q, 0-100%, lower is better), 8 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	Education control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **0** (1.97 lower to 1.97 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Disease activity: Swollen joint count (0-28, lower is better), 6 months

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	21	22	-	MD **0.94 lower** (3.72 lower to 1.84 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Disease activity: Tender joint count (0-28, lower is better), 6 months

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	21	22	-	MD **0** (2.45 lower to 2.45 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Disease activity: Patient perception of global assessment of disease activity (0-10, lower is better), 6 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	Education control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	21	22	-	MD **0.06 lower** (1.59 lower to 1.47 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference

Disease activity: DAS28 (0-9.4, lower is better), 8 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	27	27	-	MD **0.43 lower** (0.78 lower to 0.08 lower)	⨁⨁◯◯ Low	IMPORTANT **Statistically significant favoring resistive hand exercise**

Disease activity: Swollen joint count (0-28, lower is better), 8 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCT: Resistive hand exercise	Education control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	27	27	-	MD **14.77 lower** (29.4 lower to 0.14 lower)	⊕◯◯◯ Very low	IMPORTANT **Statistically significant favoring resistive hand exercise**

Disease activity: Tender joint count (0-28, lower is better), 8 weeks

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	27	27	-	MD **0.89 lower** (2.66 lower to 0.88 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Disease activity: VAS (0-100, lower is better), 8 weeks

1	randomised trials	serious[a]	not serious	serious[c]	serious[b]	none	27	27	-	MD **7.27 lower** (18.65 lower to 4.11 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

a. unblinded or unclear blinding of participants and assessors
b. Wide confidence interval

c. surrogate measure of the outcome

Table 6. Additional data on Resistive hand exercise compared to Education control

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1155 Obrien 2006 (5)	RCT	6 months	67 participants with RA -Hand strengthening (mean age=62.3 years, mean disease duration=17.7 years, 71% female) -Hand stretching (mean age=57.3 years, mean disease duration=13.2 years, 63% female) -Education control (mean age=59.5 years, mean disease duration=9.7 years, 73% female)	Interventions were performed at home for 6 months. **-Hand strengthening and mobilization home exercise (n=21)** -Hand stretching (Active control) (n=24) **-Education Control (n=22)**	Performance-based functional status (Jebsen-Taylor hand function, Lower scores indicate quicker time in seconds) Change from 0-12 weeks [median change scores (IQR)] • **Hand strengthening group (n=18): -7.62 (15.97)** • Hand stretching group (n=17): -5.47 (13.16) • **Joint protection information group (n=19): -4.75 (11.82)** Change from 0-6 months [median change scores (IQR)] • **Hand strengthening group (n=18): -7.92 (16.56)** • Hand stretching group (n=16): -3.38 (15.26) • **Joint protection information group (n=18): -3.46 (13.73)**

Table 7. Non-randomized study: Resistive hand exercises vs. no exercises(8)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Non-randomized: Resistive hand exercises	no exercises	Relative (95% CI)	Absolute (95% CI)		

Disease Activity as inferred from Color Fraction (indicates blood flow in synovial tissue, range 0-1, lower is better), 8 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Non-randomized: Resistive hand exercises	no exercises	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious[a]	not serious	very serious[b]	serious[c]	none	18	18	-	MD **0.02 lower** (0.07 lower to 0.03 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

a. non-randomized

b. surrogate measure of disease activity

c. wide confidence interval

REFERENCES:

1. Hoenig H, Groff G, Pratt K, Goldberg E, Franck W. A randomized controlled trial of home exercise on the rheumatoid hand. J Rheumatol. 1993;20(5):785-9.

2. Buljina AI, Taljanovic MS, Avdic DM, Hunter TB. Physical and exercise therapy for treatment of the rheumatoid hand. Arthritis Rheum. 2001;45(4):392-7.

3. Dellhag B, Wollersjö I, Bjelle A. Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. Arthritis Care Res. 1992;5(2):87-92.

4. Lamb SE, Williamson EM, Heine PJ, Adams J, Dosanjh S, Dritsaki M, et al. Exercises to improve function of the rheumatoid hand (SARAH): a randomised controlled trial. Lancet. 2015;385(9966):421-9.

5. O'Brien AV, Jones P, Mullis R, Mulherin D, Dziedzic K. Conservative hand therapy treatments in rheumatoid arthritis--a randomized controlled trial. Rheumatology (Oxford). 2006;45(5):577-83.

6. Dimple Choudhry JY, Harpreet Singh, Savarna, Poonam Dhankher. Effects of Hand Strengthening Exercise on Various Hand Functions in Female Patients of Chronic Rheumatoid Arthritis: A Case-control Study. Journal of Clinical and Diagnostic Research. 2021;15(4):YC01-YC4.

7. Ellegaard K, von Bülow C, Røpke A, Bartholdy C, Hansen IS, Rifbjerg-Madsen S, et al. Hand exercise for women with rheumatoid arthritis and decreased hand function: an exploratory randomized controlled trial. Arthritis Research & Therapy. 2019;21(1):158.

8. Ellegaard K, Torp-Pedersen S, Lund H, Pedersen K, Henriksen M, Danneskiold-Samsøe B, et al. The effect of isometric exercise of the hand on the synovial blood flow in patients with rheumatoid arthritis measured by color Doppler ultrasound. Rheumatol Int. 2013;33(1):65-70.

Bracing/splinting/orthoses

PICO 9: Should patients with RA and hand/wrist impairment/deformity use splinting/orthoses/compression?

Summary: Literature searches identified five randomized controlled trial (RCT) [1-5] addressing this PICO question; they made 3 comparisons:

- 3 studies compared wrist/hand splints/orthoses to no splints/orthoses[1-3]
- 1 study compared compression gloves to placebo gloves [4]. We separated this from the first comparison because compression employs different mechanisms of action than splints/orthoses.
- 1 compared a splint/orthosis for the thumb to no thumb splint/orthosis [5]. We separated this from the first comparison because thumb-specific splint/orthoses are categorically unique.

The critical outcomes for this PICO question were pain and function.

For the first comparison, both Adams et al. [1] and Silva et al. [2] assessed static resting hand splints which immobilized the wrist, fingers and thumb. Participants wore the orthoses/splints either at night or during rest periods [1] or at night while sleeping [2]. Adams et al. [1] found no significant group differences in pain or function (MHQ) at 12 months. However, Silva et al. [2] reported significant group differences favoring orthoses/splints for pain and function (HAQ and DASH) at 90 days. The Veehof et al. [3] study compared static wrist splints/orthoses (which immobilized the wrist but allowed motion of the fingers and thumb) to no splints/orthoses. The splints/orthoses were to be worn during the day during activity for 4 weeks; controls did not receive the splint. There were significance differences for changes in pain scores favoring the splint/orthoses group but no group differences for function.

In Hammond et al.'s RCT [4], adults with RA wore compression gloves (23-32 mmHg pressure) or loose-fitting placebo gloves with little to no pressure (control), and they were instructed to wear the gloves during the day or night only but not to wear the glove 24 hours/day. At 12 weeks, day and night hand pain was slightly reduced but there was not a significant difference between the groups. Small improvements in function as measured by the MHQ were observed but there were no statistically significant or clinical relevant differences between the groups. Adverse events were reported in 51% of the intervention group and 36% of the control groups. The most common events reported were that the gloves made the hands feel hot and itchy, and feelings of pins and needles and numbness in the fingertips.

The RCT by Silva et al [5] compared orthoses/splints for the thumb vs no splints/orthoses, in people with RA with type I and type II boutonniere deformities. The intervention group wore the thumb orthosis, which stabilized the metacarpophalangeal joint and prevented hyperextension of the interphalangeal joint, at home. The same type of orthosis was fabricated for the control group who were allowed to wear it only during the evaluation sessions at evaluation (to equate the position of the metacarpophalangeal and interphalangeal joints of the thumb for both groups). At 90 days, there was a significant reduction in pain in the orthosis group compared to the controls but no significant between-group differences for function (as measured by the HAQ).

All 5 studies were RCTs [1-5] were rated of moderate quality (thus we invariably downgraded for serious risk of bias). Studies followed strict protocols. Except for the Hammond et al. [1] and Adams et al. [2] study, sample sizes were relatively small (<100) in three studies; only two studies used the same outcome; and follow up time periods ranged from 1 month (4 weeks) to 12 months. In general, it was impossible to blind participants as to splint/orthosis intervention; however, blinding of compression gloves is theoretically possible due to unobservable pressure differences.

Quality of evidence across all critical outcomes: Very Low

Table 1. Hand splint/orthosis compared to No splint/orthosis

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hand splint/orthosis	No splint/orthosis	Relative (95% CI)	Absolute (95% CI)		
VAS pain, 4 wks (range 0 no to 10 severe)												
2	randomized trials	serious[a]	serious	serious	serious[b]	none	42	41	-	SMD **0.52 SD lower** (0.97 lower to 0.07 lower) This corresponds to MD 1.95 (3.6 lower to 0.65 lower) on a 0-10 scale	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring orthosis**
VAS pain, 12 wks (range 0 no to 10 severe)												
1	randomized trials	serious[a]	not serious	serious[b]	serious[b]	none	25	25	-	MD **2.39 lower** (3.77 lower to 1.01 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring orthosis**
Health Assessment Questionnaire (HAQ, 6 wks (range 0 low to 3 high disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25	25	-	MD **0.17 lower** (0.46 lower to 0.12 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
HAQ, 12 wks (range 0 low to 3 high disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25	25	-	MD **0.55 lower** (0.82 lower to 0.28 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring orthosis**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hand splint/orthosis	No splint/orthosis	Relative (95% CI)	Absolute (95% CI)		
Disabilities of the Arm, Shoulder, Hand (DASH) Q2, 6 wks (range 0 low – 100 more disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25	25	-	MD **16.93 lower** (28.77 lower to 5.09 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring orthosis**
DASH Q2, 12 wks (range 0 low to 100 more disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25	25	-	MD **34.72 lower** (44.34 lower to 25.1 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring orthosis**
DASH Q3, 6 wks (range 0 low to 100 more disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25	25	-	MD **6.11 lower** (15.43 lower to 3.21 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
DASH Q3, 12 wks (range 0 low to 100 more disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25	25	-	MD **21.07 lower** (30.15 lower to 11.99 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring orthosis**
Sequential Occupational Dexterity Assessment-S pain (SODA-S pain), 4 wks (range 0 no to 6 activities cause pain)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	17	16	-	MD **0.8 lower** (1.91 lower to 0.31 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Hand splint/orthosis	No splint/orthosis	Relative (95% CI)	Absolute (95% CI)		
SODA score, 4 wks (0 cannot do, very difficult to 48 performs as requested, not difficult)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	17	16	-	MD **1.6 higher** (1.62 lower to 4.82 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
DASH, 4 wks (range 0 low to 100 more disability)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	17	16	-	MD **2.6 lower** (10.48 lower to 5.28 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
Michigan Hand Questionnaire (MHQ), 52 wks (0 very good, not difficult to 100 very poorly, very difficult)												
1	randomized trials	serious[a]	not serious	not serious	serious[b]	none	56	60	-	MD **3 lower** (10.21 lower to 4.21 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

CI: confidence interval; **MD:** mean difference; **SMD:** standardized mean difference

Explanations

a. moderate quality

b. small sample size

Table 2: Isotoner High Compression Gloves compared to Control placebo gloves

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Isotoner High Compression Gloves	Control placebo gloves	Relative (95% CI)	Absolute (95% CI)		
Pain, Dominant Hand Numerical Rating Scale (NRS), 12 wks (range 0 no to 10 severe pain)												
1	randomized trials	not serious[a]	not serious	not serious	serious	none	84	79	-	MD **0.2 higher** (0.51 lower to 0.91 higher)	⨁⨁⨁◯ Moderate	CRITICAL No statistically significant difference
Nighttime pain, dominant hand, 12 wks (range 0 no to 10 severe pain)												
1	randomized trials	serious	not serious	not serious	serious	none	84	79	-	MD **0.2 higher** (0.58 lower to 0.98 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
Measure of Activity Performance –H (MAPHAND), 12 wks (range 1 no difficulty to 4 unable to perform)												
1	randomized trials	serious[a]	not serious	not serious	serious	none	84	79	-	MD **0** (0.18 lower to 0.18 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
Health Assessment Questionnaire (HAQ), 12 wks (range 0 low to 3 high disability)												
1	randomized trials	serious[a]	not serious	not serious	serious	none	84	79	-	MD **0.1 higher** (0.13 lower to 0.33 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
Michigan Hand Questionnaire (MHQ), 12 wks (0 very good, not difficult to 100 very poorly, very difficult)												
1	randomized trials	serious[a]	not serious	not serious	serious	none	84	79	-	MD **0.2 lower** (5.44 lower to 5.04 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Small sample size

Table 3 Thumb Splint/orthosis compared to No splint /orthosis

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Thumb Splint	no splint	Relative (95% CI)	Absolute (95% CI)		
VAS pain, 6 wks (range 0 l no to 10 severe pain)												
1	randomized trials	serious	not serious	not serious	serious[a]	none	20	20	-	MD **1.65 lower** (3.03 lower to 0.27 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring orthosis**
VAS pain, 12 wks (range 0 no to 10 severe)												
1	randomized trials	serious	not serious	not serious	serious[a]	none	20	20	-	MD **2.25 lower** (3.83 lower to 0.67 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring orthosis**
Health Assessment Questionnaire (HAQ), 6 wks (range 0 low to 3 high disability)												
1	randomized trials	serious	not serious	not serious	serious[a]	none	20	20	-	MD **0.43 lower** (0.81 lower to 0.05 lower)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
HAQ, 12 wks (range 0 low to 3 high disability)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Thumb Splint	no splint	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious	not serious	not serious	serious[a]	none	20	20	-	MD **0.01 higher** (0.44 lower to 0.46 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

CI: confidence interval; **MD:** mean difference;

Explanations

a. small sample size

References

1. Adams J, Burridge J, Mullee M, Hammond A, Cooper C. The clinical effectiveness of static resting splints in early rheumatoid arthritis: a randomized controlled trial. Rheumatology 2008;47:1548-53.

2. Silva AC, Jones A, Silva PG, Natour J. Effectiveness of a night-time hand positioning splint in rheumatoid arthritis: a randomized controlled trial. J Rehabil Med 2008;40:749-54.

3. Veehof MM, Taal E, Heijnsdijk-Rouwenhorst VD, van de Laar MA. Efficacy of wrist working splints in patients with rheumatoid arthritis: a randomized controlled study. Arthritis Car Res 2008; 59:1698-704.

4. Hammond A, Prior Y, Cotterill S, Sutton C, Camacho E, Heal C, et al. Clinical and cost effectiveness of arthritis gloves in rheumatoid arthritis (A-GLOVES): randomised controlled trial with economic analysis. BMC Musculoskelet Disord 2021 Dec;22(1):1-3.

5. Silva PG, Lombardi Jr I, Breitschwerdt C, Poli Araújo PM, Natour J. Functional thumb orthosis for type I and II boutonniere deformity on the dominant hand in patients with rheumatoid arthritis: a randomized controlled study. Clin Rehabil 2008;22:684-9.

PICO 10: Should patients with RA and foot/ankle involvement use bracing/orthoses/taping?

Summary: The literature search and inclusion criteria resulted in our inclusion of 11 RCTs [1,2,3,4,5,6,7,8,9,10,11]. They made 2 comparisons:

- Orthotics vs No Orthotics (5 RCTs [1,2,3,4,5])
    - Within this comparison, Chalmers [4] studied Semi-Rigid versus No Orthotics, and Soft versus No Orthotics
- Orthotics vs Placebo (6 RCTs [6,7,8,9,10,11])

We discuss these comparisons in the 2 sections below, and each received a different certainty of evidence rating (which appears at the end of each section).

Orthotics vs. No Orthotics

Orthotics vs No Orthotics*: Overall quality of evidence across critical outcomes: Very Low*

5 studies compared Orthotics versus No Orthotics [1,2,3,4,5]. 4 studies contained RevMan data only [1,2,3,4], while 1 study contained both RevMan and Non-RevMannable data [5]. In this comparison, the studies measured the effects of orthotics versus no orthotics (meaning that there was NO PLACEBO used).

< 12 Weeks

Gaino [1], Mejjad [2], and Fransen [3] all conducted RCTs of <12 weeks. Across the 3 studies, there was minimal overlap of outcomes, with the only outcome overlap being walking pain. The walking pain outcome was found to favor the intervention for both Mejjad [1] and Franssen [3]. Gaino [1] found that at 4 weeks followup, the intervention group was favored for each of the following outcomes: Foot Function Index (Pain Subscale, Activity Limitation Subscale, Disability Subscale, and Total Score), Timed-Up-and-Go Test, and the Berg Balance Scale. Mejjad [2] found that at 1-month followup, 100% of the sample had lower pain levels while walking with orthotics than without, while walking speed non-significantly favored the orthotics group. Lastly, Franssen [3] found that at 2 months followup, improvements were found in the intervention group for all outcomes measured: Health Assessment Questionnaire (HAQ), gait speed (normal and fast), pain-free walk time, non-weight bearing pain, stair pain, fatigue, and well-being.

≥12 Weeks

12 weeks

Chalmers [4] conducted an RCT of 12 weeks. They measured Semi-Rigid Orthotics versus No Orthotics, and Soft Orthotics versus No Orthotics. The only significant outcomes were Foot Pain (0-10), Toronto Activities of Daily Living Measure - Sub Walk subscale, which both favored Semi-Rigid Orthotics; for Soft orthotics, both of these outcomes were not significant. For both Semi-Rigid Orthotics and Soft orthotics, all

other outcomes measured were not significant: Robinson Bashall (Walking, Stairs, and Stand subscales), Toronto Activities of Daily Living Measure (Walking, and Stairs subscales), 50' walking, lower extremity synovitis joint count, and metatarsal phalangeal synovitis joint count.

30 months (130 weeks)

Woodburn [5] conducted a 30-month RCT, where the intervention group received custom foot orthotics with podiatry supervision. The control group received no orthotics assigned at baseline, but they were used if they were prescribed as part of usual treatment later on in the study; this is a limitation of the study because it may weaken the results if some of the control group did end up using orthotics. For the RevMan data, the intervention group improved Foot Function Index Scores for the Pain Subscale, Disability Subscale, and Total Score. There were minimal to no differences for DAS28 score, Global Pain, and Foot Function index – Activity Limitation Subscale. For the non-RevMannable data, there were no significant differences between the intervention and control group for Health Assessment Questionnaire (HAQ) Score, Larsen Index (Hands) Score, and Larsen Index (Feet) Score.

Table 1: RCTs: Orthotics compared to No Orthotics

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
Critical Outcomes ≥12 weeks (12 weeks to 30 months)												
FUNCTIONAL STATUS: Foot Function Index (Total) (Mean Change Scores - area under the curve) (30 months)												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	50	48	-	MD 218.5 lower (408.26 lower to 28.74 lower)	⨁⨁◯◯ Low	CRITICAL *Significant Favors Orthotics
FUNCTIONAL STATUS: Foot Function Index (Disability Subscale) (Mean Change Scores - area under the curve) (30 months)												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	50	48	-	MD 309.1 lower (557.05 lower to 61.15 lower)	⨁⨁◯◯ Low	CRITICAL *Significant Favors Orthotics

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
FUNCTIONAL STATUS: Foot Function Index (activity limitation Subscale) (Mean Change Scores - area under the curve) (30 months)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	50	48	-	MD 81.4 lower (249.13 lower to 86.33 higher)	⊕◯◯◯ Very low	CRITICAL NS
PAIN: Foot Function Index (Pain Subscale) (Mean Change Scores - area under the curve) (30 months)												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	50	48	-	MD 307.8 lower (548.23 lower to 67.37 lower)	⊕⊕◯◯ Low	CRITICAL *Significant Favors Orthotics
PAIN: Global Pain (0-100 VAS) (Mean Change Scores - area under the curve) (30 months)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	50	48	-	MD 77.3 lower (354.97 lower to 200.37 higher)	⊕◯◯◯ Very low	CRITICAL NS
PAIN: Foot Pain (0-10) (Rigid Orthotics) (Mean Change Scores) 12 weeks												
1	randomised trials	serious[g]	not serious	not serious	serious[c]	none	24	24	-	MD 1.92 lower (3.34 lower to 0.5 lower)	⊕⊕◯◯ Low	CRITICAL *Significant Favors Orthotics
PAIN: Foot Pain (0-10) (Soft Orthotics) (Mean Change Scores) 12 weeks												
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0.06 lower (1.55 lower to 1.43 higher)	⊕◯◯◯ Very low	CRITICAL NS
FUNCTIONAL STATUS: Robinson Bashall Walking (Rigid Orthotics) (12 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0.5 higher (11.19 lower to 12.19 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Robinson Bashall Walking (Soft Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 2.5 higher (8.67 lower to 13.67 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Robinson Bashall Stairs (Rigid Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 1.5 higher (5.07 lower to 8.07 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Robinson Bashall Stairs (Soft Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0.7 higher (5.78 lower to 7.18 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Robinson Bashall Stand (Rigid Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 9.3 lower (62.35 lower to 43.75 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Robinson Bashall Stand (Soft Orthotics) (12 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 13.6 higher (49.24 lower to 76.44 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Toronto Activities of Daily Living Measure - Walking (Rigid Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0.1 higher (0.53 lower to 0.73 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Toronto Activities of Daily Living Measure - Walking (Soft Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0.1 higher (0.54 lower to 0.74 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Toronto Activities of Daily Living Measure - Stairs (Rigid Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0 (0.12 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Toronto Activities of Daily Living Measure - Stairs (Soft Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0 (0.12 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Toronto Activities of Daily Living Measure - Sub Walk (Rigid Orthotics) (12 weeks)

1	randomised trials	serious[g]	not serious	not serious	serious[c]	none	24	24	-	MD 0.9 higher (0.24 higher to 1.56 higher)	⨁⨁◯◯ Low	CRITICAL ***Significant** **Favors Orthotics**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
FUNCTIONAL STATUS: Toronto Activities of Daily Living Measure - Sub Walk (Soft Orthotics) (12 weeks)												
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0.6 higher (0.09 lower to 1.29 higher)	⊕◯◯◯ Very low	CRITICAL NS
Function as inferred from 50' Walking (12 weeks)												
1	randomised trials	serious[g]	not serious	serious	very serious[b]	none	24	24	-	MD 0.2 higher (2.18 lower to 2.58 higher)	⊕◯◯◯ Very low	CRITICAL NS
FUNCTIONAL STATUS: 50' Walking (12 weeks)												
1	randomised trials	serious[g]	not serious	not serious	very serious[b]	none	24	24	-	MD 0 (2.29 lower to 2.29 higher)	⊕◯◯◯ Very low	CRITICAL NS

Critical Outcomes <12 weeks (1 to 2 months)

FUNCTIONAL STATUS: Health Assessment Questionnaire (0-3) (Mean Change Scores) (2 months)												
1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	15	13	-	MD 0.2 lower (0.36 lower to 0.04 lower)	⊕⊕◯◯ Low	CRITICAL *Significant Favors Orthotics
FUNCTIONAL STATUS: Gait - Normal Velocity (Mean Change Scores) (2 Months)												
1	randomised trials	serious[d]	not serious	not serious	very serious[b]	none	15	13	-	MD 7.5 higher (15.17 lower to 30.17 higher)	⊕◯◯◯ Very low	CRITICAL NS
FUNCTIONAL STATUS: Gait - Fast Velocity (Mean Change Scores) (2 Months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[d]	not serious	not serious	very serious[b]	none	15	13	-	MD 7.9 higher (17.87 lower to 33.67 higher)	⊕◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Walking Speed (1 Month)

1	randomised trials	serious[e]	not serious	not serious	very serious[b]	none	16	16	-	MD 0.22 higher (0.37 lower to 0.81 higher)	⊕◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Timed-Up-and-Go Test (4 weeks)

1	randomised trials	not serious	not serious	not serious	serious[c]	none	40	41	-	MD 0.99 lower (1.88 lower to 0.1 lower)	⊕⊕⊕◯ Moderate	CRITICAL ***Significant** **Favors Orthotics**

FUNCTIONAL STATUS: Berg balance scale 4 weeks

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	40	41	-	MD 1.35 higher (0.88 lower to 3.58 higher)	⊕⊕◯◯ Low	CRITICAL NS

PAIN: Foot function index (pain subscale) (4 weeks)

1	randomised trials	not serious	not serious	not serious	serious[c]	none	40	41	-	MD 1.7 lower (2.76 lower to 0.64 lower)	⊕⊕⊕◯ Moderate	CRITICAL ***Significant** **Favors Orthotics**

FUNCTIONAL STATUS: Foot function index (total score) (4 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[c]	none	40	41	-	MD 1.34 lower (2.19 lower to 0.49 lower)	⨁⨁⨁◯ Moderate	CRITICAL ***Significant** **Favors Orthotics**

FUNCTIONAL STATUS: inferred from Fatigue (0-100) (Mean Change Scores) (2 months)

1	randomised trials	serious[d]	not serious	serious[f]	very serious[b]	none	15	13	-	MD 14.8 lower (31.71 lower to 2.11 higher)	⨁◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Foot function index (activity limitation subscale) (4 weeks)

1	randomised trials	not serious	not serious	not serious	serious[c]	none	40	41	-	MD 1.25 lower (2.13 lower to 0.37 lower)	⨁⨁⨁◯ Moderate	CRITICAL ***Significant** **Favors Orthotics**

PAIN: Pain-Free Walk Time (up to 60 min) (Mean Change Scores) (2 months)

1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	15	13	-	MD 18.2 higher (8.15 higher to 28.25 higher)	⨁⨁◯◯ Low	CRITICAL ***Significant** **Favors Orthotics**

PAIN: Pain During Walking (1 Month)

1	randomised trials	serious[e]	not serious	not serious	serious[c]	none	16	16	-	MD 23.19 lower (32.97 lower to 13.41 lower)	⨁⨁◯◯ Low	CRITICAL ***Significant** **Favors Orthotics**

PAIN: Walk Pain (0-100) (Mean Change Scores) (2 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	15	13	-	MD 18.7 lower (28.67 lower to 8.73 lower)	⊕⊕◯◯ Low	CRITICAL *Significant Favors Orthotics

PAIN: Non-weight bearing pain (0-100) (Mean Change Scores) (2 months)

1	randomised trials	serious[d]	not serious	not serious	very serious[b]	none	15	13	-	MD 5 lower (15.4 lower to 5.4 higher)	⊕◯◯◯ Very low	CRITICAL NS

PAIN: Stair Pain (0-100) (Mean Change Scores) (2 months)

1	randomised trials	serious[d]	not serious	not serious	serious[c]	none	15	13	-	MD 22 lower (33.12 lower to 10.88 lower)	⊕⊕◯◯ Low	CRITICAL *Significant Favors Orthotics

Important Outcomes ≥12 weeks (12 weeks to 30 months)

DISEASE ACTIVITY: Disease Activity Score (DAS28) (Mean Change Scores - area under the curve) (30 months)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	50	48	-	MD 6.6 higher (8.97 lower to 22.17 higher)	⊕◯◯◯ Very low	IMPORTANT NS

DISEASE ACTIVITY: inferred from Lower Extremity Synovitis Joint Count (12 Weeks)

1	randomised trials	serious[g]	not serious	serious[f]	very serious[b]	none	24	24	-	MD 2.6 lower (7.3 lower to 2.1 higher)	⊕◯◯◯ Very low	IMPORTANT NS

DISEASE ACTIVITY: inferred from Lower Extremity Synovitis Joint Count (12 Weeks)

1	randomised trials	serious[g]	not serious	serious[f]	very serious[b]	none	24	24	-	MD 2.4 lower (7.31 lower to 2.51 higher)	⊕◯◯◯ Very low	IMPORTANT NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	No Orthotics	Relative (95% CI)	Absolute (95% CI)		
DISEASE ACTIVITY: inferred from Metatarsal Phalangeal Synovitis Joint Count (12 Weeks)												
1	randomised trials	serious[g]	not serious	serious[f]	very serious[b]	none	24	24	-	MD 0.2 lower (1.87 lower to 1.47 higher)	⊕◯◯◯ Very low	IMPORTANT NS
DISEASE ACTIVITY: inferred from Metatarsal Phalangeal Synovitis Joint Count (12 Weeks)												
1	randomised trials	serious[g]	not serious	serious[f]	very serious[b]	none	24	24	-	MD 0.2 lower (1.9 lower to 1.5 higher)	⊕◯◯◯ Very low	IMPORTANT NS

Important Outcomes <12 weeks (2 months)

QOL: Well-Being (0-100) (2 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	15	13	-	MD 12.5 lower (27.93 lower to 2.93 higher)	⊕⊕◯◯ Low	IMPORTANT NS

CI: confidence interval; **MD:** mean difference

Explanations

a. 5134 Revman Bias Table: 4L, 1H, 1U. Participants not blinded.

b. Single study, and confidence interval for effect size spans across the null value.

c. Single study.

d. 4255 Revman Bias Table: 3U, 2L, 1H. Impossible to blind, and several factors unspecified.

e. 2574 Revman Bias Table: 3U, 2L, 1H. Impossible to blind, and several factors unspecified.

f. Outcome is a surrogate measure.

g. Chalmers Revman Bias Table: 3L, 2U, 1H. Impossible to blind, and several factors unspecified.

Table 2: Additional data for Orthotics vs No Orthotics

Summary of findings: For the 3 outcomes below, there were no significant differences between the intervention and control group at 30 month follow-up.

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results n = 50 for intervention; n = 48 for control Timepoint is 30 months Data presented as Median (IQR) Change from Baseline
5134, Woodburn, 2002	RCT	30 months	Total n = 98 Patients w RA Intervention: mean age: 54.0 y +/- 11.8; 68.0% female; disease duration (median and IQR): 3 (1,7) Control: mean age: 53.1 +/- 11.1; 60.8% female; disease duration (median and IQR): 3 (2,6)	Intervention: custom foot orthotics w podiatry supervision Control: no orthotics assigned at baseline, but they were used if they were prescribed as part of usual treatment later on in the study	**Health Assessment Questionnaire (0-3) (Negative)** Intervention: 0 (-7.5,0.8) Control: 0 (-6.5,0.7) **Larsen Index (Hands) (0-150) (Negative)** Intervention: 54 (0,99) Control: 57 (31,169) **Larsen Index (Feet) (0-50) (Negative)** Intervention: 60 (7,155) Control: 62 (28,149)

All outcomes are negative, meaning that lower scores are better.

Orthotics vs. Placebo

***Orthotics vs Placebo**: Overall quality of evidence across critical outcomes: Very Low*

6 RCTs compared Orthotics versus Placebo Orthotics [6,7,8,9,10,11]. 5 studies contained RevMan data only [6,7,8,9,10] while 1 study contained non-RevMannable data only [11].

≥ 12 Weeks

All studies had at least a 12-week followup. The most overlap in outcomes occurred for the Foot Function Index (FFI) (Total Score and 3 subscales). For the FFI Total Score, FFI Pain Subscale, and FFI Disability/Difficulty Subscale, each of the 3 studies favored the intervention group – however, the total effect for each outcome was not significant [6,7,8]. For the FFI Activity Limitation Subscale, Reina-Bueno [6] and Conrad [8] favored the control group, while Moreiera [7] favored the intervention group; the total effect was not significant. Novak [9] and Rome [10] also reported on mean change scores for FFI Pain Subscale, and favored the intervention group, although this was non-significant. Rome [10] also found that mean change scores for FFI Disability Subscale significantly favored the intervention group while for FFI Activity Limitations Subscale there was no effect.

Another outcome for which there was considerable overlap was foot pain. After combining the *standardized mean differences* of Reina-Bueno [6], Moreiera [7], and Conrad [8], there was a slight effect (non-significant) favoring the intervention group. Moreiera [7] also found a significant effect favoring the intervention group for foot pain while walking. There were non-significant differences for foot pain days (out of 90) [6] and painful foot joint count [8]. Finally, the 6-minute walk test showed non-significant improvements for the intervention group as measured by post-data [7] and mean absolute change scores [9].

Other Functional Status outcomes measured included the following validated instruments: FHSQ, SF36, HAQ, SF12, Manchester Foot Pain and Disability Index, and Total Disability. The Foot Health Status Questionnaire (FHSQ) found very few differences across each of the 8 subscales [7]. There was a non-significant favoring of the intervention group for General Foot Health, Foot Function, Foot Pain, Physical Activity, Social Capacity, and Foot Health subscales; and non-significant favoring of placebo for Vigour and General Health subscales. The SF36 showed even fewer differences across the 8 subscales, with non-significant favoring of the intervention group for the Physical Role, Bodily Pain, Social Role Functioning, Emotional Role Functioning, and Mental Health subscales; and minimal to no differences were found for the Physical Functioning, General Health State, and Vitality subscales [7].

For the Health Assessment Questionnaire (HAQ), Moreiera found a non-significant effect favoring the intervention group [7]. Reina-Bueno found minimal differences between the intervention and placebo for the Physical SF12, Mental SF12, and the Manchester Foot Pain and Disability Index [6], as did Conrad with total disability [8].

Finally, remaining outcomes included total painful joint count [8] and quality-adjusted life years [10], which slightly favored control and intervention groups, respectively. Additionally, Budiman [11] conducted a 5-year RCT and reported non-RevMannable results. A limitation was that no quantitative data was reported at followup, but the findings mentioned that there were minimal or no improvements in the intervention group compared with the placebo group.

Table 3: RCTs: Orthotics compared to Placebo

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
Critical Outcomes ≥12 weeks (3 months to 3 years)												
FUNCTIONAL STATUS: FHSQ Foot Health (6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 0.6 higher (8.79 lower to 9.99 higher)	⨁⨁◯◯ Low	CRITICAL NS
FUNCTIONAL STATUS: SF36 Physical Functioning (6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 0.1 higher (10.44 lower to 10.64 higher)	⨁⨁◯◯ Low	CRITICAL NS
FUNCTIONAL STATUS: SF36 Physical Role (6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 10.6 higher (9.01 lower to 30.21 higher)	⨁⨁◯◯ Low	CRITICAL NS
FUNCTIONAL STATUS: Health Assessment Questionnaire (6 Months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 0.15 lower (0.38 lower to 0.08 higher)	⨁⨁◯◯ Low	CRITICAL NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
FUNCTIONAL STATUS: 6 Min Walk Test (6 Months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 13.6 higher (22.04 lower to 49.24 higher)	⨁⨁◯◯ Low	CRITICAL NS
FUNCTIONAL STATUS: 6 Min Walk Test (Mean Absolute Change) (6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	20	19	-	MD 15.55 higher (17.35 lower to 48.45 higher)	⨁⨁◯◯ Low	CRITICAL NS
FUNCTIONAL STATUS: Total Disability (AIMS) (3 year)												
1	randomised trials	serious[j]	not serious	not serious	very serious[b]	none	44	44	-	MD 1.1 lower (8.13 lower to 5.93 higher)	⨁◯◯◯ Very low	CRITICAL NS
FUNCTIONAL STATUS: Foot Function Index (disability/difficulty subscale) (mean change scores) (16 weeks)												
1	randomised trials	serious[f]	not serious	not serious	serious[g]	none	20	21	-	MD 12.5 lower (24.96 lower to 0.04 lower)	⨁⨁◯◯ Low	CRITICAL ***Significant Favors Orthotics**
FUNCTIONAL STATUS: Foot Function Index (Activity Limitation Subscale) (12+ weeks)												
3	randomised trials	serious[c]	serious[i]	not serious	very serious[e]	none	109	107	-	MD 1.06 higher (3.47 lower to 5.58 higher)	⨁◯◯◯ Very low	CRITICAL NS
FUNCTIONAL STATUS: Foot Function Index (activity limitation subscale) (mean change score) (16 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[f]	not serious	not serious	very serious[b]	none	20	21	-	MD 1.3 lower (10 lower to 7.4 higher)	⊕◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Foot Function Index (Total Score) (12+ weeks)

3	randomised trials	serious[c]	not serious	not serious	serious[d]	none	109	107	-	MD 3.83 lower (9.71 lower to 2.06 higher)	⊕⊕◯◯ Low	CRITICAL NS

FUNCTIONAL STATUS: FHSQ General Foot Health (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 2.9 higher (9.83 lower to 15.63 higher)	⊕⊕◯◯ Low	CRITICAL NS

FUNCTIONAL STATUS: FHSQ Foot Function (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 1.3 higher (9.75 lower to 12.35 higher)	⊕⊕◯◯ Low	CRITICAL NS

FUNCTIONAL STATUS: Physical SF-12 (QOL) (90 days)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	28	25	-	MD 1.28 lower (6.49 lower to 3.93 higher)	⊕◯◯◯ Very low	CRITICAL NS

FUNCTIONAL STATUS: Foot Function Index (Disability/Difficulty Subscale) (12+ weeks)

3	randomised trials	serious[c]	not serious	not serious	very serious[e]	none	109	107	-	MD 6.61 lower (14.32 lower to 1.1 higher)	⊕◯◯◯ Very low	CRITICAL NS

PAIN: Foot Function Index Pain Subscale (Mean Absolute Change) (12+ weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious[h]	not serious	not serious	serious[d]	none	40	40	-	MD 10.06 lower (19.04 lower to 1.08 lower)	⊕⊕◯◯ Low	CRITICAL *Significant Favors Orthotics

PAIN: Foot function index (pain) (12+ weeks)

3	randomised trials	serious[c]	not serious	not serious	serious[d]	none	109	107	-	MD 5.36 lower (12.5 lower to 1.79 higher)	⊕⊕◯◯ Low	CRITICAL NS

PAIN: FHSQ Foot Pain (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 5.4 higher (6.08 lower to 16.88 higher)	⊕⊕◯◯ Low	CRITICAL NS

PAIN: SF36 Bodily Pain (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 2.6 higher (6.59 lower to 11.79 higher)	⊕⊕◯◯ Low	CRITICAL NS

PAIN: Manchester foot pain and disability index 90 days

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	28	25	-	MD 0.46 lower (6.15 lower to 5.23 higher)	⊕◯◯◯ Very low	CRITICAL NS

PAIN: Foot Pain (standardized mean difference) (12+ weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials	serious[c]	very serious[k]	not serious	very serious[e]	none	109	107	-	SMD 0.16 lower (0.42 lower to 0.11 higher) CALCULATED SMD: 0.43 lower (1.13 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL NS

PAIN: Foot Pain while Walking (VAS) (R Foot) (6 months)

1	randomised trials	not serious	not serious	not serious	serious[g]	none	37	38	-	MD 2.2 lower (3.35 lower to 1.05 lower)	⨁⨁⨁◯ Moderate	CRITICAL *Significant Favors Orthotics

PAIN: Foot pain days (90 days)

1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	28	25	-	MD 0.21 lower (1.66 lower to 1.24 higher)	⨁◯◯◯ Very low	CRITICAL NS

PAIN: Painful Foot Joint Count (3 year)

1	randomised trials	serious[j]	not serious	not serious	very serious[b]	none	44	44	-	MD 0.2 higher (0.55 lower to 0.95 higher)	⨁◯◯◯ Very low	CRITICAL NS

PAIN: Total Painful Joint Count (3 year)

1	randomised trials	serious[j]	not serious	not serious	very serious[b]	none	44	44	-	MD 1.1 higher (2.34 lower to 4.54 higher)	⨁◯◯◯ Very low	CRITICAL NS

Important Outcomes ≥12 weeks (3 to 6 months)

QOL: FHSQ Physical Activity (6 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 2.2 higher (8.33 lower to 12.73 higher)	⨁⨁◯◯ Low	IMPORTANT NS

QOL: FHSQ Vigour (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 2.3 lower (12.81 lower to 8.21 higher)	⨁⨁◯◯ Low	IMPORTANT NS

QOL: SF36 General Health State (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 1 higher (9.44 lower to 11.44 higher)	⨁⨁◯◯ Low	IMPORTANT NS

QOL: SF36 Vitality (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 1 higher (10.02 lower to 12.02 higher)	⨁⨁◯◯ Low	IMPORTANT NS

QOL: SF36 Social Role Functioning (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 6.8 higher (6.82 lower to 20.42 higher)	⨁⨁◯◯ Low	IMPORTANT NS

QOL: SF36 Emotional Role Functioning (6 months)

1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 2.8 higher (17.89 lower to 23.49 higher)	⨁⨁◯◯ Low	IMPORTANT NS

QOL: FHSQ Social Capacity (6 months)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Orthotics	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 6 higher (6.24 lower to 18.24 higher)	⨁⨁◯◯ Low	IMPORTANT NS
QOL: FHSQ General Health (6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 1.6 lower (13.54 lower to 10.34 higher)	⨁⨁◯◯ Low	IMPORTANT NS
QOL: QALY (16 weeks)												
1	randomised trials	serious[f]	not serious	not serious	very serious[b]	none	20	21	-	MD 0.04 higher (0.01 lower to 0.09 higher)	⨁◯◯◯ Very low	IMPORTANT NS
MENTAL HEALTH: SF36 Mental Health (6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	37	38	-	MD 1.1 higher (9.15 lower to 11.35 higher)	⨁⨁◯◯ Low	IMPORTANT NS
MENTAL HEALTH: Mental SF-12 (QOL) (90 days)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	28	25	-	MD 1.04 higher (4.59 lower to 6.67 higher)	⨁◯◯◯ Very low	IMPORTANT NS

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. 1897 Revman Bias Table: 4L, 1H, 1U. Possible attrition bias.

b. Single study, and confidence interval for effect size spans across the null value.

c. 2 studies have serious classification, and 1 not serious.

d. The confidence intervals of effect sizes for all studies span across the null value.

e. The confidence intervals of effect sizes for all studies span across the null value, plus wide CI(s).

f. 907 Revman Bias Table: 4L, 2H. Assessor blinding not reported, and significant differential attrition.

g. Single study.

h. 1 study has serious classification, and 1 not serious.

i. Inconsistency in effect direction.

j. 3093 Revman Bias Table: 3L, 2U, 1H. Demographics not reported, and several other factors not reported.

k. Inconsistency in effect direction and magnitude.

Table 4: Additional data for Orthotics vs Placebo

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
698, Budiman-Mak et al., 1955	Double-blind randomized controlled trial	5 years	RA patients = 102 Foot orthoses n = 52 Age, mean: 60.2 (SD = 10.6) Male: 46 Female: 6 Control n = 50 Age, mean: 58.8 (SD = 11.9) Male: 43 Female: 7	Patients were randomized to wear a specialized foot orthoses designed to maintain an angular , anatomic relationship between the forefoot and hindfoot during ambulation, or a placebo fabricated as a thin flexible leather shell molded over a plaster impression of the patients foot.	Measures of pain, disability, and function of the foot showed little or no benefit from the orthosis.

References

1. Gaino JZ, Bértolo MB, Nunes CS, et al. The effect of foot orthoses on balance, foot function, and mobility in rheumatoid arthritis: A randomized controlled clinical trial. Clinical Rehabilitation. 2021;35(7):1011-1020.
2. Mejjad O, Vittecoq O, Pouplin S, et al. Foot orthotics decrease pain but do not improve gait in rheumatoid arthritis patients. Joint Bone Spine. 2004;71:542-545.
3. Fransen M, Edmonds J. Off-the-shelf orthopedic footwear for people with rheumatoid arthritis. 1997;10(4):250-256.
4. Chalmers AC, Busby C, Goyert J, Porter B, Schulzer M. Metatarsalgia and rheumatoid arthritis – a randomized, single blind, sequential trial comparing 2 types of foot orthoses and supportive shoes. J Rheumatology. 2000;27(7):1643-1647.
5. Woodburn J, Barker S, Helliwell PS. A randomized controlled trial of foot orthoses in rheumatoid arthritis. J Rheum. 2002;29(7):1377-1383.
6. Reina-Bueno M, Vázquez-Bautista MDC, Pérez-Garcia S, Rosende-Bautista C, Sáez-Díaz A, Munuera-Martínez PV. Effectiveness of custom-made foot orthoses in patients with rheumatoid arthritis: a randomized controlled trial. Clinical Rehabilitation. 2019;33(4):661-669.
7. Moreira E, Jones A, Oliveira HA, Jennings F, Fernandes ARC, Natour J. Effectiveness of insole use in rheumatoid feet: a randomized controlled trial. Scandinavian Journal of Rheumatology. 2016;1-8.
8. Conrad KJ, Budiman-Mak, Roach KE, Hedeker D. Impacts of foot orthoses on pain and disability in rheumatoid arthritis. J Clin Epidemiol. 1996;49(1):1-7.
9. Novak P, Burger H, Tomsic M, Marincek C, Vidmar G. Influence of foot orthoses on plantar pressures, foot pain and walking ability of rheumatoid arthritis patients–a randomised controlled study. Disability and Rehabilitation. 2009;31(8):638-645.
10. Rome K, Clark H, Gray J, McMeekin P, Plant M, Dixon J. Clinical effectiveness and cost-effectiveness of foot orthoses for people with established rheumatoid arthritis: an exploratory clinical trial. 2016;1-7.
11. Budiman-Mak E, Conrad KJ, Roach KE, et al. Can foot orthoses prevent hallux valgus deformity in rheumatoid arthritis? A randomized clinical trial. J Clin Rheumatol. 1995;1:313-321.

PICO 11. Should patients with RA and knee involvement use bracing/orthoses?

No studies met inclusion criteria for this question.

Rehabilitation

PICO 12: Should patients with RA use joint protection techniques?

Summary: Literature searches identified 3 small, randomized control trials [1,2,3] addressing this question. Joint protection techniques administered included education on rheumatoid arthritis, mechanisms of pain and stress, home exercise programs, rest to avoid joint overload, principles of joint protection and energy conservation, and assistive technical equipment design to reduce joint forces such as modified handles on utensils.

Masiero et al,[3] found statistically significantly better functional status measures after 12 weeks (see Table 1) for those receiving joint protection programs than the control group. The difference was modest (e.g., AIMS2 physical function scale was only 1.7 points better on a 0 to 10 scale). Neither pain nor disease activity were statistically significantly different between groups.

Neither article by Hammond et al,[1,2] found statistically significant between-group differences in pain, disease activity, or functional status.

No harms were reported in any of the three studies, suggesting that education and joint protection are not harming patients.

The evidence was rated low quality of evidence due to low number of total studies [1,2,3] looking at use of joint protection to alter pain, disease activity, and functional status.

Quality of evidence across all critical outcomes: Low

Table 1. Data from randomized controlled trials

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Experimental (Drugs + Educational–behavioral Joint Protection training group)	Control (Drugs)	Relative (95% CI)	Absolute (95% CI)		
Pain > 12 weeks (3 months to 6 months)												
1	randomised trials	Serious[c]	not serious	not serious	Serious[a]	none	36	34	-	MD **5.1 lower** (15.31 lower to 5.11 higher)	⨁⨁◯◯ Low	CRITICAL Not statistically significant
Functional Status (AIMS2 - physical function) > 12 weeks (3 months to 6 months)												
1	randomised trials	Serious[c]	serious[a]	not serious	serious[a]	none	36	34	-	MD **1.7 lower** (2.5 lower to 0.9 lower)	⨁◯◯◯ Very Low	CRITICAL **Statistically significant in favor of joint protection**
Functional Status (AIMS2 - Psychological) > 12 weeks (3 months to 6 months)												
1	randomised trials	Serious[c]	serious[a]	not serious	Serious[a]	none	36	34	-	MD **1 lower** (1.96 lower to 0.04 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of joint protection**
Functional Status (AIMS2 - Symptoms) > 12 weeks (3 months to 6 months)												
1	randomised trials	Serious[c]	serious[c]	not serious	not serious	none	36	34	-	MD **1.3 lower** (2.24 lower to 0.36 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of joint protection**
Functional Status (AIMS2 - Social Interaction) > 12 weeks (3 months to 6 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RCTs: Experimental (Drugs + Educational–behavioral Joint Protection training group)	Control (Drugs)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	Serious[c]	not serious	not serious	serious[a]	none	36	34	-	MD **1.6 lower** (2.66 lower to 0.54 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of joint protection**

Functional Status (AIMS2 - Work) > 12 weeks (3 months to 6 months)

1	randomised trials	Serious[c]	serious[ac]	not serious	not serious	none	36	34	-	MD **1 lower** (1.54 lower to 0.46 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of joint protection**

Functional Status (HAQ) > 12 weeks (3 months to 6 months)

1	randomised trials	Serious[c]	serious[c]	not serious	not serious	none	36	34	-	MD **0.31 lower** (0.54 lower to 0.08 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant in favor of joint protection**

Disease Activity (RAI) > 12 weeks (3 months to 6 months)

1	randomised trials	Serious[c]	not serious	serious[b]	serious[a]	none	36	34	-	MD **4.2 lower** (8.51 lower to 0.11 higher)	⨁◯◯◯ Very Low	IMPORTANT Not statistically significant

CI: confidence interval; **MD:** mean difference

Explanations

a. Low N resulting in wide CIs across 0
b. Indirect measure
c. Personnel or concealment bias

Table 2. Additional Data from RCT and Observational Studies

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
891 Hammond 2002	RCT	6 months	30 RA patients	Joint protection program (intervention) vs. No treatment (waiting control)	There were no significant differences in any secondary outcome measures knowledge, pain, strength, fatigue, HAQ, self-efficacy, and RAI between the two groups at 3 months ("3-month numerical data were not reported
1274 Hammond 1999	RCT	24 weeks	27 RA patients	Joint protection program + home visit (intervention) vs. No treatment (control)	Visual Analogue Scale for pain: Median (IQR) Intervention: 62.00 (40.50-72.50) Control: 24.00 (4.50-54.25) No significant changes in measures of pain, functional disability, grip strength, self-efficacy or helplessness occurred post-education.

References:

1. Hammond, A., Lincoln, N., & Sutcliffe, L. (1999). A crossover trial evaluating an educational–behavioural joint protection programme for people with rheumatoid arthritis. *Patient Education and Counseling*, *37*(1), 19-32.

2. Hammond, A., Jeffreson, P., Jones, N., Gallagher, J., & Jones, T. (2002). Clinical applicability of an educational-behavioural joint protection programme for people with rheumatoid arthritis. *British Journal of Occupational Therapy*, *65*(9), 405-412.

3. Masiero, S., Boniolo, A., Wassermann, L., Machiedo, H., Volante, D., & Punzi, L. (2007). Effects of an educational–behavioral joint protection program on people with moderate to severe rheumatoid arthritis: a randomized controlled trial. *Clinical Rheumatology*, *26*(12), 2043-2050.

PICO 13. Should patients with RA use activity pacing/energy conservation/activity modification/fatigue management techniques?

No studies met inclusion criteria for this question.

PICO 14. Should patients with RA use assistive devices?

No studies met inclusion criteria for this question.

PICO 15. Should patients with RA use adaptive equipment?

No studies met inclusion criteria for this question.

PICO 16. Should patients with RA use environmental adaptations?

No studies met inclusion criteria for this question.

Psychosocial and vocational

PICO 17: Should patients with RA participate in comprehensive occupational therapy?
Evidence Summary: We included eight randomized controlled trials (RCTs)[1-8] addressing this PICO question.

- Six RCTs[1-5,7] compared an **occupational therapy program** to a control group.
- One RCT(Ayhan et al.)[6] compared an **inpatient rehabilitation model** versus a home exercise model
- One RCT(Shearn et al.[8]) compared a **stress management program** versus treatment as usual.

We categorized the latter two as comprehensive occupational therapy interventions, but analyzed them separately from the other six studies. Critical outcomes for this PICO were pain and functional status, and we classified each outcome as either short term (<12 weeks) or long terms (>=12 weeks).

Six RCTS[1-5,7] evaluated an occupational therapy program. Programs varied across studies, but all included some form of occupational therapy led by a rheumatologist, sometimes combined with educational or cognitive behavioral training sessions. Here we summarize the results for the critical outcomes:

- Three studies [2,5,7] evaluated the effect of the program on pain assessment at 12 weeks or later, and none found that there was a statistically significant difference from the control group.
- Functional status was assessed through various measurements, such as the Health Assessment Questionnaire (HAQ), the Arthritis Impact Measurement Scale (AIMS2), McGill indexes, the Visual Analogue Scale (VAS), and the Canadian Occupational Performance Measure (COPM). Results for the three long-term studies were mixed, but generally showed an improvement in measures such as mobility, grip, arm movement, and tender and swollen joint count.
- In short-term studies, one study, Tonga et al.[3], found at 4 weeks, patients receiving OT showed statistically significant improvement in pain assessment and the various measures of functional status based on the HAQ, the AIMS2, McGill indexes, the VAS, and the COPM compared to the control group. Another study, Helewa et al.[4], found that at 6 weeks an occupational therapy program had a statistically significant effect on functional status in the treatment group compared to the control group.

The overall certainty of evidence for pain and functional status from these studies was very low, primarily due to concerns around small sample size, statistically nonsignificant between-group differences, and a lack of patient and outcome assessor blinding.

One RCT (Ayhan et al.[6]) had implemented an inpatient rehabilitation model. Both groups received education and disease info, joint protections, energy conservation, sleep hygiene, relaxation training, physical activity recommendations and management of fatigue, pain, flares. Patient enrolled in the inpatient rehabilitation program also received 15 daily sessions of physical therapy, occupational therapy over the course of the

program. They found that patients enrolled in an inpatient rehabilitation model had a statistically significant improvement in HAQ and disease activity compared with those enrolled in a home exercise model. The certainty of evidence for this study however was very low, primarily due to concerns about high attrition, small sample size, lack of blinding, and large baseline differences between groups.

The final study included for this PICO (Shearn et al.[8]) focused on a stress management program. Patients were randomized to either a stress management program led by a psychologist, with a focus on self-responsibility, building relationships, and decrease social isolation, or they were assigned to a control group receiving treatment as usual. The study found that, in the long term, there was no significant difference between patients treated with a stress management program versus control for either pain assessments or functional status. The certainty of evidence was very low due to small sample size, high attrition, and a lack of patient or outcome assessor blinding.

Quality of evidence across all critical outcomes: Very low

Table 1: Occupational Therapy compared to control

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
Pain: Change in VAS pain (0-100, higher score indicates better health) 6 - 24 months												
2	randomised trials	not serious	serious[d]	not serious	serious[a]	none	178	180	-	MD **5.68 lower** (11.58 lower to 0.21 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
Function: Change in Arthritis Impact Measurement Scales II Pain (0-10; higher score indicates more problems) 6 - 24 months												
2	randomised trials	not serious	serious[d]	not serious	serious[a]	none	178	180	-	MD **0.22 lower** (0.66 lower to 0.21 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference
Function: Change in Arthritis Impact Measurement Scales II Work ability (0-10; higher score indicates more problems) 24 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	65	64	-	MD **0.12 higher** (0.83 lower to 1.07 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

Function: Change in Arthritis Impact Measurement Scales II Physical Function (0-10; higher score indicates more problems) 24 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	162	164	-	MD **0.09 higher** (0.18 lower to 0.36 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

Function: Change in Arthritis Impact Measurement Scales II Affect scale (0-10; higher score indicates more problems) 24 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	162	164	-	MD **0.12 lower** (0.38 lower to 0.14 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

Function: Change in HAQ 24 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	162	164	-	MD **0.03 higher** (0.15 lower to 0.21 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

Function: Change in Arthritis Helplessness Index 6 - 24 months

2	randomised trials	not serious	serious[d]	not serious	serious[a]	none	178	180	-	MD **0.58 lower** (1.59 lower to 0.43 higher)	⨁⨁◯◯ Low	CRITICAL No statistically significant difference

Function as inferred from COPM Satisfaction (1-10, higher score is better) 6 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	16	16	-	MD **3.83 higher** (2.24 higher to 5.42 higher)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**

COPM Performance (1-10, higher score is better) 6 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	16	16	-	MD **3.38 higher** (2.17 higher to 4.59 higher)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant difference favoring OT**

Change in EuroQol Global (0-100, higher score indicates better health) 6 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	16	16	-	MD **27.29 higher** (7.43 higher to 47.15 higher)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant difference favoring OT**

Change in EuroQol Index (0-1, higher score indicates better health) 6 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	16	16	-	MD **0.28 higher** (0.06 higher to 0.5 higher)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from COPM Satisfaction (1-10, higher score is better) 6 months												
1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	16	16	-	MD **3.83 higher** (2.24 higher to 5.42 higher)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from change in Arthritis Impact Measurement Scales II Satisfaction with health (0-10; higher score indicates more problems) 24 months												
1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	162	164	-	MD **1.27 higher** (0.26 lower to 2.8 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
HAQ 3 months												
1	randomised trials	serious[f]	not serious	not serious	very serious[a,b]	none	30	30	-	MD **0.16 lower** (0.29 lower to 0.03 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**
Pain: McGill VAS (1-15; lower score is better) 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **1.15 lower** (2 lower to 0.3 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**
McGill Affective Index (1-15; lower score is better) 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **0.65 lower** (0.92 lower to 0.38 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
Function: RAQL (0-30, lower score indicates higher quality of life) 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **6.1 lower** (9.02 lower to 3.18 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**
Function: COPM performance (1-10, higher score is better) 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **1.9 higher** (1.15 higher to 2.65 higher)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**
Function: AIMS2 Arthritis Pain (0-10; higher score indicates more problems) 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **1.84 lower** (2.63 lower to 1.05 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significa difference favoring C**
Function: HAQ Total 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **0.51 lower** (0.7 lower to 0.32 lower)	⨁⨁◯◯ Low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from AIMS2 Mobility (0-10; higher score indicates more problems) 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.98 lower** (1.68 lower to 0.28 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from HAQ Rising 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.85 lower** (1.18 lower to 0.52 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from HAQ Eating 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.55 lower** (0.88 lower to 0.22 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from HAQ Walking 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.05 lower** (1.39 lower to 0.71 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from HAQ Grip 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.4 lower** (0.75 lower to 0.05 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from HAQ Activities 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.95 lower** (1.38 lower to 0.52 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**
Function as inferred from AIMS2 Walking and Bending (0-10; higher score indicates more problems) 1 month												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.75 lower** (2.64 lower to 0.86 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**

Function as inferred from AIMS2 Hand and finger (0-10; higher score indicates more problems) 1 month

1	randomised trials	serious[g]	not serious	serious[c]	very serious[a,b]	none	20	20	-	MD **0.2 lower** (0.98 lower to 0.58 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference

Function as inferred from AIMS2 Self care (0-10; higher score indicates more problems) 1 month

1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.09 lower** (1.83 lower to 0.35 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**

Function as inferred from AIMS2 Arm function (0-10; higher score indicates more problems) 1 month

1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.23 lower** (2.19 lower to 0.27 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**

Function as inferred from AIMS2 Household tasks (0-10; higher score indicates more problems) 1 month

1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.75 lower** (2.54 lower to 0.96 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**

Function as inferred from AIMS2 arthritis impact (0-10; higher score indicates more problems) 1 month

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[g]	not serious	serious[c]	serious[a,b]	none	20	20	-	MD **0.27 lower** (0.95 lower to 0.41 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference

Function as inferred from COPM Satisfaction (1-10, higher score is better) 1 month

1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **3.25 higher** (2.33 higher to 4.17 higher)	⨁◯◯◯ Very low	CRITICAL **Statistically significant difference favoring OT**

Change in number of doctor visit for arthritis 24 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	162	164	-	MD **0.1 higher** (0.41 lower to 0.61 higher)	⊕◯◯◯ Very low	IMPORTANT No statistically significant difference

Change in Jebsen test (seconds) 24 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	162	164	-	MD **0.92 higher** (1.03 lower to 2.87 higher)	⊕⊕◯◯ Low	IMPORTANT No statistically significant difference

Change in Total self efficacy scale (0-100) 24 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	162	164	-	MD **1.21 higher** (2.09 lower to 4.51 higher)	⊕◯◯◯ Very low	IMPORTANT No statistically significant difference

Function: Mobility (per dutch health questionnaire) 6 months

1	randomised trials	very serious[e]	not serious	not serious	very serious[a,b]	none	28	19	-	MD **0.9 lower** (5.14 lower to 3.34 higher)	⊕◯◯◯ Very low	IMPORTANT No statistically significant difference

Self care (per dutch health questionnaire) 6 months

1	randomised trials	very serious[e]	not serious	not serious	very serious[a,b]	none	28	19	-	MD **0.4 lower** (4.54 lower to 3.74 higher)	⊕◯◯◯ Very low	IMPORTANT No statistically significant difference

Anxiety (per dutch health questionnaire) 6 months

1	randomised trials	very serious[e]	not serious	not serious	very serious[a,b]	none	28	19	-	MD **0.5 lower** (4.47 lower to 3.47 higher)	⊕◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**

Depression (per dutch health questionnaire) 6 months

1	randomised trials	very serious[a]	not serious	not serious	very serious[a,b]	none	28	19	-	MD **0.1 lower** (2.41 lower to 2.21 higher)	⊕◯◯◯ Very low	IMPORTANT No statistically significant difference

Function: Change in Health Assessment Questionnaire disability index 6 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	16	16	-	MD **0.44 lower** (0.79 lower to 0.09 lower)	⊕⊕⊕◯ Moderate	IMPORTANT **Statistically significant difference favoring OT**

Change in RA Work Instability Scale 6 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	16	16	-	MD **2.8 lower** (5.22 lower to 0.38 lower)	⊕⊕⊕◯ Moderate	IMPORTANT **Statistically significant difference favoring OT**

Change in Visual Analog Scale Work Performance 6 months

1	randomised trials	not serious	not serious	not serious	serious[a]	none	16	16	-	MD **38.51 lower** (66.03 lower to 10.99 lower)	⊕⊕⊕◯ Moderate	IMPORTANT **Statistically significant difference favoring OT**

Change in Visual Analog Scale Work Satisfaction 6 months

1	randomised trials	not serious	not serious	serious[c]	serious[a]	none	16	16	-	MD **26.81 lower** (51.55 lower to 2.07 lower)	⊕⊕◯◯ Low	IMPORTANT **Statistically significant difference favoring OT**

Change in Work days missed per month 6 months

1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	16	16	-	MD **3.43 lower** (7.28 lower to 0.42 higher)	⊕⊕◯◯ Low	IMPORTANT **Statistically significant difference favoring OT**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
Change in Days missed/days worked per month, % 6 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	16	16	-	MD **0.18 lower** (0.38 lower to 0.02 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference
Change in Arthritis Impact Measurement Scales II Tension (0-10; higher score indicates more problems) 6 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	16	16	-	MD **0.57 lower** (1.47 lower to 0.33 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference
Change in Arthritis Impact Measurement Scales II Mood (0-10; higher score indicates more problems) 6 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	16	16	-	MD **0.06 higher** (0.76 lower to 0.88 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference
Disease activity: Change in DAS28 6 months												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	16	16	-	MD **1.05 lower** (1.93 lower to 0.17 lower)	⨁⨁⨁◯ Moderate	IMPORTANT **Statistically significant difference favoring OT**
Change in Patient global assessment 6 months												
1	randomised trials	not serious	not serious	not serious	very serious[a,b]	none	16	16	-	MD **18 lower** (40.87 lower to 4.87 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference
Disease activity as inferred from Change in early morning stiffness (mins) 24 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	162	164	-	MD **14.71 higher** (20.05 lower to 49.47 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference

Function as inferred from Change in Power grip (kg) 24 months

1	randomised trials	not serious	not serious	serious[c]	very serious[a,b]	none	162	164	-	MD **0.13 lower** (1.33 lower to 1.07 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference

Disease activity as inferred Change in Fatigue 24 months

1	randomised trials	not serious	not serious	serious[f]	very serious[b,c]	none	16	16	-	MD **0.06 lower** (0.59 lower to 0.47 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference

Disease activity as inferred from Change in 28 tender joint count 6 - 24 months

2	randomised trials	not serious	serious[d]	not serious	serious[a]	none	178	180	-	MD **1.08 lower** (2.51 lower to 0.34 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference

Disease activity as inferred Change in 28 swollen joint count 6 - 24 months

2	randomised trials	not serious	serious[d]	not serious	serious[a]	none	178	180	-	MD **0.52 lower** (1.86 lower to 0.81 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
AIMS2 Level of tension (0-10; higher score indicates more problems) 1 month												
1	randomised trials	serious[g]	not serious	not serious	very serious[a,b]	none	20	20	-	MD **0.63 lower** (1.27 lower to 0.01 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference
McGill Sensory Index (1-15; lower score is better) 1 month												
1	randomised trials	serious[g]	not serious	not serious	serious[a]	none	20	20	-	MD **3.6 lower** (4.72 lower to 2.48 lower)	⨁⨁◯◯ Low	IMPORTANT **Statistically significant difference favoring OT**
AIMS2 Health Perceptions (0-10; higher score indicates more problems) 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.18 lower** (2.04 lower to 0.32 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**
Function as inferred from AIMS2 Satisfaction (0-10; higher score indicates more problems) 1 month												
1	randomised trials	serious[g]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.63 lower** (2.27 lower to 0.99 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**
Function as inferred from AIMS2 mood (0-10; higher score indicates more problems) 1 month												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Occupational Therapy	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.82 lower** (1.59 lower to 0.05 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**

Function as inferred from HAQ Hygiene 1 month

1	randomised trials	serious[a]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.45 lower** (0.82 lower to 0.08 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**

Function as inferred from HAQ Dressing and grooming 1 month

1	randomised trials	serious[a]	not serious	serious[c]	serious[a]	none	20	20	-	MD **0.45 lower** (0.85 lower to 0.05 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**

Function as inferred from HAQ Reach 1 month

1	randomised trials	serious[a]	not serious	serious[c]	serious[a]	none	20	20	-	MD **1.1 lower** (1.45 lower to 0.75 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring OT**

CI: confidence interval; **MD:** mean difference

Explanations

a. Small sample size

b. Wide confidence interval

c. Surrogate measure

d. Significant difference of effect between studies

e. No patient or outcome assessor blinding, very little complete data provided. P-values for within group changes reported for select outcomes, without point estimate sizes provided for any items aside from health questionnaire items.

f. Blinding and randomization reported but methods not clearly defined.

g. Outcome assessor blinding not reported, blinding of patients not possible due to nature of intervention

Table 2: Inpatient Rehabilitation Model compared to Home Exercise Model

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Inpatient Rehabilitation Model	Home Exercise Model	Relative (95% CI)	Absolute (95% CI)		
Function: HAQ 15 months												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	32	28	-	MD **0.2 higher** (0.08 higher to 0.32 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Disease activity: DAS28 15 months												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	32	28	-	MD **0.5 higher** (0.4 higher to 0.6 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. High attrition bias, self reports and no outcome assessor blinding reported, large baseline differences

b. Small sample size

Table 3: Stress management compared to Support

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stress management	Support	Relative (95% CI)	Absolute (95% CI)		
Pain (1-15) 8 months												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b,c]	none	26	25	-	MD **0.2 higher** (2.31 lower to 2.71 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Functional disability (higher score means more disability) 8 months												
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	26	25	-	MD **0.12 higher** (0.25 lower to 0.49 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Function as inferred from Time to walk 50 feet (seconds) 8 months												
1	randomised trials	very serious[a]	not serious	serious[d]	very serious[b,c]	none	26	25	-	MD **0.3 lower** (2.81 lower to 2.21 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Function as inferred from Grip strength (mm Hg) 8 months												
1	randomised trials	very serious[a]	not serious	serious[d]	serious[b]	none	26	25	-	MD **15.6 lower** (31.42 lower to 0.22 higher)	⨁◯◯◯ Very low	Important No statistically significant difference
Disease activity as inferred from Morning stiffness (hours) 8 months												
1	randomised trials	very serious[a]	not serious	serious[d]	very serious[b,c]	none	26	25	-	MD **0.32 higher** (0.69 lower to 1.33 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference
Disease activity as inferred from ESR (mm/hour) 8 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stress management	Support	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a,b]	not serious	serious[d]	very serious[b,c]	none	26	25	-	MD **4.8 lower** (22.3 lower to 12.7 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. No patient or outcome assessor blinding reported, high attrition

b. Small sample size

c. Wide confidence interval

d. Surrogate measure

Table 4. Additional Data from RCT and Observational Studies

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
2034 Helewa 1991	Randomized controlled trial	6 weeks	105 Patients with RA, ages 18-70, with impairment of physical function	Patient-specific program of occupational therapy at home for 6 weeks	Functional score improved from baseline in the experimental group, while Beck depression scale and HAQ did not. There was a statistically significant improvement in pooled index (active joints, grip strength, ESR, morning stiffness, and functional change) at 6 weeks. When subjective measures were removed from the pooled index (so as to include only active joints, grip strength and ESR) the change did not reach statistical significance.

References:

1. Mathieux R, Marotte H, Battistini L, Sarrazin A, Berthier M, Miossec P. Early occupational therapy programme increases hand grip strength at 3 months: results from a randomised, blind, controlled study in early rheumatoid arthritis. *Ann Rheum Dis*. 2009;68(3):400-403. doi.org/10.1136/ard.2008.094532
2. Kraaimaat FW, Brons MR, Geenen R, Bijlsma JW. The effect of cognitive behavior therapy in patients with rheumatoid arthritis. *Behav Res Ther*. 1995;33(5):487-495.
3. Tonga E, Duger T, Karatas M. Effectiveness of Client-Centered Occupational Therapy in Patients With Rheumatoid Arthritis: Exploratory Randomized Controlled Trial. *Arch Rheumatol*. 2016;31(1):6-13. doi.org/10.5606/ArchRheumatol.2016.5478
4. Helewa A, Goldsmith CH, Lee P, et al. Effects of occupational therapy home service on patients with rheumatoid arthritis. *Lancet*. 1991;337(8755):1453-1456.
5. Macedo AM, Oakley SP, Panayi GS, Kirkham BW. Functional and work outcomes improve in patients with rheumatoid arthritis who receive targeted, comprehensive occupational therapy. Comment in (CIN). *Arthritis Rheumatol*. 2009;61(11):1522-1530. doi.org/10.1002/art.24563
6. Figen A, Gecene M, Gunduz R, Borman P, Yorgancioglu R. Long-term effects of comprehensive inpatient rehabilitation on function and disease activity in patients with chronic rheumatoid arthritis and ankylosing spondylitis. *Turk J Rheumatol*. 2011;26(2):135-144. doi.org/10.5152/tjr.2011.020
7. Hammond A, Young A, Kidao R. A randomised controlled trial of occupational therapy for people with early rheumatoid arthritis. *Ann Rheum Dis*. 2004;63(1):23-30.
8. Shearn MA, Fireman BH. Stress management and mutual support groups in rheumatoid arthritis. *Am J Med*. 1985;78(5):771-775.

PICO 18: Should patients with RA participate in a comprehensive physical therapy program?

Summary: This PICO question was addressed by 5 RCTs (1-5) and one non-randomized comparative study (6); they made five comparisons:

- Two RCTs: Comprehensive PT compared to Usual Care (1, 2)
- RCT: Comprehensive PT compared to conventional therapy (4)
- RCT: Health education compared to no education (5). Health education in this study was considered a type of comprehensive PT because it included medication, diet and exercise skills.
- RCT: Specially trained PT compared to Traditional PT (3)
- Non-randomized study: Community rehabilitation package that included water exercise, a self-help course, a stress management group and informal social activities compared to controls, who only attended orientation meeting (6)

For comparison to usual care, of the 14 outcomes included (Table 1), results favored comprehensive PT over usual care for 5 outcomes, favored usual care for overall quality of life as measured by the EuroQol, and were statistically non-significant for the other 8 outcomes.

When comparing comprehensive PT compared to conventional therapy (Table 2), all 4 outcomes favored conventional therapy. (4). We considered conventional therapy as different from usual care because it used physiotherapy measures for symptomatic treatments.

Comparing health education to no education, the results were in favor of health education (5).

In one RCT the outcomes for community rehabilitation program observational study were beneficial to rehabilitation programs than to controls (6).

In one RCT the outcomes for specially trained PT were more beneficial as compared to traditional PT, but the results are imprecise (3).

Overall Quality of Evidence: Very Low.

Table 1: Comprehensive PT compared to Usual Care (1, 2)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	Usual Care	Relative (95% CI)	Absolute (95% CI)		Statistical significance
Pain: VAS (0-100 scale), 6 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	Usual Care	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	not serious	serious[b]	none	35	31	-	MD **0.1 lower** (13.84 lower to 13.64 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference

Function as inferred from Performance-based test: 1-min sit to stand(number of complete rises), 6 months

1	randomised trials	not serious	not serious	serious[a]	serious[c]	None	35	31	-	MD **7.8 higher** (4.2 higher to 11.4 higher)	⊕⊕◯◯ Low	Critical **Statistically significant difference favoring the comprehensive PT group**

Function as inferred from Fatigue: BRAF-MDQ total (0 – 70 scale), 6 months

1	randomised trials	not serious	not serious	serious[a]	serious[c]	none	35	31	-	MD **6.2 lower** (12.26 lower to 0.14 lower)	⊕⊕◯◯ Low	Critical **Statistically significant difference favoring the comprehensive PT group**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	Usual Care	Relative (95% CI)	Absolute (95% CI)		Statistical significance
Function as inferred from Fatigue: VAS (0 – 100 scale), 6 months												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	35	31	-	MD **9.1 lower** (21.17 lower to 2.97 higher)	⊕⊕◯◯ Low	Critical No statistically significant difference
Pain: VAS (0-100 scale), 6 weeks												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	69	58	-	MD **0.2 lower** (9.16 lower to 8.76 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference
Disease Activity: DAS28 (score), 6 months												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	35	31	-	MD **0.3 lower** (0.66 lower to 0.06 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference
Mental Health: Stress VAS (0-100 scale), 6 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	Usual Care	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	not serious	serious[b]	none	35	31	-	MD **7.3 lower** (19.86 lower to 5.26 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference

Mental Health: HADS Anxiety (0 – 21 scale), 6 months

1	randomised trials	not serious	not serious	not serious	serious[c]	none	35	31	-	MD **1.5 lower** (2.66 lower to 0.34 lower)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring the comprehensive PT group**

Mental Health: HADS Depression (0-21 scale), 6 months

1	randomised trials	not serious	not serious	not serious	serious[b]	none	35	31	-	MD **1.3 lower** (2.68 lower to 0.08 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference

Quality of Life: EuroQol VAS (0-100 scale), 6 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	Usual Care	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	not serious	serious[c]	None	35	31	-	MD **13.2 higher** (3.65 higher to 22.75 higher)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring the usual care**

Self-efficacy: Arthritis Self-Efficacy Scale (10-100 scale), 6 weeks[a]

1	randomised trials	not serious	not serious	not serious	serious[b]	None	76	74	-	MD **5.1 higher** (0.76 lower to 10.96 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference

Self-efficacy: Arthritis Self-Efficacy Scale (10-100 scale), 6 months

1	randomised trials	not serious	not serious	not serious	serious[c]	none	35	31	-	MD **7.5 higher** (0.75 higher to 14.25 higher)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring the comprehensive PT group**

Disease activity: Tender joint count (number), 6 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	Usual Care	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	not serious	serious[b]	none	69	58	-	MD **0.9 lower** (5.31 lower to 3.51 higher)	⨁⨁⨁◯ Moderate	Important No statistically significant difference

Disease activity: Morning stiffness time (minutes), 6 weeks

1	randomised trials	not serious	not serious	not serious	serious[c]	none	69	58	-	MD **60.5 lower** (116.88 lower to 4.12 lower)	⨁⨁⨁◯ Moderate	Important **Statistically significant difference favoring the comprehensive PT group**

CI: confidence interval; **MD:** mean difference

Explanations

a. Indirect outcome

b. Wide CI crosses no-effect and significant effect lines

c. Number of patients in each group less than 200

Table 2: Comprehensive PT compared to conventional therapy (4)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	control	Relative (95% CI)	Absolute (95% CI)		
Functional Status Index - Mobility Assistance , 6 months												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	14	14	-	MD **1.71 higher** (0.8 higher to 2.63 higher)	⊕⊕◯◯ Low	Critical **Statistically significant difference favoring controls**
Functional Status Index - Mobility pain, 6 months												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	14	14	-	MD **2.0 higher** (1.5 higher to 2.5 higher)	⊕⊕◯◯ Low	Critical **Statistically significant difference favoring controls**
Functional Status Index Mobility difficulty, 6 months												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	14	14	-	MD **1.85 higher** (1.45 higher to 2.25 higher)	⊕⊕◯◯ Low	Critical **Statistically significant difference favoring controls**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comprehensive PT	control	Relative (95% CI)	Absolute (95% CI)		
EQ-5D, 6 months												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	14	14	-	MD **41.56 higher** (30.43 higher to 52.69 higher)	⨁⨁◯◯ Low	Important **Statistically significant difference favoring controls**

CI: confidence interval;

Explanations

a. No blinding involved

b. Less than 200 patients in each group

Table 3: Health education compared to no education (5)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Health education	no education	Relative (95% CI)	Absolute (95% CI)		
HAQ, 24 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Health education	no education	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[b]	none	46	46	-	MD **0** (0.3 lower to 0.3 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference

HAQ, 12 weeks

1	randomised trials	not serious	not serious	not serious	serious[b]	none	46	46	-	MD **0.12 lower** (0.48 lower to 0.24 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference

DAS28, 24 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	46	46	-	MD **0.87 lower** (1.55 lower to 0.19 lower)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring education group**

Self-efficacy, 24 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Health education	no education	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[a]	none	46	46	-	MD **12.17 higher** (5.31 higher to 19.03 higher)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring the education group**

DAS28, 12 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	46	46	-	MD **0.76 lower** (1.43 lower to 0.09 lower)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring the education group**

Self-efficacy, 12 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	46	46	-	MD **17 higher** (9.59 higher to 24.41 higher)	⊕⊕⊕◯ Moderate	Important **Statistically significant difference favoring the education group**

CI: confidence interval; **MD:** mean difference

Explanations

a. Less than 200 patients in each group

b. Less than 200 patients in each group and wide CI crosses no-effect and significant effect lines

CI: confidence interval; **MD:** mean difference

Explanations

a. Less than 200 patients in each group and wide CI crosses no-effect and significant effect lines

Table 4: Specially trained PT compared to Traditional PT (3)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Specially trained PT	Traditional PT	Relative (95% CI)	Absolute (95% CI)		
Change in functional capacity (0-20 score) at 4 months												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	54	36	-	MD **0.56 lower** (3.6 lower to 2.48 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Specially trained PT	Traditional PT	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from Grip strength (mm Hg), 4 months												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	54	36	-	MD **4.4 higher** (12.61 lower to 21.41 higher)	⊕⊕◯◯ Low	Critical No statistically significant difference
Tender joints (number), 4 months												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	54	36	-	MD **1.07 lower** (4.14 lower to 2 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference
Morning stiffness (min), 4 months												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	54	36	-	MD **16.7 lower** (48.84 lower to 15.44 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Less than 200 patients in each group, wide CI crosses no-effect and significant effect lines

Table 5: Observational study: Community rehabilitation compared to control (6)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Observational study: Community rehab	control	Relative (95% CI)	Absolute (95% CI)		
Pain (0-100 scale), 9 months												
1	observational studies	serious[a]	not serious	not serious	serious[b]	none	29	16	-	MD **0.39 higher** (12.72 lower to 13.5 higher)	⊕◯◯◯ Very low	Critical No statistically significant difference
Function as inferred from Fatigue (0-5 scale), 9 months												
1	observational studies	serious[a]	not serious	serious[c]	serious[b]	none	29	16	-	MD **0.14 lower** (0.7 lower to 0.42 higher)	⊕◯◯◯ Very low	Critical No statistically significant difference
Self-efficacy (perform self-management behaviors) (1-10 scale), 9 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Observational study: Community rehab	control	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	serious[a]	not serious	not serious	serious[b]	none	29	16	-	MD **0.89 higher** (0.01 higher to 1.77 higher)	⊕◯◯◯ Very low	Important **Statistically significant difference favoring community rehab**

Self efficacy (manage disease in general) (1-10 scale), 9 months

1	observational studies	serious[a]	not serious	not serious	serious[b]	none	29	16	-	MD **0.41 higher** (0.81 lower to 1.63 higher)	⊕◯◯◯ Very low	Important No statistically significant difference

Self-efficacy (active health outcome) (1-10 scale), 9 months

1	observational studies	serious[a]	not serious	not serious	serious[b]	none	29	16	-	MD **0.54 higher** (2.81 lower to 3.89 higher)	⊕◯◯◯ Very low	Important No statistically significant difference

Cognitive symptoms management (0-5), 9 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Observational study: Community rehab	control	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	serious[a]	not serious	not serious	serious[b]	none	29	16	-	MD **0.93 higher** (0.43 higher to 1.43 higher)	⊕◯◯◯ Very low	Important **Statistically significant difference favoring community rehab**

CI: confidence interval; MD**:** mean difference

Explanations

a. Non-randomized study

b. Less than 200 patients in each group

c. Indirect outcome

References:

1. Bell MJ, Lineker SC, Wilkins AL, Goldsmith CH, Badley EM. A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis. The Journal of rheumatology. 1998;25(2):231-7.
2. Feldthusen C, Dean E, Forsblad-d'Elia H, Mannerkorpi K. Effects of Person-Centered Physical Therapy on Fatigue-Related Variables in Persons With Rheumatoid Arthritis: A Randomized Controlled Trial. Archives of physical medicine and rehabilitation. 2016;97(1):26-36.
3. Helewa A, Smythe H, Goldsmith C. Can specially trained physiotherapists improve the care of patients with rheumatoid arthritis? A randomized health care trial. The Journal of rheumatology. 1994;21:70-9.

4. Shinde S, Varadharajulu G. Effect of Therapeutic Exercise Programme in Adults with Early Rheumatoid Arthritis. Indian Journal of Physiotherapy and Occupational Therapy - An International Journal. 2017;11:76.
5. Zhao S, Chen H. Effectiveness of health education by telephone follow-up on self-efficacy among discharged patients with rheumatoid arthritis: A randomised control trial. Journal of Clinical Nursing. 2019;28(21-22):3840-7.
6. Siu AM, Chui DY. Evaluation of a community rehabilitation service for people with rheumatoid arthritis. Patient education and counseling. 2004;55(1):62-9.

PICO 19: Should patients with rheumatoid arthritis (RA) use a standardized, evidence-based self-management program?

Literature searches identified 16 randomized controlled trials (RCTs) [1-16] and 5 non-randomized comparative studies [17-21] addressing this PICO question. Data on the critical outcomes for this PICO, functional status and pain, are summarized in Table 1. Table 2 includes all important outcomes (disease activity, quality of life, self-efficacy, work status, and mood). Table 3 provides results from the four studies (out of 21) with reporting that did not permit calculation of effect sizes (e.g., no dispersion reported).

All 21 studies were judged to make the same overall comparison: self-management program vs control/usual care.

Functional status was assessed by self-reported outcome measures, including the Health Assessment Questionnaire (HAQ), Disabilities of the Arm, Shoulder and Hand (DASH) [22], and the Activities Limitation Scales [23]. In addition, functional status was assessed by performance-based tests such as grip strength and timed eating/dressing, as well as fatigue as surrogate measures (Table 1).

Five RCTs evaluated self-management programs on functional status measured by the HAQ (0-3; lower = better outcome) in people with RA [2-5,11]. These studies included 731 participants (Intervention: 362; Control: 369). Combining the five studies, a small effect was found in the HAQ favouring self-management programs, with a mean difference (MD) of -0.13 units (95% CI: -0.28; 0.04) on the 0-3 scale at 6-24 months. The results correspond to a standardized mean difference (SMD) of -0.18 (95% CI: -0.4, 0.05), favoring the intervention. Below are the specific study results:

- Mayoux et al. [2] evaluated an 8-week group-education program consisting of information on RA, treatment, and lifestyle (total 6 hours). A small effect was found (MD in change-from-baseline: -0.1; 95% CI: -0.23, 0.03).
- Lorig et al. [11] used an online version of *Arthritis Self-Management Program (ASMP)* (MD in change-from-baseline: -0.03; 95% CI: -0.15, 0.09).
- Conn et al. [3] delivered the *ASMP* in-person to a primarily African American population (MD post-intervention: 0.07; 95% CI: -0.22, 0.36).
- Zhao et al. [5] studied a health education intervention delivered by a rheumatology nurse by telephone. While they found a significant improvement in self-efficacy in managing RA, minimal change was found in the HAQ (MD post-intervention: 0; 95% CI: -0.3, 0.3).
- In a 2021 study, Shao et al. [4] evaluated an 8-week individualized self-management program, consisting of peer storytelling, goal-setting, self-monitoring, and education on joint protection and physical activity. The Intervention Group (n = 112) also received a DVD on self-management techniques and a booklet to record goals, activities, symptoms and their thoughts on the program. The Control Group received usual care (n = 112). At 6-month follow up, a moderate effect was found in the Modified HAQ (20-80; lower = better outcome) favouring the Intervention Group (MD post-intervention: -3.31; 95% CI: -4.98, -1.64).

One study each assessed functional status using the DASH, the Activity Limitation Scale, and a self-reported measure of physical activity:

- Manning et al. [1] developed and studied the 12-week *EXTRA (Education, Self-Management, and Upper Extremity Exercise Training in People with Rheumatoid Arthritis)* program against usual care. EXTRA included four 1-hour group education, self-management, and global upper extremity exercise training sessions (in the first 2 weeks) supplementing the individualized home exercise regimen. They found a small effect on the DASH (primary outcome) favouring the intervention (MD in change-from-baseline: -2.2 points on a 0-100 scale; 95% CI: -11.18, 6.78).
- Lorig et al. [11] included the Activity Limitation Scale as a secondary outcome and found a moderate effect favouring the intervention (MD in change-from-baseline: -0.5 points on a 0-4 scale; 95% CI: -0.79, -0.21).
- Mayoux et al. [2] found a small effect in a self-reported measure of physical activity (Baecke Questionnaire) favouring the intervention (MD in change-from-baseline: 1.49 points on a 2-10 scale; 95% CI: -0.42, 3.4).

Five RCTs assessed fatigue as a surrogate for functional status [1,2,7,11,15]. These studies included 630 participants (Intervention: 301; Control: 329). Pooling data from the five studies, a small effect was found in the VAS for fatigue favouring the intervention, with a MD of 0.22 units (95% CI: - 0.58; 0.17) on a 0-10 scale at 6-12 months. The result correspond to a SMD of -0.09 (95% CI: -0.24, 0.07) favouring the intervention. In addition, Manning et al. [1] included performance-based measures as a surrogate for assessing functional status. Small effects were found in those measures favouring the intervention (Table 1).

Pain was measured by eight studies (Table 1) with small effects found in six [1,2,7,11,14,16]. In a follow-up of a pilot study on the 3-month *Learning about RA* program, Neuberger et al. [13] found a moderate effect favouring the intervention (SMD: -0.54; 95% CI: -1.36, 0.28). Yousefi et al. [8] compared an 8-week small group education program (n = 100) with usual care (n = 106). Topics included the disease, pain management and a variety of non-pharmacological treatments. Compared to usual care, the intervention group had a mean of 13 points lower (95% CI: 15.77, 10.23) in pain measured in a visual analogue scale (0 – 100). Combining the eight studies, a small effect was found in the pain measure favouring the intervention (SMD: -0.36; 95% CI: -0.73, 0.02). This corresponds to a MD of 0.92 units (95% CI: 1.86 lower; 0.05 higher) on a 10-point pain VAS.

The RCTs assessing functional status were rated as low quality evidence, except for the Activity Limitation Scale and self-reported physical activity – each included one study (Table 1). The RCTs assessing pain were also rated as low quality evidence.

Quality of evidence across all critical outcomes: Low.

Table 1. Data from Randomized Controlled Trials – Critical Outcomes

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Self-Management Program	Control / Usual Care	MD (95% CI)	SMD (95% CI)		
Functional Status: HAQ (0-3; lower = better outcome) – 6 months to 24 months												
5	randomised trials	not serious	serious[a]	not serious	serious[b]	none	362	369	MD **0.13 lower** (0.28 lower to 0.04 higher)	SMD **0.18 lower** (0.4 lower to 0.05 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
Functional status: DASH (0-100; lower = better outcome) – 36 weeks (8.3 months)												
1	randomised trials	not serious[c]	not serious	not serious	very serious[f]	none	52	56	MD **2.2 lower** (11.18 lower to 6.78 higher)		⨁⨁◯◯ Low	CRITICAL No significant difference
Functional Status - Activity Limitation Scale (0-4; lower = better outcome) – change score at 6 months												
1	randomised trials	not serious	not serious	not serious	not serious	None	72	72	MD **0.5 lower** (0.79 lower to 0.21 lower)		⨁⨁⨁⨁ High	CRITICAL **Significant difference in favor of self-management program**
Functional status inferred from timed eating (minutes; lower = better outcome) – 36 weeks (8.3 months)												
1	randomised trials	not serious[c]	not serious	serious[d]	serious[b]	None	52	56	MD **0.5 lower** (1.68 lower to 0.6 higher)		⨁⨁◯◯ Low	CRITICAL No significant difference
Functional status inferred from timed dressing (minutes; lower = better outcome) – 36 weeks (8.3 months)												
1	randomised trials	not serious[c]	not serious	serious[d]	serious[b]	None	52	56	MD **1.9 lower** (5.07 lower to 1.27 higher)		⨁⨁◯◯ Low	CRITICAL No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Self-Management Program	Control / Usual Care	MD (95% CI)	SMD (95% CI)		
Functional status inferred from handgrip (dominant side) measured by a hand-grip dynamometer (peak force [N]; higher = better) – 36 weeks (8.3 months)												
1	randomised trials	not serious[c]	not serious	serious[d]	serious[b]	none	52	56	MD **22.6 higher** (22.66 lower to 67.86 higher)		⨁⨁◯◯ Low	CRITICAL No significant difference
Functional status inferred from grip ability test (seconds; lower = better outcome) – 36 weeks (8.3 months)												
1	randomised trials	not serious[c]	not serious	serious[d]	serious[b]	none	52	56	MD **2.2 lower** (5.59 lower to 1.19 higher)		⨁⨁◯◯ Low	CRITICAL No significant difference
Functional status inferred from fatigue measured by Visual Analogue Scale (0-10; lower = better outcome) – 6 months to 12 months												
5	randomised trials	not serious	not serious	serious[d]	serious[b]	none	301	329	MD **0.22 lower** (0.58 lower to 0.17 higher)	SMD **0.09 lower** (0.24 lower to 0.07 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
Pain measured by Visual Analogue Scale (0-10; lower = better outcome) – 16 weeks (3.7 months) – 15 months												
8	randomised trials	serious[g]	serious[a]	not serious	not serious	none	442	449	MD **0.92 lower** (1.86 lower to 0.05 higher)	SMD **0.36 SD lower** (0.73 lower to 0.02 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

CI: confidence interval; **SMD:** standardised mean difference

Explanations

a. Inconsistency in direction and magnitude of effect
b. The CI overlaps "0" but includes the possibility of benefit favoring the Intervention.
c. One study - participants were not blinded.
d. Surrogate measure

f. The CI overlaps "0" and includes the possibility of benefit and harm.
g. Four of the 8 studies have at least 1 high ROB. In addition, 3 of the studies have at least 2 unclear ROB.

Table 2. Data from Randomized Controlled Trials – Important Outcomes

Certainty assessment							№ of patients		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Self-Management Program	Control / Usual Care	MD (95% CI)	SMD (95% CI)	Certainty	Importance
Disease Activity: DAS28 (0-28; lower = better outcome) – 6 months to 24 months												
5	randomised trials	serious[a]	serious	not serious	not serious	None	320	327	MD **2.3 lower** (0.63 lower to 0.02 lower)	SMD **0.17 lower** (0.32 lower to 0.01 lower)	⨁⨁◯◯ Low	IMPORTANT **Significant difference in favor of self-management program**
Disease Activity inferred from Total Number of Swollen Joints (lower = better outcome) – 18 months												
1	randomised trials	not serious	not serious	serious	serious[b]	None	34	41	MD **0.9 lower** (2.5 lower to 0.7 higher)		⨁⨁◯◯ Low	IMPORTANT No significant difference
Disease Activity inferred from Total Number of Tender Joints (lower = better outcome) – 18 months												
1	randomised trials	not serious	not serious	serious	serious[b]	None	34	41	MD **0.9 lower** (3.57 lower to 1.77 higher)		⨁⨁◯◯ Low	IMPORTANT No significant difference
Quality of Life – SF-36 Physical (0-100; higher = better outcome) – 6 months to 15 months												
3	randomised trials	not serious	serious[c]	not serious	serious[b]	none	243	261	MD **11.81 higher** (13.31 lower to 36.92 higher)	SMD **0.63 higher** (0.71 lower to 1.97 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Self-Management Program	Control / Usual Care	MD (95% CI)	SMD (95% CI)		
Quality of Life – SF-36 Mental (0-100; higher = better outcome) – 6 months to 15 months												
3	randomised trials	not serious	serious[c]	not serious	very serious[d]	none	243	261	MD **8.09 higher** (9.33 lower to 25.63 higher)	SMD **0.65 higher** (0.75 lower to 2.06 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference
Quality of Life – RAQoL (0-30; lower = better outcome) - 36 weeks												
1	randomised trials	not serious	not serious	not serious	very serious[e]	none	52	56	MD **0.6 higher** (2.03 lower to 3.23 higher)		⨁⨁◯◯ Low	IMPORTANT No significant difference
Quality of Life Scale – QLS [24] (16-112; higher = better outcome) - 41 weeks												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	43	45	MD **4.9 higher** (0.96 lower to 10.76 higher)		⨁⨁⨁◯ Moderate	IMPORTANT No significant difference
Quality of Life: Health Distress Scale (10-50; lower = better outcome) – 6 months												
1	randomised trials	not serious	not serious	not serious	not serious	none	72	72	MD **0.45 lower** (0.78 lower to 0.13 lower)		⨁⨁⨁⨁ High	IMPORTANT **Significant difference in favor of self-management program**
Self-efficacy – Pain (10-100, higher = better outcome) – 4 months to 9 months												
5	randomised trials	not serious	not serious	not serious	serious[b]	none	268	258	MD **4.86 higher** (10.25 lower to 19.86 higher)	SMD **0.28 higher** (0.59 lower to 1.14 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Self-Management Program	Control / Usual Care	MD (95% CI)	SMD (95% CI)		
Self-efficacy - RA/Symptoms (10-100, higher = better outcome) – 6 months to 24 months												
6	randomised trials	not serious	not serious	not serious	not serious	none	373	401	MD **5.35 higher** (0.32 higher to 10.22 higher)	SMD **0.33 higher** (0.02 higher to 0.63 higher)	⊕⊕⊕⊕ High	IMPORTANT **Significant difference in favor of self-management program**
Self-efficacy - Function (10-100, higher = better outcome) – 4 months – 36 weeks (8.3 months)												
2	randomised trials	not serious	not serious	not serious	serious[b]	none	89	95	MD **14.16 higher** (4.79 lower to 33.1 higher)	SMD **0.71 higher** (0.24 lower to 1.66 higher)	⊕⊕⊕◯ Moderate	IMPORTANT No significant difference
Work status measured by EMIR (French version of AIMS2; 0-10; lower = better outcome) - 12 months												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	79	72	MD **0.2 higher** (0.63 lower to 1.03 higher)		⊕⊕⊕◯ Moderate	IMPORTANT No significant difference
Depression measured by CES-D (0-60; lower = better outcome) – 16 weeks to 12 months												
4	randomised trials	not serious	serious[c]	not serious	serious[b]	none	186	186	MD **0.92 lower** (3.93 lower to 2.08 higher)	SMD **0.08 lower** (0.34 lower to 0.18 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Anxiety measured by the STAI Anxiety Scale (0-80; lower = better outcome) – 6 months to 12 months												
2	randomised trials	not serious	serious[c]	not serious	very serious[e]	none	130	130	MD **1.52 higher** (2.41 lower to 5.45 higher)	SMD **0.24 higher** (0.38 lower to 0.86 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

CI: confidence interval; **SMD:** standardised mean difference

Explanations

a. Two of the 6 studies have 1 item rated high risk of bias (ROB) and 1 study has 4 items.
b. The CI overlaps "0" but includes the possibility of benefit favoring the Intervention.
c. Inconsistency in direction and magnitude of effect
d. The CI overlaps "0" but includes the possibility of benefit favoring the Intervention. Wide CI.
e. The CI overlaps "0" and includes the possibility of benefit and harm.
f. The CI overlaps "0" and includes the possibility of benefit and harm.

Table 3. Additional Data from RCT and Non-randomized Studies

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
8 Helliwell 1999	RCT	4 week education program, final assessment at 12 months	**Education Group** (n = 34,mean age = 55 yrs, 62.8 % female, disease duration 3 yrs) **Control** (n = 43, mean age = 56.5 yrs, 70.5 % female, disease duration 3.5 yrs)	**Patient education** consisted of four 2 hr sessions covering pathophysiology of RA, medications, local treatments, pain, stress, exercise, rest, joint protection, task allocation, splinting and assistive devices Control – no education but would be eligible at end of study if classes found to be of benefit	Groups similar in demographic and baseline info. Education group at 12 months (median and ranges) Larsen 39.5 (1-92) HAQ 0.875 (0-2.125) RAI 7 (0-20) SF-36 physical function 45 (0-95) SF-36 Mental function 76 (32-100) Control group at 12 months (median and ranges) Larsen 43 (5-101) HAQ 1.0 (0-2.75) RAI 6.5 (0-20) SF-36 physical function 42.5 (5-95) SF-36 Mental function 80 (16-100)
2747, Taal, 1993	field-experimental design with experiment	14 months	75 RA patients	Group education program for RA patients consisting of 5 weekly 2-hour sessions with 6-8 patients (partners were invited as well). Groups were led by professionals with expertise on	Mean change scores at 14 months in control (n=30) and experimental group (n=27) **Health Status** Physical activities: C -0.48, E -0.16

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
	al and control groups			rheumatic diseases or leading groups. Leaders received 2 days of training and a manual. Patient received a packet with a workbook, self-help guide, education on RA, and audiotape with relaxation exercises. Program included: contracting/goal setting, self-management and problem-solving, information on RA and treatment, pain management and relaxation, physical exercises, communication skills, and coping with depression. Participants in the control group did not receive information or materials.	Dexterity: C 0.00, E -0.16 Household activities: C -0.26, E 0.00 Pain: C -0.33, E -0.02 Depression: C -0.60, E -0.25 Anxiety: C -0.26, -0.54 Social activities: C -0.47, E -0.06 Arthritis impact: C -0.25, E -0.47 Disability (M-HAQ): C 0.15, E 0.09 Joint tenderness: C 1.63, E 0.77 **Lab Tests** ESR: C 9.50, E 3.58 **Behavior** Relaxation: C 0.00, E 0.74 Physical exercises: C -2.12, 1.91 **(p<.001)** Endurance exercises: C 0.27, E 0.59 Self-management activities: C 0.07, E 0.23 **Outcome Expectations Self-efficacy**: C 0.08, E 0.20 Self-efficacy pain: C 0.15, E 0.33 Self-efficacy function: C -0.06, E 0.17 (**p<.05)** Self-efficacy other symptoms: C 0.11, E 0.10
4946 Lindroth 1997	RCT	3-month education program. Assessment at 3 and 12 months	100 participants (12 men, 84 women); 27 - 77 years old **Intervention:** n = 49 (male/female : 5/44; age 54 [SD 15] years) **Control:** n = 47 (male/female : 7/40; age 56 [SD 12] years)	**Intervention Group:** Received Rheumatoid Arthritis School. Patients receives a handbook. During 8 sessions, 2.5 hours once a week, group discussions were led by a multi-disciplinary team. Each group consisted of 5 to 7 patients with RA. **Control Group:** Waiting list	Groups similar in demographic and baseline info. **Intervention Group** at 12 months - mean Pain (mm on VAS) 47.8 HAQ 1.3 Depressed feelings (# reported) 10 **Control Group** at 12 months - mean Pain (mm on VAS) 47.2 HAQ 1.1 Depressed feelings (# reported) 16

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
2267 Westone 1985	Randomized controlled trial	24-34 days	36 patients with RA	Computer-based education, delivered as a series of case studies, factual data, directed advice, with accompanying multiple choice questions	Results are reported as the number of patients who had increased belief, decreased belief, and no change in belief. Patients receiving computer-based therapy showed less belief that chance plays a role in their health. P<0.05. Please see table.

		Control			Computer-based	
Scale	Increased	Decreased	No Change	Increased	Decreased	No change
Internal	5	12	0	9	9	0
Powerful Others	9	7	1	8	9	1
Chance	10	6	1	3	13	2

References

1. Manning VL, Hurley MV, Scott DL, Coker B, Choy E, Bearne LM (2014) Education, self-management, and upper extremity exercise training in people with rheumatoid arthritis: a randomized controlled trial. Arthritis Care Res (Hoboken) 66: 217-227.

2. Giraudet-Le Quintrec JS, Mayoux-Benhamou A, Ravaud P, Champion K, Dernis E, Zerkak D, Ouslimani A, Courpied JP, Revel M, Kahan A, Dougados M (2007) Effect of a collective educational program for patients with rheumatoid arthritis: a prospective 12-month randomized controlled trial. J Rheumatol 34: 1684-1691.

3. Conn DL, Pan Y, Easley KA, Comeau DL, Carlone JP, Culler SD, Tiliakos A (2013) The effect of the Arthritis Self-Management Program on outcome in African Americans with rheumatoid arthritis served by a public hospital. Clin Rheumatol 32: 49-59.

4. Shao JH, Yu KH, Chen SH (2021) Effectiveness of a self-management program for joint protection and physical activity in patients with rheumatoid arthritis: A randomized controlled trial. Int J Nurs Stud 116: 103752.

5. Zhao S, Chen H (2019) Effectiveness of health education by telephone follow-up on self-efficacy among discharged patients with rheumatoid arthritis: A randomised control trial. J Clin Nurs 28: 3840-3847.

6. Anvar N, Matlabi H, Safaiyan A, Allahverdipour H, Kolahi S (2018) Effectiveness of self-management program on arthritis symptoms among older women: A randomized controlled trial study. Health Care Women Int 39: 1326-1339.

7. Zuidema R, van Dulmen S, Nijhuis-van der Sanden M, Meek I, van den Ende C, Fransen J, van Gaal B (2019) Efficacy of a Web-Based Self-Management Enhancing Program for Patients with Rheumatoid Arthritis: Explorative Randomized Controlled Trial. J Med Internet Res 21: e12463.

8. Yousefi H, Chopra A, Farrokhseresht R, Sarmukaddam S, Noghabi FA, Bedekar N, Madani A (2015) Epidemiological evaluation quality of life in patients suffering from early rheumatoid arthritis: a pragmatic, prospective, randomized, blind allocation controlled of a modular program group intervention. Epidemiol Health 37: e2015048.

9. Wetstone SL, Sheehan TJ, Votaw RG, Peterson MG, Rothfield N (1985) Evaluation of a computer based education lesson for patients with rheumatoid arthritis. J Rheumatol 12: 907-912.

10. Taal E, Riemsma RP, Brus HL, Seydel ER, Rasker JJ, Wiegman O (1993) Group education for patients with rheumatoid arthritis. Patient Educ Couns 20: 177-187.

11. Lorig KR, Ritter PL, Laurent DD, Plant K (2008) The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 59: 1009-1017.

12. Lindroth Y, Brattstrom M, Bellman I, Ekestaf G, Olofsson Y, Strombeck B, Stenshed B, Wikstrom I, Nilsson JA, Wollheim FA (1997) A problem-based education program for patients with rheumatoid arthritis: evaluation after three and twelve months. Arthritis Care Res 10: 325-332.

13. Neuberger GB, Smith KV, Black SO, Hassanein R (1993) Promoting self-care in clients with arthritis. Arthritis Care Res 6: 141-148.

14. Shigaki CL, Smarr KL, Siva C, Ge B, Musser D, Johnson R (2013) RAHelp: An Online Intervention for Individuals With Rheumatoid Arthritis. Arthritis Care & Research 65: 1573-1581.

15. Pot-Vaucel M, Aubert MP, Guillot P, Glémarec J, Berthelot JM, Le Goff B, Maugars Y (2016) Randomised study versus control group of customised therapeutic education for patients in follow-up for rheumatoid arthritis. Joint Bone Spine 83: 199-206.

16. Kirwan JR, Hewlett S, Cockshott Z, Barrett J (2005) Clinical and psychological outcomes of patient education in rheumatoid arthritis. Musculoskeletal Care 3: 1-16.

17. Lindroth Y, Bauman A, Barnes C, McCredie M, Brooks PM (1989) A controlled evaluation of arthritis education. Br J Rheumatol 28: 7-12.

18. Lindroth Y, Bauman A, Brooks PM, Priestley D (1995) A 5-year follow-up of a controlled trial of an arthritis education programme. British Journal of Rheumatology 34: 647-652.

19. Shao JH, Yu KH, Chen SH (2020) Feasibility and Acceptability of a Self-Management Program for Patients With Rheumatoid Arthritis. Orthop Nurs 39: 238-245.

20. Mollard E, Michaud K (2018) A Mobile App With Optical Imaging for the Self-Management of Hand Rheumatoid Arthritis: Pilot Study. JMIR Mhealth Uhealth 6: e12221.

21. Oh H, Seo W (2003) Decreasing pain and depression in a health promotion program for people with rheumatoid arthritis. J Nurs Scholarsh 35: 127-132.

22. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C (2001) Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther 14: 128-146.

23. Lorig K, Stewart A, Ritter P, Gonzalez VM, Laurent D, Lynch J (1996) Outcome measures for health education and other health care interventions. Thousand Oaks: Sage Publications.

24. Burckhardt CS, Anderson KL (2003) The Quality of Life Scale (QOLS): Reliability, Validity, and Utilization. Health and Quality of Life Outcomes 1: 60.

PICO 20: Should patients with RA use mind-body approaches?

Summary: We included 23 studies for this PICO; all were randomized controlled trials (RCTs) (1-23). In total, there were 15 different comparisons involving 12 mind-body approaches:

- Cognitive behavioral therapy (CBT) vs control (Table 1)
- Meditation vs control (Table 2)
- CBT vs Meditation (Table 3)
- Mindfulness vs control (Table 4)
- Progressive muscle relaxation vs control (Table 5)
- Standard group therapy (SGT) vs control (Table 6)
- Yoga vs control (Table 7)
- Whole body vibration (WBV) vs control (Table 8)
- CBT vs arthritis education (Table 9)
- Relaxation response vs arthritis education (Table 10)
- Behavioral therapy with family support vs behavioral therapy (Table 11)
- Behavioral therapy with family support vs control (Table 12)
- Stress management vs support (Table 13)
- Motivational interviewing vs control (Table 14)
- CBT vs SGT (Table 15)

The most common comparison was cognitive behavioral therapy (CBT) versus control (5, 8-10, 13, 18-21, 23). The other comparisons comprised only 1-2 studies each.

For CBT, in most studies, CBT was statistically significantly protective against depression, anxiety, and fatigue (after a more than 12-week follow-up) versus the control group (5, 8, 10, 13, 18, 20). CBT also was also borderline associated with very good sleep quality (8, 9). For other outcomes (pain levels, disease activity, mobility, disability, AIMS Physical Functioning score, self-efficacy, and quality of life), there were no statistically significant differences between CBT groups and control groups.

In addition to CBT, some other mind-body interventions improved some outcomes at follow-up. Below, we list which comparisons/outcomes had at least one statistically significant difference; in all cases there were additional outcomes with statistically nonsignificant differences.

- Mindfulness was statistically significantly associated with higher well-being, lower depression, lower disease activity, and lower anxiety when compared with the control group at 12 or more weeks (3, 16). .

- Progressive muscle relaxation was statistically significantly associated with higher sleep quality and lower fatigue versus the control group at <12 weeks (11).
- When comparing those who received the standard group therapy (SGT) with the control group we found that at 12 weeks or more follow-up, SGT was associated with a statistically significantly lower pain behavior score, disease activity (Rheumatoid Activity Index), and anxiety (2) as compared to a control group.
- Grip strength was statistically significantly higher at follow-up (12 or more weeks) in the Yoga group versus the control group (4).
- Whole Body Vibration (WBV) Therapy was statistically significantly associated with lower disability versus the control group at follow-up (12 weeks or more). (17)
- Self-efficacy was higher and functional status (health assessment questionnaire (HAQ) for disability) was lower in those who received the motivational interviewing/self-regulation at follow-up (12 weeks or more) compared with those who did not (12).

CBT performed statistically significantly better than SGT for 3 measures of disease activity (Rheumatologist or Nurse Assessment of Disease Activity, articular index, and rheumatoid activity index) (2).

For other comparisons, no outcomes were statistically significant, specifically comparing meditation to a control group, meditation to CBT, CBT to arthritis education, relaxation response to arthritis education, behavioral therapy with family support to behavioral therapy, behavioral therapy with family support versus control, stress management versus support (1, 3, 6, 16, 18, 22, 23). Many of these comparisons only involved single studies.

Quality of evidence across all critical outcomes:

- Low for CBT, meditation, progressive muscle relaxation, SGT, WBV, relaxation response, family support for behavioral therapy, and behavioral therapy with family support.
- Very low for yoga, stress management, and motivational interviewing.

Table 1: Cognitive behavioral therapy (2, 4, 5, 8-10, 13, 18-21, 23)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
Pain, 12 weeks or more (9 studies follow-up ranged from 18 weeks to 24 months) (0-10 scale where lower is better)												
9	randomised trials	not serious	serious[a]	not serious	serious[b]	none	457	445	-	MD **0.35 lower** (0.93 lower to 0.22 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
Fatigue, 12 weeks or more (5 studies follow-up ranged from 18 weeks to 24 months)												
5	randomised trials	not serious	serious[a]	serious[d]	not serious	none	402	403	-	MD **1.85 lower** (2.71 lower to 0.99 lower)	⨁⨁◯◯ Low	CRITICAL **Significant difference favoring CBT**
Depression, 12 weeks or more (7 studies follow-up ranged from 18 weeks to 24 months)												
7	randomised trials	not serious	serious[a]	not serious	not serious	none	421	409	-	MD **1.24 lower** (2.1 lower to 0.43 lower)	⨁⨁⨁◯ Moderate	IMPORTANT **Significant difference favoring CBT versus Control**

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
Anxiety, 12 weeks or more (5 studies follow-up ranged from 18 weeks to 24 months)												
5	randomised trials	not serious	serious[a]	not serious	not serious	none	367	350	-	MD **0.93 lower** (1.63 lower to 0.28 lower)	⨁⨁⨁◯ Moderate	IMPORTANT **Significant difference favoring CBT versus Control**
Disease Activity (ESR), 12 weeks or more (2 studies follow-up ranged from 6 to 18 months)												
2	randomised trials	not serious	serious[a]	serious[c]	serious[b]	none	44	44	-	MD **5.17 lower** (11.2 lower to 0.86 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference
Disease Activity (CRP), 12 weeks or more (18 months)												
1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	27	26	-	MD **5.2 lower** (18.68 lower to 8.28 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
Functional Status (Mobility), 12 weeks or more (2 studies follow-up ranged from 6 to 12 months)												
2	randomised trials	not serious	serious[a]	serious[c]	serious[b]	none	51	59	-	MD **0.99 higher** (1.37 lower to 3.34 higher)	⨁◯◯◯ Very low	IMPORTANT No significant difference
Disease Activity, 12 weeks or more (5 studies follow-up ranged from 18 weeks to 24 months)												
5	randomised trials	not serious	serious[a]	not serious	serious[b]	none	375	374	-	MD **0.09 lower** (0.29 lower to 0.11 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference
Quality of life, 12 weeks or more (4 studies follow-up ranged from 18 weeks to 24 months)												
4	randomised trials	not serious	serious[a]	not serious	serious[b]	none	357	335	-	MD **1.29 lower** (3.29 lower to 0.71 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
Self-efficacy, 12 weeks or more (3 studies follow-up ranged from 18 weeks to 24 months)												
3	randomised trials	not serious	not serious	not serious	serious[b]	none	325	307	-	MD **1.06 higher** (1.63 lower to 3.75 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No significant difference
Sleep (Very good quality), 12 weeks or more (2 studies follow-up ranged from 18 weeks to 24 months)												
2	randomised trials	not serious	serious[a]	serious[c]	not serious	none	30/281 (10.7%)	16/263 (6.1%)	not estimable	**50 fewer per 1,000** (from 90 fewer to 0 fewer)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring CBT**
Disability, 12 weeks or more (5 studies follow-up ranged from 18 weeks to 24 months)												
5	randomised trials	not serious	not serious	not serious	serious[b]	none	375	354	-	MD **0.1 lower** (0.25 lower to 0.05 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No significant difference
Functional Status (AIMS Physical Functioning), 12 weeks or more (2 months) (scale range not reported)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	14	15	-	MD **0.65 lower** (7.99 lower to 6.69 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Disease Activity (Joint exam swelling severity), 12 weeks or more (2 months)

1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	14	15	-	MD **12.37 lower** (29.31 lower to 4.57 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Disease Activity (Joint exam # of swollen joints 2 months), 12 weeks or more (2 months)

1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	14	15	-	MD **7.49 lower** (17.95 lower to 2.97 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Pain, <12 weeks (8 weeks) (0-100 scale where lower is better)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	34	28	-	MD **3.80 higher** (-18.97 lower to 26.57 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Depression, <12 weeks (8 weeks) (1-5 scale where lower is better)

1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	17	14	-	MD **0.15 higher** (0.46 lower to 0.76 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Disease Activity (joint swelling), <12 weeks (8 weeks) (0-84 scale where lower is better)

1	randomised trials	not serious	not serious	serious[c]	serious[b]	none	17	14	-	MD **2.84 higher** (2.9 lower to 8.58 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Disease Activity (joint tenderness), <12 weeks (8 weeks) (0-84 scale where lower is better)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[c]	Very serious[b]	none	17	14	-	MD **10.96 higher** (4.85 lower to 26.77 higher)	⨁◯◯◯ Very Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. High heterogeneity

b. Wide confidence intervals

c. Enrolled patients are not typical

d. Indirect measure of the critical outcome of functional status

Table 2. Meditation compared to control (23)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Meditation	Control	Relative (95% CI)	Absolute (95% CI)		
Pain, <12 weeks (8 weeks) (0-100 scale where lower is better)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Meditation	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	30	-	MD **4.16 lower** (13.04 lower to 4.72 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Depression, < 12 weeks (8 weeks) (1-5 scale where lower is better)

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	30	-	MD **0.05 lower** (0.48 lower to 0.38 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Disease Activity (Joint Swelling), <12 weeks (8 weeks) (0-84 scale where lower is better)

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	30	-	MD **1.89 higher** (2.06 lower to 5.84 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Disease Activity (Joint Tenderness), <12 weeks (8 weeks) (0-84 scale where lower is better)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Meditation	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	30	-	MD **1.66 higher** (9.22 lower to 12.54 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 3. Meditation compared to CBT. (23)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Meditation	CBT	Relative (95% CI)	Absolute (95% CI)		
Pain, <12 weeks (8 weeks) (0-100 scale where lower is better)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Meditation	CBT	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	35	-	MD **6.96 lower** (15.54 lower to 1.62 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Disease Activity (Joint Swelling),<12 weeks (8 weeks) (0-84 scale where lower is better)

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	35	-	MD **0.14 lower** (3.97 lower to 3.69 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Disease Activity (Joint tenderness), <12 weeks (8 weeks) (0-84 scale where lower is better)

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	41	35	-	MD **4.24 higher** (6.27 lower to 14.75 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 4. Mindfulness compared to control (3, 16)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness	Control	Relative (95% CI)	Absolute (95% CI)		
Depression, 12 weeks or more (both studies followed participants for 6 months)												
2	randomised trials	not serious	not serious	serious[a]	serious[b]	none	45	46	-	MD **0.86 lower** (1.60 lower to 0.13 lower)	⊕⊕◯◯ Low	IMPORTANT **Significant difference favoring mindfulness**
Well-Being, 12 weeks or more (6 months) (scale range 42-252 where higher is better)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	31	32	-	MD **11.02 higher** (1.57 higher to 20.47 higher)	⊕⊕◯◯ Low	IMPORTANT **Significant difference favoring mindfulness**
Disease Activity (DAS 28), 12 weeks or more (6 months)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	31	32	-	MD **0.18 lower** (0.64 lower to 0.28 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Disease Activity (Tender Joint Count) 12 weeks or more (6 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	not serious	none	21	21	-	MD **2.9 lower** (3.57 lower to 2.23 lower)	⊕⊕⊕◯ Moderate	IMPORTANT **Significant difference favoring mindfulness**

Disease Activity (Change Swollen Joint Count), 12 weeks or more) (6 months)

1	randomised trials	not serious	not serious	serious[a]	not serious	none	21	21	-	MD **0.96 higher** (0.51 higher to 1.41 higher)	⊕⊕⊕◯ Moderate	IMPORTANT **Significant difference favoring control**

Anxiety, 12 weeks or more (6 months) (scale range 0-42 where lower is better)

1	randomised trials	not serious	not serious	serious[a]	not serious	none	14	14	-	MD **3.43 lower** (5.38 lower to 1.48 lower)	⊕⊕⊕◯ Moderate	IMPORTANT **Significant difference favoring mindfulness**

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 5. Progressive muscle relaxation compared to control. (11)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Progressive muscle relaxation	control	Relative (95% CI)	Absolute (95% CI)		
Fatigue (Total Fatigue Severity Scale), <12 weeks (6 weeks) Assessed using Fatigue Severity Scale. Higher score = more fatigue. pathological fatigue = score of 4 and above												
1	randomised trials	not serious	not serious	very serious[b]	not serious	none	35	37	-	MD **4.42 lower** (5.01 lower to 3.83 lower)	⊕⊕◯◯ Low	CRITICAL **Significant difference favoring progressive muscle relaxation**
Sleep Quality (PSQI), <12 weeks (6 weeks) Assessed using Pittsburgh Sleep Quality Index. Score ranges from 0 to 21, 21=poor sleep quality												
1	randomised trials	not serious	not serious	serious[a]	not serious	none	35	37	-	MD **7.17 lower** (8.8 lower to 5.54 lower)	⊕⊕⊕◯ Moderate	IMPORTANT **Significant difference favoring progressive muscle relaxation**

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Enrolled patients are not typical, or the outcome is an indirect measure of the critical outcome of functional status

Table 6. SGT compared to control. (2)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SGT	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Pain Behavior Score), 12 weeks or more (6 months) (scale range unclear, because authors adjusted for pretreatment scores, but lower scores are better)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **5.78 lower** (10.6 lower to 0.96 lower)	⨁⨁◯◯ Low	CRITICAL **Significant difference favoring SGT**
Pain (Pain Intensity Rating), 12 weeks or more (6 months) assessed using 10-cm visual analog scale ratings. The score ranges from 0 to 10, 0= no pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **0.95 higher** (0.13 lower to 2.03 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SGT	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Pain unpleasantness index), 12 weeks or more (6 months) assessed using 10-cm visual analog scale ratings. The score ranges from 0 to 10, 0= no pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **0.31 higher** (1.09 lower to 1.71 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Disease Activity (Rheumatoid Activity Index), 12 weeks or more (6 months) Range from 0–10, 0=no disease activity												
1	randomised trials	not serious	not serious	serious[a]	serious	none	18	18	-	MD **80.54 lower** (93.13 lower to 67.95 lower)	⊕⊕◯◯ Low	IMPORTANT **Significant difference favoring SGT**
Disease Activity (Rheumatologist or nurse assessment of disease activity), 12 weeks or more (6 months),scale range unclear, because authors adjusted for pretreatment scores, but lower scores are better												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **2.52 higher** (3.58 lower to 8.62 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Disease Activity (Articular Index), 12 weeks or more (6 months) Assessed using number of tender joints												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SGT	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **0.45 higher** (3.32 lower to 4.22 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Functional Status (Grip Strength), 12 weeks or more (6 months) (nurse evaluation, scale range not reported)

1	randomised trials	not serious	not serious	very serious[c]	serious[b]	none	18	18	-	MD **0.03 lower** (1.87 lower to 1.81 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Disease Activity (ESR), 12 weeks or more (6 months) Assessed using erythrocyte sedimentation rates (Westergren), mm Hg

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **1.89 higher** (2.93 lower to 6.71 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Anxiety, 12 weeks or more (6 months) Assessed using Trait Form of the State-Trait Anxiety Inventory. Score ranges from 20 to 80, 20= no anxiety

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SGT	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **12.97 lower** (22.38 lower to 3.56 lower)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring SGT**

Depression, 12 weeks or more (6 months) Assessed using Depression Adjective Checklist, scale range not reported

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	18	18	-	MD **5.86 higher** (14.74 lower to 26.46 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

c. Enrolled patients are not typical, or the outcome is an indirect measure of the critical outcome of functional status

Table 7. Yoga compared to control. (4)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Yoga	control	Relative (95% CI)	Absolute (95% CI)		
Functional Status (Left Grip strength) (kg), <12 weeks (30 days) Assessed using hand grip dynamometer (kg)												
1	randomised trials	not serious	not serious	very serious[a]	serious[b]	none	20	20	-	MD **12.5 higher** (7.87 higher to 17.13 higher)	⊕◯◯◯ Very low	CRITICAL **Significant difference favoring Yoga**
Functional Status (Right grip strength) (kg), <12 weeks (30 days) Assessed using hand grip dynamometer (kg)												
1	randomised trials	not serious	not serious	very serious[a]	serious[b]	none	20	20	-	MD **12.8 higher** (8.53 higher to 17.07 higher)	⊕◯◯◯ Very low	IMPORTANT **Significant difference favoring Yoga**

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical, or the outcome is an indirect measure of the critical outcome of functional status

b. Wide confidence intervals

Table 8: Whole Body Vibration (WBV) Therapy (17) vs control

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Whole Body Vibration (WBV) Therapy	Control	Relative (95% CI)	Absolute (95% CI)		
Fatigue, 12 weeks or more (6 months) Assessed using a Likert scale anchored at 0 (not tired at all) and 5 (the most tired I have ever felt)												
	randomised trials	not serious	not serious	very serious[c]	serious[b]	none	16	15	-	MD **0** (0.49 lower to 0.49 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference
Pain, 12 weeks or more (6 months) Assessed using a Likert scale anchored at 0 (no pain) and 5 (unbearable pain)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	16	15	-	MD **0** (0.53 lower to 0.53 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Disability, 12 weeks or more (6 months) assessed using modified Health Assessment Questionnaire (mHAQ). Score ranges from 0 to 3, 0= no disability.												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Whole Body Vibration (WBV) Therapy	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	16	15	-	MD **0.25 lower** (0.39 lower to 0.11 lower)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring WBV Therapy**

Disease Activity, 12 weeks or more (6 months) assessed using Clinical Disease Activity Index.

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	16	15	-	MD **0.84 higher** (0.53 lower to 2.21 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

c. Enrolled patients are not typical, and the outcome is an indirect measure of the critical outcome of functional status

Table 9: CBT compared to Arthritis Education. (1)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	Arthritis Education	Relative (95% CI)	Absolute (95% CI)		
Pain (VAS from RASQ), 12 weeks or more (12 months) Score ranges from 0 to 10, 0=no pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	55	50	-	MD **0.1 higher** (0.8 lower to 1 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Mobility (AIMS-2), 12 weeks or more (12 months) Score ranges from 1 to 5, 0=no mobility problem												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	55	50	-	MD **0.4 higher** (0.28 lower to 1.08 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Mental Health(MHI) Depression, 12 weeks or more (12 months) Assessed using Rand Mental Health Inventory (MHI). Score ranges from 0 to 20, 0=no depression												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	55	50	-	MD **0.5 higher** (0.92 lower to 1.92 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 10: Relaxation response compared to Arthritis Education. (1)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Relaxation Response RR	arthritis education	Relative (95% CI)	Absolute (95% CI)		
Pain (VAS from RASQ), 12 weeks or more (12 months) Score ranges from 0 to 10, 0=no pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	55	50	-	MD **0.1 higher** (0.91 lower to 1.11 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Mobility (AIMS-2), 12 weeks or more (12 months) Score ranges from 1 to 5, 0=no mobility problem												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	55	50	-	MD **0.3 lower** (0.92 lower to 0.32 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Mental health (MHI) Depression, 12 weeks or more (12 months) Assessed using Rand Mental Health Inventory (MHI). Score ranges from 0 to 20, 0=no depression												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Relaxation Response RR	arthritis education	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	55	50	-	MD **0.2 higher** (1.22 lower to 1.62 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 11: Behavioral Therapy with Family Support compared to Behavioral Therapy. (18)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Behavioral Therapy	Relative (95% CI)	Absolute (95% CI)		

Pain (AIMS), 12 weeks or more (2 months) Assessed using Arthritis Impact Measurement Pain Subscale. Score range from 4 to 24, 4 = No pain

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Behavioral Therapy	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	14	-	MD **0.12 higher** (1.69 lower to 1.93 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Physical Functioning (AIMS), 12 weeks or more (2 months) Mobility, Physical Activity, Dexterity, Household Activities, and Activities of Daily Living subscales of the AIMS were used to assess functional impairment. Higher score = more functional impairments

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	14	-	MD **2.51 lower** (9.72 lower to 4.7 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Behavioral Therapy	Relative (95% CI)	Absolute (95% CI)		
Pain (Joint exam pain), 12 weeks or more (2 months) 60 joints were evaluated on a scale from 0 to 3, 0=no pain/tenderness												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	14	-	MD **4.53 higher** (7.04 lower to 16.1 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference
Depression (CES-D), 12 weeks or more (2 months) Measured using 20-item the Center for Epidemiological Studies-Depression Scale. Score ranges from 0 to 60, 0=no depression												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	14	-	MD **2.19 lower** (8.78 lower to 4.4 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference
Disease Activity (Joint exam swelling severity), 12 weeks or more (2 months) 60 joints were evaluated on a scale from 0 to 3, 0=no swelling												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	14	-	MD **0.76 lower** (15.86 lower to 14.34 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Behavioral Therapy	Relative (95% CI)	Absolute (95% CI)		
Disease Activity (Joint exam Number of swollen joints), 12 weeks or more (2 months)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	14	-	MD **1.04 lower** (10.13 lower to 8.05 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 12: Behavioral Therapy with Family Support compared to control. (18)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Joint exam pain), 12 weeks or more (2 months) 60 joints were evaluated on a scale from 0 to 3, 0=no pain/tenderness												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	15	-	MD **1.73 lower** (15.9 lower to 12.44 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Pain (AIMS), 12 weeks or more (2 months). Assessed using Arthritis Impact Measurement Pain Subscale. Score range from 4 to 24, 4 = No pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	15	-	MD **0.74 lower** (2.35 lower to 0.87 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Physical Functioning (AIMS), 12 weeks or more (2 months). Mobility, Physical Activity, Dexterity, Household Activities, and Activities of Daily Living subscales of the AIMS were used to assess functional impairment. Higher score = more functional impairments												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	15	-	MD **3.16 lower** (9.74 lower to 3.42 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Control	Relative (95% CI)	Absolute (95% CI)		
Depression (CES-D), 12 weeks or more (2 months) Measured using 20-item the Center for Epidemiological Studies-Depression Scale. Score ranges from 0 to 60, 0=no depression												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	15	-	MD **2.67 lower** (9.78 lower to 4.44 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Disease Activity (Joint exam swelling severity), 12 weeks or more (2 months). 60 joints were evaluated on a scale from 0 to 3, 0=no swelling												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	15	-	MD **13.13 lower** (30.97 lower to 4.71 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Disease Activity (Joint exam Number of swollen joints), 12 weeks or more (2 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Behavioral Therapy with Family Support	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	15	15	-	MD **8.53 lower** (18.52 lower to 1.46 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

Table 13: Stress management compared to Support. (22)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stress management	Support	Relative (95% CI)	Absolute (95% CI)		
Pain, <12 weeks (10 weeks) Assessed using a 15 cm analog scale. The score ranges from 0 to 0 to 100, 0= No pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	26	25	-	MD **0.2 higher** (2.31 lower to 2.71 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference
Disease Activity (Tender joints), <12 weeks (10 weeks) Number of tender joints												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	26	25	-	MD **0.05 lower** (2.81 lower to 2.71 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Disease Activity (Morning stiffness), <12 weeks (10 weeks) assessed as patient reported duration of morning stiffness (hours)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	26	25	-	MD **0.32 higher** (0.69 lower to 1.33 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference
Disability, <12 weeks (10 weeks) assessed by self-administered, validated scale developed by Fries et al [6], containing nine categories. High score = more disability												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stress management	Support	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	26	25	-	MD **0.12 higher** (0.25 lower to 0.49 higher)	⊕⊕◯◯ Low	CRITICAL No significant difference

Functional Status (Grip Strength), <12 weeks (10 weeks) mm Hg.

1	randomised trials	not serious	not serious	very serious[c]	serious[b]	none	26	25	-	MD **15.6 lower** (31.42 lower to 0.22 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference

Functional Status (Time to walk 50 feet in seconds), <12 weeks (10 weeks) Time in Seconds

1	randomised trials	not serious	not serious	very serious[c]	serious[b]	none	26	25	-	MD **0.3 lower** (2.81 lower to 2.21 higher)	⊕◯◯◯ Very low	CRITICAL No significant difference

Disease Activity (ESR), <12 weeks (10 weeks) mm/hour

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Stress management	Support	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	26	25	-	MD **4.8 lower** (22.3 lower to 12.7 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

c. Enrolled patients are not typical, and the outcome is an indirect measure of the critical outcome of functional status

Table 14: Motivational interviewing/self-regulation compared to control. (12)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Motivational interviewing/self-regulation	no MI/SR	Relative (95% CI)	Absolute (95% CI)		
Functional status (assessed using HAQ) (0-3 scale), 0=no disability												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	38	40	-	MD **0.30 lower** (0.60 lower to 0.01 lower)	⨁⨁◯◯ low	CRITICAL **Significant difference favoring motivational interviewing/self-regulation**
Fatigue, 12 weeks or more (6 months) assessed using Checklist of Individual Strength. The score ranges from 20 to 140, 20=no fatigue												
1	randomised trials	not serious	not serious	very serious[c]	serious[b]	none	38	40	-	MD **2.70 lower** (8.90 lower to 3.50 higher)	⨁◯◯◯ Very low	CRITICAL No significant difference
Self-efficacy (12 weeks or more) (6 months) assessed using an 18-item questionnaire from Bandura. The score ranges from 0–180, 0=low self-efficacy												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Motivational interviewing/self-regulation	no MI/SR	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	38	40	-	MD **19.00 higher** (5.80 higher to 32..20 higher)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring motivational interviewing/self-regulation**

Disease Activity (12 weeks or more) (6 months) assessed using Rheumatoid Arthritis Disease Activity Index. The score ranges from 0–10, 0=no disease activity

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	38	40	-	MD **0.50 higher** (0.03 higher to 0.97 higher)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring no motivational interviewing/self-regulation**

Depression (12 weeks or more) (6 months) assessed using Brief Symptom Inventory. Score ranges from 0–4, 0=no symptoms

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	38	40	-	MD **0.03 lower** (0.15 lower to 0.09higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

CI: confidence interval; **SMD:** standardised mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

c. Enrolled patients are not typical, and the outcome is an indirect measure of the critical outcome of functional status

Table 15: CBT compared to SGT. (2)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	SGT	Relative (95% CI)	Absolute (95% CI)		
Pain (Pain Behavior Score), 12 weeks or more (6 months)												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **6.38 lower** (11.29 lower to 1.47 lower)	⨁⨁◯◯ Low	CRITICAL **Significant difference favoring CBT**
Pain (Pain Intensity Rating), 12 weeks or more (6 months) assessed using 10-cm visual analog scale ratings. The score ranges from 0 to 10, 0= no pain												
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **1.69 lower** (2.78 lower to 0.6 lower)	⨁⨁◯◯ Low	CRITICAL **Significant difference favoring CBT**
Pain (Pain Unpleasantness Rating), 12 weeks or more (6 months) assessed using 10-cm visual analog scale ratings. The score ranges from 0 to 10, 0= no pain												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	SGT	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **1.29 lower** (2.7 lower to 0.12 higher)	⨁⨁◯◯ Low	CRITICAL No significant difference

Disease Activity (Rheumatologist or Nurse Assessment of Disease Activity), 12 weeks or more (6 months), scale range unclear, because authors adjusted for pretreatment scores, but lower scores are better

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **8.99 lower** (15.19 lower to 2.79 lower)	⨁⨁◯◯ Low	IMPORTANT Significant difference favoring CBT

Disease Activity (Articular Index), 12 weeks or more (6 months) Assessed using number of tender joints

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **4.89 lower** (8.73 lower to 1.05 lower)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring CBT**

Functional Status (Grip Strength), 12 weeks or more (6 months), nurse evaluation, scale range not reported

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	SGT	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	very serious[c]	serious[b]	none	17	18	-	MD **1.44 lower** (3.34 lower to 0.46 higher)	⊕◯◯◯ Very low	IMPORTANT No significant difference

Disease Activity (ESR), 12 weeks or more (6 months) – Assessed using erythrocyte sedimentation rates (Westergren), mm Hg

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **4.39 lower** (9.24 lower to 0.46 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Anxiety, 12 weeks or more (6 months) Assessed using Trait Form of the State-Trait Anxiety Inventory. Score ranges from 20 to 80, 20= no anxiety

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **1.41 higher** (8.15 lower to 10.97 higher)	⊕⊕◯◯ Low	IMPORTANT No significant difference

Depression, 12 weeks or more (6 months) Assessed using Depression Adjective Checklist

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	SGT	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **14.43 lower** (35.32 lower to 6.46 higher)	⨁⨁◯◯ Low	IMPORTANT No significant difference

Disease Activity (Rheumatoid Activity Index), 12 weeks or more (6 months) Range from 0–10, 0=no disease activity

1	randomised trials	not serious	not serious	serious[a]	serious[b]	none	17	18	-	MD **29.58 lower** (42.44 lower to 16.72 lower)	⨁⨁◯◯ Low	IMPORTANT **Significant difference favoring CBT**

CI: confidence interval; **MD:** mean difference

Explanations

a. Enrolled patients are not typical

b. Wide confidence intervals

c. Enrolled patients are not typical, and the outcome is an indirect measure of the critical outcome of functional status

Table 16. Additional Data from RCTs (7, 14, 15)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
2062 Parker 1995 (15)	Randomized control study comparing 3 10 week long interventions	10 week intervention, End point of all 3 groups is 15 months	Stress management (n=47, mean age 60.0 years, disease duration 118.0 months (9.8 years), 40% female Attention control (n=49, mean age 59.0 years, disease duration 109.0 months (9.1 years), 45% female) Standard care control (n=45, mean age 60.0 years, disease duration 119.0 months (9.9 years), 42% female)	Stress management (SM)– outpatient stress management program provided individually at 10 weekly visits each lasting 1.5 hours, consisted of relaxation training and CBT, coping strategies, goals, pain, stress, self esteem and social relationships. After 10 weeks was a 15 month maintenance program in which participants were seen once every 3 months to identify program and assist with application of CBT Attention control (AC)– participated in a computer assisted education program based on materials.	Graphs report medians, no variances. Medians determined from Webplot ditizer 15 month Arthritis Self Efficacy Scale median score SM, 221.5 CN, 188.6 AC, 180.2
7898 Parker 1988 (14)	RCT	12 months	83 patients with RA	There were 3 groups –Cognitive behavioral (CB) comprehensive pain management: Basic RA education program (AP): Routine Care (CN)	Coping strategies Questionnaire was the only measure with a sig. difference at 6 and 12 months between groups. At 6 months: Control over pain: CB group 3.5; Placebo group 3.6; Ability to decrease pain: CB group 3.3; Placebo group 3.3; Ignoring pain sensations: CB group 2.4; Placebo group 1.9; Increasing activity level: CB group 3.3; Placebo group 2.9. At 12 months: Control over pain: CB group 3.7; Placebo group 3.2; Ability to decrease pain: CB group 3.3; Placebo group 3.0; Ignoring pain sensations: CB group 2.5; Placebo group 1.9; Increasing activity level: CB group 3.2; Placebo group 2.8. After this, a High Adherence (HA) subgroup was pulled from the CB group and compared to the other groups at 12 months for the rest of the outcomes, which were not initially significant. VAS: HA group 2.7; AP group 4.4

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>% of body pain: HA group 9.6; AP group 10.1
Present pain intensity: HA group 1.4; AP group 2.2</td></tr>
<tr><td>6534, Freeman, 2002 (7)</td><td>RCT</td><td>6 months</td><td>64 patients with newly diagnosed (<1 month) RA</td><td>Cognitive behavioral therapy with emphasis on coping strategies and joint protection</td><td>Median change scores at 3 months
<table>
<tr><th>Outcome (Median values reported)</th><th>CBT</th><th>Control</th><th>Between group p-value</th></tr>
<tr><td>Change in early morning stiffness, 3 months</td><td>-10</td><td>-20</td><td>0.2</td></tr>
<tr><td>Change in 28 tender and swollen joint count, 3 months</td><td>-3</td><td>-6</td><td>0.03</td></tr>
<tr><td>Change in ESR, 3 months</td><td>3</td><td>1</td><td>0.7</td></tr>
<tr><td>Change in Pain VAS, 3 months</td><td>-18</td><td>-4</td><td>0.2</td></tr>
<tr><td>Change in physical function (per AIMS2), 3 months</td><td>0.3</td><td>0</td><td>0.01</td></tr>
<tr><td>Change in affect (per AIMS2), 3 months</td><td>0.6</td><td>-0.4</td><td>0.01</td></tr>
<tr><td>Change in helplessness index, 3 months</td><td>-1.0</td><td>-2.0</td><td>0.003</td></tr>
<tr><td>Change in self-efficacy, 3 months (note – I strongly suspect paper had a typo with the control change score here)</td><td>1.4</td><td>80</td><td>0.1</td></tr>
</table></td></tr>
</table>

References:

1. Barsky AJ, Ahern DK, Orav EJ, Nestoriuc Y, Liang MH, Berman IT, et al. A randomized trial of three psychosocial treatments for the symptoms of rheumatoid arthritis. Seminars in arthritis and rheumatism. 2010;40(3):222-32.
2. Bradley LA, Young LD, Anderson KO, Turner RA, Agudelo CA, McDaniel LK, et al. Effects of psychological therapy on pain behavior of rheumatoid arthritis patients. Treatment outcome and six-month followup. Arthritis and rheumatism. 1987;30(10):1105-14.
3. Dalili Z, Bayazi MH. The effectiveness of Mindfulness-Based Cognitive Therapy on the illness perception and Psychological Symptoms in patients with Rheumatoid Arthritis. Complementary therapies in clinical practice. 2019;34:139-44.
4. Dash M, Telles S. Improvement in hand grip strength in normal volunteers and rheumatoid arthritis patients following yoga training. Indian journal of physiology and pharmacology. 2001;45(3):355-60.

5. Ferwerda M, van Beugen S, van Middendorp H, Spillekom-van Koulil S, Donders ART, Visser H, et al. A tailored-guided internet-based cognitive-behavioral intervention for patients with rheumatoid arthritis as an adjunct to standard rheumatological care: results of a randomized controlled trial. Pain. 2017;158(5):868-78.
6. Fogarty FA, Booth RJ, Gamble GD, Dalbeth N, Consedine NS. The effect of mindfulness-based stress reduction on disease activity in people with rheumatoid arthritis: a randomised controlled trial. Annals of the rheumatic diseases. 2015;74(2):472-4.
7. Freeman K, Hammond A, Lincoln NB. Use of cognitive-behavioural arthritis education programmes in newly diagnosed rheumatoid arthritis. Clinical rehabilitation. 2002;16(8):828-36.
8. Hewlett S, Almeida C, Ambler N, Blair PS, Choy E, Dures E, et al. Group cognitive-behavioural programme to reduce the impact of rheumatoid arthritis fatigue: the RAFT RCT with economic and qualitative evaluations. Health technology assessment (Winchester, England). 2019;23(57):1-130.
9. Hewlett S, Almeida C, Ambler N, Blair PS, Choy EH, Dures E, et al. Reducing arthritis fatigue impact: two-year randomised controlled trial of cognitive behavioural approaches by rheumatology teams (RAFT). Annals of the rheumatic diseases. 2019;78(4):465-72.
10. Hewlett S, Ambler N, Almeida C, Cliss A, Hammond A, Kitchen K, et al. Self-management of fatigue in rheumatoid arthritis: a randomised controlled trial of group cognitive-behavioural therapy. Annals of the rheumatic diseases. 2011;70(6):1060-7.
11. Kılıç N, Parlar Kılıç S. The effect of progressive muscle relaxation on sleep quality and fatigue in patients with rheumatoid arthritis: A randomized controlled trial. International journal of nursing practice. 2021:e13015.
12. Knittle K, De Gucht V, Hurkmans E, Peeters A, Ronday K, Maes S, et al. Targeting motivation and self-regulation to increase physical activity among patients with rheumatoid arthritis: a randomised controlled trial. Clinical rheumatology. 2015;34(2):231-8.
13. Kraaimaat FW, Brons MR, Geenen R, Bijlsma JWJ. The effect of cognitive behavior therapy in patients with rheumatoid arthritis. Behaviour Research and Therapy. 1995;33(5):487-95.
14. Parker JC, Frank RG, Beck NC, Smarr KL, Buescher KL, Phillips LR, et al. Pain management in rheumatoid arthritis patients. A cognitive-behavioral approach. Arthritis and rheumatism. 1988;31(5):593-601.
15. Parker JC, Smarr KL, Buckelew SP, Stucky-ropp RC, Hewett JE, Johnson JC, et al. Effects of stress management on clinical outcomes in rheumatoid arthritis. Arthritis & Rheumatism. 1995;38(12):1807-18.
16. Pradhan EK, Baumgarten M, Langenberg P, Handwerger B, Gilpin AK, Magyari T, et al. Effect of Mindfulness-Based Stress Reduction in rheumatoid arthritis patients. Arthritis and rheumatism. 2007;57(7):1134-42.
17. Prioreschi A, Makda MA, Tikly M, McVeigh JA. In Patients with Established RA, Positive Effects of a Randomised Three Month WBV Therapy Intervention on Functional Ability, Bone Mineral Density and Fatigue Are Sustained for up to Six Months. PloS one. 2016;11(4):e0153470.
18. Radojevic V, Nicassio PM, Weisman MH. Behavioral intervention with and without family support for rheumatoid arthritis. Behavior Therapy. 1992;23(1):13-30.
19. Sharpe L, Schrieber L. A blind randomized controlled trial of cognitive versus behavioral versus cognitive-behavioral therapy for patients with rheumatoid arthritis. Psychotherapy and psychosomatics. 2012;81(3):145-52.

20. Sharpe L, Sensky T, Timberlake N, Ryan B, Allard S. Long-term efficacy of a cognitive behavioural treatment from a randomized controlled trial for patients recently diagnosed with rheumatoid arthritis. Rheumatology (Oxford, England). 2003;42(3):435-41.
21. Sharpe L, Sensky T, Timberlake N, Ryan B, Brewin CR, Allard S. A blind, randomized, controlled trial of cognitive-behavioural intervention for patients with recent onset rheumatoid arthritis: preventing psychological and physical morbidity. Pain. 2001;89(2-3):275-83.
22. Shearn MA, Fireman BH. Stress management and mutual support groups in rheumatoid arthritis. The American journal of medicine. 1985;78(5):771-5.
23. Zautra AJ, Davis MC, Reich JW, Nicassario P, Tennen H, Finan P, et al. Comparison of cognitive behavioral and mindfulness meditation interventions on adaptation to rheumatoid arthritis for patients with and without history of recurrent depression. Journal of consulting and clinical psychology. 2008;76(3):408-21.

PICO 21. Should patients with RA, who are currently employed or want to become employed, use vocational rehabilitation?

No studies met inclusion criteria for this question.

PICO 22: Should patients with RA, who are currently employed or want to become employed, receive work site evaluations and modifications?

Evidence Summary: We included one randomized controlled trial (RCT) on worksite modification reported in two studies, the first included data from baseline and 6 month follow up, and the other included 12 month follow up on the same sample. (1,2) Critical outcomes for this PICO are pain, function, and work outcomes.

The total number of participants was 150 (75 in control: 84% female, age: 49.6 years, RA duration: 10.0 years, DAS: 2.7, and 75 in intervention: 84% female, age: 49.8 years, RA duration: 10.9 years, DAS: 2.7). All patients received usual rheumatologist-led care, which meant they were treated according to the current guidelines in The Netherlands.
Intervention: The patients in the intervention arm received the Care for Work intervention program, which consisted of two components: integrated care and a participatory workplace intervention. Integrated care was delivered by a multidisciplinary team, which consisted of a trained clinical occupational physician (who acted as care manager), a trained occupational therapist and the patients' own rheumatologist. The care manager was responsible for the planning and coordination of care and for communication between members of the multidisciplinary team. The care manager started the intervention with the intake of the patient. The care manager started with history taking and physical examination with the goal to identify functional limitations at work and factors that could influence functioning at work. The care manager proposed a treatment plan at the end of the first consultation. After the patient's consent, the care manager sent the treatment plan to the other members of the multidisciplinary team. The patients visited the care manager again after 6 and 12 weeks to evaluate, and, if necessary, adjust the treatment plan. After the occupational therapist received the treatment plan from the care manager, the occupational therapist started the participatory workplace intervention, which is based on active participation and strong commitment of both the patient and supervisor. The workplace intervention was based on methods used in participatory ergonomics. The intervention was delivered 3 times, at baseline, 6 weeks, and 12 weeks. Controls received no intervention. Main outcomes were at-work productivity loss, work limitations questionnaire, and work instability. Secondary outcomes include quality of life/ RANDS 36.

The trial found statistical significance in two critical outcomes including the Work Limitations Questionnaire Time management demands (at 6 months) and at work productivity loss (at 12 months). Sixteen other outcomes (including Work Instability and quality of life) showed no statistically significant differences between the groups. However, this does not imply that no difference exists. For some outcomes, the confidence interval was wide enough to include important effects (e.g., 17 points on the 0-100 scale for quality-of-life physical role limitations).

We graded the evidence as low certainty based on downgrades for serious risk of bias and imprecision. The study did not have any serious bias related to inconsistency or indirectness.
Quality of Evidence Across Critical Outcomes: Low

Table 1: Workplace integrated care compared to Usual care

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Workplace integrated care	Usual care	Relative (95% CI)	Absolute (95% CI)		
WLQ- Time management demands Timepoint: 6 months Range of scores: 0-100 (0 no limitation- 100 highest limitation)												
1	randomised trials	serious	not serious	not serious	serious	none	75	75	-	MD **7.2 higher** (0.91 higher to 13.49 higher)	⨁⨁◯◯ Low	Critical Statistically significant difference favoring control group
WLQ- Time management demands Timepoint: 12 months Range of scores: 0-100 (0 no limitation- 100 highest limitation)												
1	randomised trials	serious	not serious	not serious	serious	none	71	72	-	MD **6.2 higher** (0.68 lower to 13.08 higher)	⨁⨁◯◯ Low	Critical NS
WLQ Mental demands Timepoint: 6 months Range of scores: 0-100 (0 no limitation- 100 highest limitation)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Workplace integrated care	Usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious	not serious	not serious	serious	none	75	75	-	MD **3.9 higher** (1.4 lower to 9.2 higher)	⨁⨁◯◯ Low	Critical NS

WLQ Mental demands
Timepoint: 12 months
Range of scores: 0-100 (0 no limitation- 100 highest limitation)

1	randomised trials	serious	not serious	not serious	serious	none	71	72	-	MD **5 higher** (1.17 lower to 11.17 higher)	⨁⨁◯◯ Low	Critical NS

WLQ physical demands
Timepoint: 6 months
Range of scores: 0-100 (0 no limitation- 100 highest limitation)

1	randomised trials	serious	not serious	not serious	serious	none	75	75	-	MD **2.9 lower** (9.61 lower to 3.81 higher)	⨁⨁◯◯ Low	Critical NS

WLQ physical demands
Timepoint: 12 months
Range of scores: 0-100 (0 no limitation- 100 highest limitation)

1	randomised trials	serious	not serious	not serious	serious	none	71	72	-	MD **4.3 higher** (2.24 lower to 10.84 higher)	⨁⨁◯◯ Low	Critical NS

WLQ outputs demand
Timepoint: 6months
Range of scores: 0-100 (0 no limitation- 100 highest limitation)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Workplace integrated care	Usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious	not serious	not serious	not serious	none	75	75	-	MD **0.1 higher** (6.08 lower to 6.28 higher)	⨁⨁⨁◯ Moderate	Critical NS

WLQ outputs demand
Timepoint: 12 months
Range of scores: 0-100 (0 no limitation- 100 highest limitation)

1	randomised trials	serious	not serious	not serious	serious	none	71	72	-	MD **6.1 higher** (0.11 lower to 12.31 higher)	⨁⨁◯◯ Low	Critical NS

Work Instability Scale (WIS)
Timepoint: 6 months
Range of scores: 0-23

1	randomised trials	serious	not serious	not serious	not serious	none	75	75	-	MD **0.7 higher** (0.99 lower to 2.39 higher)	⨁⨁⨁◯ Moderate	Critical NS

Work Instability Scale (WIS)
Timepoint: 12 months
Range of scores: 0-23

1	randomised trials	serious	not serious	not serious	not serious	none	73	72	-	MD **1.6 higher** (0.23 lower to 3.43 higher)	⨁⨁⨁◯ Moderate	Critical NS

At work-productivity loss
Timepoint: 6 months
Range of scores: Total hours

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Workplace integrated care	Usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious	not serious	not serious	not serious	none	75	75	-	MD **0.8 higher** (0.1 lower to 1.7 higher)	⨁⨁⨁◯ Moderate	Critical NS

At work-productivity loss
Timepoint: 12 months
Range of scores: Total hours

1	randomised trials	serious	not serious	not serious	not serious	none	71	72	-	MD **1.1 higher** (0.23 higher to 1.97 higher)	⨁⨁⨁◯ Moderate	Critical Statistically significant difference favoring control group

QoL- Mental Health
Timepoint: 6 months
Range of scores: 0-100 (100 indicates better health)

1	randomised trials	serious	not serious	not serious	not serious	none	75	75	-	MD **1.4 lower** (6.16 lower to 3.36 higher)	⨁⨁⨁◯ Moderate	Important NS

QoL- Mental Health
Timepoint: 12 months
Range of scores: 0-100 (100 indicates better health)

1	randomised trials	serious	not serious	not serious	serious	none	71	72	-	MD **4.2 lower** (9.36 lower to 0.96 higher)	⨁⨁◯◯ Low	Important NS

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Workplace integrated care	Usual care	Relative (95% CI)	Absolute (95% CI)		
QoL- Physical role limitations: RAND 36 Timepoint: 6 months Range of scores: 0-100 (100 indicates better health)												
1	randomised trials	serious	not serious	not serious	very serious	none	75	75	-	MD **4.1 lower** (17.44 lower to 9.24 higher)	⨁◯◯◯ Very low	Important NS
QoL- Physical role limitations: RAND 36 Timepoint: 12 months Range of scores: 0-100 (100 indicates better health)												
1	randomised trials	serious	not serious	not serious	very serious	none	71	72	-	MD **12 lower** (25.43 lower to 1.43 higher)	⨁◯◯◯ Very low	Important NS
QoL- Physical Functioning Timepoint: 6 months Range of scores: 0-100 (100 indicates better health)												
1	randomised trials	serious	not serious	not serious	not serious	none	75	75	-	MD **2 higher** (4.37 lower to 8.37 higher)	⨁⨁⨁◯ Moderate	Important NS
QoL- Physical Functioning Timepoint: 12 months Range of scores: 0-100 (100 indicates better health)												
1	randomised trials	serious	not serious	not serious	serious	none	71	72	-	MD **3.3 lower** (10.01 lower to 3.41 higher)	⨁⨁◯◯ Low	Important NS

CI: confidence interval; **MD:** mean difference

References

1. van Vilsteren, Myrthe, et al. "Effectiveness of an integrated care intervention on supervisor support and work functioning of workers with rheumatoid arthritis." *Disability and Rehabilitation* 39.4 (2017): 354-362.

2. Van Vilsteren, M., et al. "One year effects of a workplace integrated care intervention for workers with rheumatoid arthritis: results of a randomized controlled trial." *Journal of occupational rehabilitation* 27.1 (2017): 128-136.

Additional integrative interventions

PICO 23: Should patients with RA use acupuncture?

Summary of findings: The studies included for this PICO question had interventions such as acupuncture, moxibustion, "triple strong" technique, and electroacupuncture (1-10) as add-on treatments to anti-rheumatic drugs. For several outcomes, data favored acupuncture as compared to controls with statistical significance(3), but results were inconclusive between electroacupuncture and placebo (statistically non-significant) (2). One study on a triple-strong technique (which included bloodletting, cupping and moxibustion) in RA patients had more favorable outcomes as compared to controls (6). One study comparing the effects of moxibustion both with and without ARD treatment versus only ARD, had more favorable outcomes in groups with moxibustion (7). Three RCTs have a data for outcomes with effect sizes not computable in review manager, that were slightly less favorable for acupuncture, but the results are very imprecise (8-10).

The tables below summarize the evidence on five comparisons:

- Five RCTs: Acupuncture compared to Controls. (1-5) Traditional acupuncture compared to sham acupuncture (2), Laser acupuncture and telerehabilitation sessions, which consisted of aerobic exercise and virtual reality training compared to telerehabilitation sessions only (1), acupuncture combined with DMARD therapy versus DMARD therapy only (3), acupuncture compared to superficial acupuncture at non-acupuncture points (4), acupuncture + MTX+LEF compared to MTX+LEF (5).
- One RCT: Electroapuncture compared to sham acupuncture (2)
- One RCT: Triple strong technique (bloodletting, cupping and moxibustion) in addition to ARD (diclofenac, MTX, folic acid) versus control (only diclofenac, MTX, folic acid) (6)
- One RCT: Moxibustion + ARD compared to ARD only (7)
- One RCT: Moxibustion Only compared to ARD only (7)
- Three RCTs with additional data with not computable effect sizes compared Acupucture to Placebo (9), Acupuncture combined with MTX and telerehabilitation to MTX and telerehabilitation (8), and acupuncture and intra-articular GC to intra-articular GC (10).

The GRADE tables appear in Tables 1 through 5, and additional data where effect sizes were not computable appear in Table 6.

Overall Quality of Evidence: Low

Table 1: Acupuncture versus no acupuncture (1-5)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
VAS pain, 12 weeks (4, 5) (scale range not reported)												
2	randomised trials	not serious	not serious	not serious	serious[a]	none	34	36	-	MD **0.76 lower** (2.18 lower to 0.66 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference
Function, HAQ change 13 weeks (2, 4) (scale range not reported)												
2	randomised trials	not serious	not serious	not serious	serious[a]	none	26	28	-	MD **0.07 lower** (0.45 lower to 0.31 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference
Patient Global VAS (1-10 scale), 10 weeks (2) (authors did not specify what was being rated, but could involve pain)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	12	12	-	MD **0** (1.92 lower to 1.92 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference
Physician Global VAS (1-10 scale), 10 weeks (2)(authors did not specify what was being rated, but could involve pain)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	not serious	serious[a]	none	12	12	-	MD **0.1 lower** (1.86 lower to 1.66 higher)	⊕⊕⊕◯ Moderate	Critical No statistically significant difference

Joint pain (1-10 scale), 8 weeks (2, 3)

2	randomised trials	not serious	not serious	not serious	serious[a]	none	112	112	-	MD **0.59 lower** (0.68 lower to 0.50 lower)	⊕⊕⊕◯ Moderate	Critical Statistically significantly favors acupuncture

Function as inferred from Gripping power (kpa), 8 weeks (3)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **3.04 higher** (2.38 higher to 3.7 higher)	⊕⊕◯◯ Low	Critical Statistically significantly favors acupuncture

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
Function as inferred from Walking time to finish a walk of 20 meters (in seconds), 8 weeks (3)												
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **7.74 lower** (9.01 lower to 6.47 lower)	⨁⨁◯◯ Low	Critical Statistically significantly favors acupuncture
Function, HAQ, 4 weeks. (1) (scale range not reported)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	30	30	-	MD **0** (0.22 lower to 0.22 higher)	⨁⨁⨁◯ Moderate	Critical No statistically significant difference
Joint pain, 4 weeks (3) (scale range not reported)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	100	100	-	MD **0.74 lower** (0.9 lower to 0.58 lower)	⨁⨁⨁◯ Moderate	Critical Statistically significant favors acupuncture

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
Function as inferred from Gripping power (kpa), 4 weeks (3)												
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **2.83 higher** (1.97 higher to 3.69 higher)	⊕⊕◯◯ Low	Critical Statistically significantly favors acupuncture
Function as inferred from Walking time (seconds), 4 weeks (3)												
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **4.39 lower** (5.73 lower to 3.05 lower)	⊕⊕◯◯ Low	Critical Statistically significantly favors acupuncture
Disease activity as inferred from Swollen joint count (number), 12-13 weeks (4, 5)												
2	randomised trials	not serious	not serious	serious[b]	serious[a]	none	34	36	-	MD **2.98 lower** (3.66 lower to 2.3 lower)	⊕⊕◯◯ Low	Important Statistically significant favors acupuncture
Disease activity as inferred from Tender joint count (number), 12-13 weeks (4, 5)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
2	randomised trials	not serious	not serious	serious[b]	serious[a]	none	34	36	-	MD **1.96 lower** (3.60 lower to 0.32 lower)	⊕⊕◯◯ Low	Important Statistically significant favors acupuncture

Disease activity DAS28 (score), 12 weeks (2, 4, 5)

3	randomised trials	not serious	not serious	not serious	serious[a]	none	46	48	-	MD **0.3 lower** (0.71 lower to 0.11 higher)	⊕⊕⊕◯ Moderate	Important No statistically significant difference

Disease activity as inferred from Morning stiffness (minutes), 8 weeks (3, 5)

2	randomised trials	not serious	not serious	serious[b]	serious[a]	none	120	120	-	MD **7.17 lower** (11.71 lower to 2.63 lower)	⊕⊕◯◯ Low	Important Statistically significantly favors acupuncture

Disease activity as inferred from Joint swelling (score), 8 weeks (3)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **0.38 lower** (0.45 lower to 0.31 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors acupuncture

Remission (number), 8 weeks (3)

1	randomised trials	not serious	not serious	not serious	serious[a]	none	20/100 (20.0%)	10/100 (10.0%)	**RR 2.00** (0.99 to 4.05)	**100 more per 1,000** (from 1 fewer to 305 more)	⨁⨁⨁◯ Moderate	Important No statistically significant difference

Disease activity as inferred from High activity stage (number), 8 weeks (3) ("high" activity not defined by authors)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	10/100 (10.0%)	20/100 (20.0%)	**RR 0.50** (0.25 to 1.01)	**100 fewer per 1,000** (from 150 fewer to 2 more)	⨁⨁◯◯ Low	Important No statistically significant difference

Disease activity as inferred from Morning stiffness (minutes), 4 weeks (3)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **5.62 lower** (8.2 lower to 3.04 lower)	⊕⊕◯◯ Low	Important Statistically significantly favors acupuncture

Disease activity as inferred from Joint swelling (0-3 score where lower is better), 4 weeks (3)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	100	100	-	MD **0.11 lower** (0.19 lower to 0.03 lower)	⊕⊕◯◯ Low	Important Statistically significantly favors acupuncture

Any adverse events (number), 4 weeks (3)

1	randomised trials	not serious	not serious	not serious	serious[a]	none	7/100 (7.0%)	9/100 (9.0%)	**RR 0.78** (0.30 to 2.01)	**20 fewer per 1,000** (from 63 fewer to 91 more)	⊕⊕⊕◯ Moderate	Important No statistically significant difference

RAQoL 4 weeks. (1) (scale range not reported)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	no acupuncture	Relative (95% CI)	Absolute (95% CI)		Statistical significance
1	randomised trials	not serious	not serious	not serious	serious[a]	none	30	30	-	MD **4.47 lower** (7.86 lower to 1.08 lower)	⨁⨁⨁◯ Moderate	Important Statistically significantly favors acupuncture

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Less than 200 patients in each group
b. Surrogate measure

Table 2: Electroacupuncture vs control (2)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Electroacupuncture vs Placebo	placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (1-10 scale), 10 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Electroacupuncture vs Placebo	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	12	12	-	MD **0.6 higher** (1.09 lower to 2.29 higher)	⨁⨁◯◯ Low	Critical No statistically significant difference

Function, HAQ , 10 weeks

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	12	12	-	MD **0** (0.65 lower to 0.65 higher)	⨁⨁◯◯ Low	Critical No statistically significant difference

Patient Global VAS (1-10 scale), 10 weeks (authors did not specify what was being rated, but could involve pain)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	12	12	-	MD **0.8 higher** (1 lower to 2.6 higher)	⨁⨁◯◯ Low	Important No statistically significant difference

Physician Global VAS (1-10 scale), 10 weeks (authors did not specify what was being rated, but could involve pain)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Electroacupuncture vs Placebo	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	12	12	-	MD **0.8 lower** (2.56 lower to 0.96 higher)	⊕⊕◯◯ Low	Important No statistically significant difference
DAS28 (count), 10 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	12	12	-	MD **0** (0.84 lower to 0.84 higher)	⊕⊕◯◯ Low	Important No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Differential attrition

b. Less than 200 patients in each group

Table 3: Triple strong technique (bloodletting, cupping and moxibustion) versus control (6)

							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Triple strong technique	control	Relative (95% CI)	Absolute (95% CI)		
DAS 28, 30 days												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **1.11 lower** (1.54 lower to 0.68 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors triple strong technique
Effectiveness rate (number), 30 days												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	29/30 (96.7%)	24/30 (80.0%)	**RR 1.21** (1.00 to 1.46)	**168 more per 1,000** (from 0 fewer to 368 more)	⨁⨁◯◯ Low	Important No statistically significant difference
Quantitative grading of symptom (scores), 30 days (authors did not specify what was being rated, but could involve pain)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **1.19 lower** (5.48 lower to 3.1 higher)	⨁⨁◯◯ Low	Important No statistically significant difference

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

Explanations

a. Open-label RCT

b. Less than 200 patients in each group

Table 4: Moxibustion+ARD compared to ARD only (7)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Moxibustion+ARD	ARD only	Relative (95% CI)	Absolute (95% CI)		
Pain VAS (8 weeks) (scale range not reported)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	60	60	-	MD **2 lower** (2.36 lower to 1.64 lower)	⊕⊕⊕◯ Moderate	Critical Statistically significantly favors moxibustion
Function, HAQ (8 weeks)												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	60	60	-	MD **3.3 lower** (3.87 lower to 2.73 lower)	⊕⊕⊕◯ Moderate	Critical Statistically significantly favors moxibustion
Disease activity as inferred from Swollen Joints (number), (8 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Moxibustion+ARD	ARD only	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	60	60	-	MD **2.9 lower** (3.25 lower to 2.55 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors moxibustion

Disease activity as inferred from Tender Joints (number), (8 weeks)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	60	60	-	MD **3.4 lower** (3.98 lower to 2.82 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors moxibustion

Disease activity as inferred from Duration Morning Stiffness (minutes), (8 weeks)

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	60	60	-	MD **36.5 lower** (46.25 lower to 26.75 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors moxibustion

Disease activity, DAS-28 (8 weeks)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Moxibustion+ARD	ARD only	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[a]	none	60	60	-	MD **2.4 lower** (2.72 lower to 2.08 lower)	⨁⨁⨁◯ Moderate	Important Statistically significantly favors moxibustion

CI: confidence interval; **MD:** mean difference

Explanations

a. Less than 200 patients in each group

Table 5: Moxibustion Only compared to ARD only (7)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Moxibustion Only	ARD only	Relative (95% CI)	Absolute (95% CI)		
Pain VAS, 8 weeks (scale range not reported)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Moxibustion Only	ARD only	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious[a]	none	60	60	-	MD **1.2 lower** (1.65 lower to 0.75 lower)	⊕⊕⊕◯ Moderate	Critical Statistically significantly favors moxibustion

Function, HAQ, 8 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	60	60	-	MD **2 lower** (2.61 lower to 1.39 lower)	⊕⊕⊕◯ Moderate	Critical Statistically significantly favors moxibustion

Disease activity as inferred from Swollen Joints (number), 8 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	60	60	-	MD **2.1 lower** (2.41 lower to 1.79 lower)	⊕⊕◯◯ Low	Important Statistically significantly favors moxibustion

Disease activity as inferred from Tender Joints (number), 8 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Moxibustion Only	ARD only	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	60	60	-	MD **1.8 lower** (2.27 lower to 1.33 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors moxibustion

Disease activity as inferred from Duration Morning Stiffness (minutes), 8 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	60	60	-	MD **22 lower** (31.01 lower to 12.99 lower)	⨁⨁◯◯ Low	Important Statistically significantly favors moxibustion

DAS-28, 8 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	60	60	-	MD **1.2 lower** (1.56 lower to 0.84 lower)	⨁⨁⨁◯ Moderate	Important Statistically significantly favors moxibustion

CI: confidence interval; **MD:** mean difference

Explanations

a. Less than 200 patients in each group

b. Surrogate measure

Table 6: Additional data not used in GRADE tables

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1730 David 1999 (9)	-randomized placebo-controlled cross-over design	22 weeks	56 RA patients n = 29 for Group A (intervention); n = 27 for Group B (control)	Acupuncture (Liver 3 acupuncture point) vs. placebo (pressure on Liver 3 site with NO skin puncture) -for the first time period, Group A received acupuncture intervention and Group B received placebo. -for the second time period, this was reversed and Group A received placebo, and Group B received acupuncture intervention.	Timepoint is 12 weeks. Data presented as Median (95% CI) (median is of the change in baseline characteristics from Timepoint A-Wk 1 to Timepoint C-Wk 12): Pain • # out of 28 tender joint count (negative) o Group A: -.5 (-3 to 1.5) o Group B: -1 (-3 to .3) • visual analogue scale of pain (VAS P) (negative) o Group A: -4 (-15 to 11) o Group B: 0 (-11 to 5) RA Disease Activity • Disease Activity Score (negative) o Group A: -.2 (-.5 to .4) o Group B: -.4 (-1 to .2) • ESR (negative) o Group A: -1.5 (-6 to 2.3) o Group B: -3 (-8 to 1.2) • CRP (negative) o Group A: 0 (-2.5 to 0) o Group B: 0 (-.5 to 3.7) • # out of 28 swollen joint count (negative) o Group A: 0 (1 to 1) o Group B: 0 (-1.3 to 1) Quality-of-life • General Health Questionnaire (GHQ 28) (negative) o Total score ▪ Group A: -1 (-5 to 0)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					▪ Group B: 0 (-1.3 to 0) ○ Anxiety subscale ▪ Group A: 0 (-1.5 to .5) ▪ Group B: 0 (-1 to 0) ○ Somatic symptom subscale ▪ Group A: 0 (-1 to 0) ▪ Group B: 0 (0 to 0) ○ Socialization subscale ▪ Group A: 0 (-1 to 0) ▪ Group B: 0 (-1 to 0) ○ Depression subscale ▪ Group A: 0 (0 to 0) ▪ Group B: 0 (0 to 0) • visual analogue scale of patient's global assessment (VAS G) (pretty sure this is negative) ○ Group A: 0 (-9 to 14) ○ Group B: -2 (-16 to 6) Other • # anelgesic tablets/day (negative) ○ Group A: 0 (-.5 to 0) Group B: 0 (0 to 0)
2015 Adly, Af 2021 (8)	RCT	Four weeks	60 patients with active RA who had not been on a DMARD three months prior to the study. There were 41 females in 19 males in all were between the ages of 65 to 75.	Patients were treated 3 times a week for four weeks Acupuncture group (group A) - received Remote laser acupuncture, methotrexate, and tele-rehabilitation sessions. The tele-rehabilitation sessions consisted of	Change in HAQ: group A: 0.2350; group B: 0.0460 (non-significant p-value) Change in RAQOL: group A: 2.733; group B: 0.3 (non-significant p-value)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
				aerobic exercise and virtual reality training. Control group (group B) - received tele-rehabilitation sessions which consisted of aerobic exercise and virtual reality training, and methotrexate	
4842 Man 1974 (10)	RCT	3 months	20 RA patients	10 patients were given acupuncture and intra-articular GC and 10 controls had intra-articular GC and no acupuncture.	Pain: intervention group 90% moderate decrease, control 10% decrease; Local swelling: intervention group 10% slight increase, control no change; Local heat: intervention group no change, control no change; Range of motion: intervention group 30% slight increase, control no change; Average pain-free duration: intervention group 1-3 months, control less than 10 hours.

References:

1. Adly AS, Adly AS, Adly MS, Ali MF. A novel approach utilizing laser acupuncture teletherapy for management of elderly-onset rheumatoid arthritis: A randomized clinical trial. Journal of telemedicine and telecare. 2021;27(5):298-306.
2. Tam LS, Leung PC, Li TK, Zhang L, Li EK. Acupuncture in the treatment of rheumatoid arthritis: a double-blind controlled pilot study. BMC complementary and alternative medicine. 2007;7:35.
3. Wang X, Wang H-p, Lv X, Wang X, editors. DMARDs combined with acupuncture therapy to treat RA: study of effects on dsDNA/NETs level and mechanism analysis2020.
4. Zanette Sde A, Born IG, Brenol JC, Xavier RM. A pilot study of acupuncture as adjunctive treatment of rheumatoid arthritis. Clinical rheumatology. 2008;27(5):627-35.
5. Zeng C, Bai X, Qin H, Wang H, Rong X, Yan J. Effect of adjuvant therapy with electroacupuncture on bone turnover markers and interleukin 17 in patients with rheumatoid arthritis. Journal of traditional Chinese medicine = Chung i tsa chih ying wen pan. 2019;39(4):582-6.

6. Cao W-z, Zhao W-x, Guo H-m, Zhang X-m, Zhang M-f, Zhang X-l, et al. Rheumatoid arthritis treated with the triple strong-stimulation technique of acupuncture and moxibustion at specific acupoints: A randomized controlled trial. World Journal of Acupuncture - Moxibustion. 2018;28(4):251-6.
7. Liu D, Guo M, Hu Y, Liu T, Yan J, Luo Y, et al. Effect of sanhuangwuji powder, anti-rheumatic drugs, and ginger-partitioned acupoint stimulation on the treatment of rheumatoid arthritis with peptic ulcer: a randomized controlled study. Journal of traditional Chinese medicine = Chung i tsa chih ying wen pan. 2015;35(3):273-80.
8. Adly AS, Adly AS, Adly MS. Effects of laser acupuncture tele-therapy for rheumatoid arthritis elderly patients. Lasers in medical science. 2022;37(1):499-504.
9. David J, Townsend S, Sathanathan R, Kriss S, Doré CJ. The effect of acupuncture on patients with rheumatoid arthritis: a randomized, placebo-controlled cross-over study. Rheumatology (Oxford, England). 1999;38(9):864-9.
10. Man SC, Baragar FD. Preliminary clinical study of acupuncture in rheumatoid arthritis. The Journal of rheumatology. 1974;1(1):126-9.

PICO 24: Should patients with RA receive massage therapy?

Summary: Literature searches identified two randomized controlled trials [1, 4] addressing PICO question #24 on Massage. These studies investigated aromatherapy [1] and Swedish [4] massages. Both the studies found participants who recieved massage reported significantly lower pain compared to those who received no intervention or usual care. However, overall quality of evidence was low given there was serious risk of bias and imprecision. Below we have described the evidence of each type of massage.

Aromatheapy Massage Therapy

One study [1] included 34 participants (17 in massage and 17 in the control group) aged 18 years or older diagnosed with rheumatoid arthritis for at least 1 year, had a Visual Analog Scale (VAS) score of >=4 points and a Fatigue Severity Scale (FSS) score of >=4 points, not currently using biological drug therapy, and not currently receiving physiotherapy or using any complementary therapy modalities. Participants were excluded if they had knee and foot wounds or surgery, cancer, osteoarthritis, essential oil allergies, blood coagulation disorders such as hemophilia, were pregnant, anemic, or who had a Disease Activity Score (DAS28) > 5.1.

Participants were randomized to aromatherapy massage, reflexology or no intervention (the reflexology group was not considered because the treatment is not of interest):

- Experimental group received aromatherapy massage on both knees for 30 minutes, three times each week for a 6-week period. Specifically, the study PI (with a PhD and nursing background) gave aromatherapy massage using essential oil (5% mixture consisting of Lavandula augustifolia, Juniperus communis, Cananga odorata, and Rosmarinus officinalis in the ratio 3:3:2:2 in 100 mL of coconut carrier oil)
- Control received no intervention

The study referenced two earlier publications on aromatherapy massage,[2, 3] but these were not considered because they were not published in English.

Pain (Visual Analog Scale (VAS) score) and Fatigue (Fatigue Severity Scale (FSS) score) were primary outcomes assessed at 6 weeks. The study reported less pain and fatigue at 6 weeks in participants who received aromatherapy massage compared to those who received no intervention. Because fatigue is an indirect measure of function, we downgraded its evidence for indirectness. Both outcomes were downgraded for both serious risk of bias and serious imprecision.

Swedish Massage Therapy

One study [4] included total of 60 patients with a diagnosis of RA that affected one or more joints of the hand, shoulder, elbow, wrist, or fingers; a VAS-pain score of 4 or greater; and no history of surgery in the affected joints. Block randomization method was used to assigned participants to the controlgroup (n = 30) and Swedish massage (n = 30) group.

- Experimental group: Recieved a 30-min Swedish massage regularly for eight weeks: twice a week forthe first four weeks, and three times a week for the last four weeks. The intervention was delivered by a personnel who had a certificate in performing professional Swedish massage therapy.
- Control group: Recieved usual care and treatments included nonsteroidal anti-inflammatory drugs, disease-modifying antirheumatic drugs, and corticosteroids, as well as recommendations for daily activities and lifestyle changes.

Pain (Visual Analog Scale (VAS) score) was primary outcome assessed at 1 month post intervention. The study reported less pain at 1 month post intervention in participants who received swedish massage compared to those who received usual care. This outcome was downgraded for both serious risk of bias and serious imprecision.

The study referenced two earlier publications on massage, [5, 6] but these were not considered because it was either a case study [5] or did not include correct comparator, i.e., 2 different massages were being compared [6]

Quality of evidence across all critical outcomes: Low

Table 1: Data from Randomized Controlled Trials looking at Aromoatherapy Massage

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Experimental (Aromatherapy Massage)	Control	Relative (95% CI)	Absolute (95% CI)		
Pain Score (0-10 Scale) at 6 weeks (Lower values are better)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Experimental (Aromatherapy Massage)	Control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a,b,c,d,e]	not serious	not serious	serious[g]	none	17	17	-	MD **2.7 lower** (3.96 lower to 1.44 lower)	⨁⨁◯◯ Low	Critical **Statistically significant difference favoring the aromatherapy massage group**

Function as inferred from Fatigue Score (Score Range between 9 to 63) at 6 weeks (Lower values are better)

1	randomised trials	serious[a,b,c,d,e]	not serious	serious[f]	serious[g]	none	17	17	-	MD **1.47 lower** (2.48 lower to 0.46 lower)	⨁◯◯◯ Very Low	Critical **Statistically significant difference favoring the aromatherapy massage group**

CI: confidence interval; **MD:** mean difference

Explanations

a. Allocation concealment is Unclear, since they didn't state whether the next assignment in the randomized list was knowable by the person who was including/excluding potential participants (for example by sealed envelopes).

b. Patient/provider blinding. The key thing is whether participants in one group would have differential expectations of the treatments being compared. If so, then knowledge of treatment group can affect outcomes, and so it was determined to be High risk of bias, even if it would be impossible to blind.

c. Outcome assessor blinding. Since patients were assessing their own pain fatigue, this item is also high risk of bias

d. Selective outcome reporting. They measured outcomes using DAS, but didn't report its results, suggesting that the data they DID report are over-estimating the effect size. So this item is high risk of bias.

e. For risk of bias, based on the above, one level downgrade was done

f. For indirectness – Fatigue is a surrogate measure of function, resulting in a one level downgrade

g. For imprecision - Only 34 patients combined, which was determined to be serious imprecision

Table 2: Data from Randomized Controlled Trials looking at Swedish massage

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Swedish Massage	Usual Care	Relative (95% CI)	Absolute (95% CI)		
Pain (0-10 Scale) at 1 month after intervention (Lower values are better)												
1	randomised trials	serious[a,b,c,d]	not serious	not serious	very serious[e]		30	30	-	MD **2.5 lower** (3.01 lower to 1.99 lower)	⨁⨁◯◯ Low -	Critical **Statistically significant difference favoring the Swedish massage group**

CI: confidence interval; **MD:** mean difference

Explanations

a. Allocation concealment is Unclear, since they didn't state whether the next assignment in the randomized list was knowable by the person who was including/excluding potential participants (for example by sealed envelopes).

b. Patient/provider blinding. The key thing is whether participants in one group would have differential expectations of the treatments being compared. If so, then knowledge of treatment group can affect outcomes, and so it was determined to be High risk of bias, even if it would be impossible to blind.

c. Outcome assessor blinding. Since patients were assessing their own pain, this item is also high risk of bias

d. For risk of bias, based on the above, one level downgrade was done

e. For imprecision - Only 60 patients combined, which was determined to be serious imprecision

References

1. Gok Metin Z, Ozdemir L. The Effects of Aromatherapy Massage and Reflexology on Pain and Fatigue in Patients with Rheumatoid Arthritis: A Randomized Controlled Trial. *Pain Manag Nurs*. 2016;17(2):140-149. doi:10.1016/j.pmn.2016.01.00
2. Han, S.-H., Nam, E.-S., Uhm, D.-C., Kim, K.-S., Paik, S.-I., & Park, S.-H. (2010). Effects of aromatherapy on pain and inflammatory responses in patients with rheumatoid arthritis. Journal of Muscle and Joint Health, 17(1), 25–34.
3. Kim, M.-J., Nam, E.-S., & Paik, S.-I. (2005). The effects of aromatherapy on pain, depression, and life satisfaction of arthritis patients. TaehanKanhoHakhoe chi,35(1), 186–194.
4. Sahraei, F., Rahemi, Z., Sadat, Z., Zamani, B., Ajorpaz, N. M., Afshar, M., & Mianehsaz, E. (2022). The effect of Swedish massage on pain in rheumatoid arthritis patients: A randomized controlled trial. Complementary Therapies in Clinical Practice, 46, 101524.
5. G.B. Shetty, A. Mooventhan, N. Anagha, Effect of electro-acupuncture, massage, mud, and sauna therapies in patient with rheumatoid arthritis, J. Ayurveda Integr. Med. 6 (4) (2015) 295–299, https://doi.org/10.4103/0975-9476.172415.
6. T. Field, M. Diego, J. Delgado, D. Garcia, G.G. Funk, Rheumatoid arthritis in upper limbs benefits from moderate pressure massage therapy, Complement, Ther. Clin. Pract. 19 (2) (2013) 101–103, https://doi.org/10.1016/j.ctcp.2012.12.001.

PICO 25: Should patients with RA receive thermal modalities?

Evidence Summary: We included thirteen randomized controlled trials (RCTs)[1-12,15] and two nonrandomized controlled studies[13,14] addressing this PICO question.

- Seven RCTs[4,6,7,9-12] compared **laser therapy** to placebo in either short term (< 12 weeks) or long term (>=12 weeks).
- Two RCTs [2,5] focused on short or long term **ultrasonic hand treatments**.
- One RCT (Gunduz et al. 2019[3]) focused on short term effect of **dry heat** treatment.
- Two RCTs [1,8] focused on short term effects of **paraffin wax bath hand treatments** versus control.
- One RCT (Klemm et al 2022[15]) focused on **cryotherapy** versus a rehabilitation program alone.
- One non-RCT (Sadura-Sieklucka et al. 2019[14]) focused on **cryotherapy** versus a rehabilitation program alone.
- One non-RCT (Hamilton et al. 1959[13]) focused on **short-wave diathermy** versus **parrafin wax** versus **infrared radiation** vs sham diathermy.

We categorized all studies as thermal modality therapy interventions, but analyzed them separately by specific intervention. Primary outcomes were pain and functional status and were classified as short term (<12 weeks) or long terms (>=12 weeks).

The 7 RCTs of laser therapy [4,6,7,9-12] differed in their specific methods, but all provided some type of laser therapy to patients with RA and compared the results to a group receiving either no therapy or sham laser treatment. Only two[6,11] reported a statistically significant difference in pain or functional status favoring laser versus control. No studies found a statistically significant difference between groups for any other outcomes. The certainty of evidence was very low, mainly due to imprecision, small sample size, and high attrition.

Two RCTs[2,5] focused on ultrasonic hand treatment versus control. Patients in both studies participated in an ultrasonic therapy program versus a control group receiving no treatment. Conrad et al. 1994[5] followed patients for less than 12 weeks, while Kiraly et al. 2017[2] followed patients for greater than 12 weeks. Conrad et al. found that patients undergoing ultrasonic hand treatments had a statistically significant found that there was a statistically significant improvement in pain, tenderness, stiffness, and functional status in the treatment groups versus the control group. However, Kiraly et al. did not find any statistically significant difference between treatment and control group for any critical or important outcomes. Certainty of evidence was low, mainly due to small sample size and inconsistency between studies.

Another RCT [3] focusing on dry heat treatments versus control found no statistically significant difference between groups for any of the measures, including pain (VAS), functional status, stiffness, and disease activity. Certainty of evidence was very low due to small sample size, lack of effect, and unclear outcome assessor blinding.

Two RCTs[1,8] focused on short term effects of paraffin wax bath hand treatments versus control. Patients received either paraffin wax bath therapy for <12 weeks or no treatment. Harris et al. 1955[8] found that there was a statistically significant improvement in pain, tenderness, stiffness, and functional status in the treatment groups compared to the control groups, however Dellhag et al. 1992[1] found no statistically significant difference between groups. The certainty of evidence was low, primarily due to small sample size and inconsistency between studies.

One randomized controlled study (Klemm et al 2022[15]) and one non-randomized controlled study (Sadura-Sieklucka et al 2019[14]) had focused on a cryotherapy versus a rehabilitation program. Patients were divided into a cryotherapy group and a conventional rehabilitation group. Both groups used comprehensive rehabilitation which depended on the patient's problems. Physical therapy included electrotherapy, ultrasound, magnetic field, laser therapy, and kinesitherapy included unloading exercises, individual exercises, active exercises, and hand exercises. Patients in the cryotherapy group also performed intermittent sessions of cryotherapy for the duration of the programs. Patient outcomes were assessed at the end of the programs. The studies found that cryotherapy had a statistically significant effect on pain, disease activity, and HAQ scores compared to the control group. Certainty of evidence was low due to small sample size, and high attrition. There was no patient blinding, but this was not possible due to the nature of the study.

One old non-randomized study[13] (published in 1959) used five treatments, and four were relevant to this PICO: short-wave diathermy, hot wax, infra-red radiation and sham diathermy (which served as the control group). The other treatment, faradic stimulation, is discussed in PICO 26. Patients enrolled received either cold short-wave diathermy therapy, infra-red therapy, or a paraffinwax bath treatment versus "cold" diathermy as control. All treatments included a regime of exercise. There were no statistically significant differences between the treatment and control groups for any of the critical or important outcome measures. The certainty of evidence was very low due to small sample size, nonrandomized trial design, and statistically nonsignificant between-group differences.

Quality of evidence across all critical outcomes: Very low

Table 1: Laser compared to Placebo for arthritis

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Laser	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain VAS (0-10) 10 weeks												
1	randomised trials	very serious[d,e]	not serious	serious[f]	serious[b]	none	38	34	-	MD **0.33 lower** (0.65 lower to 0.01 lower)	⨁◯◯◯ Very low	CRITICAL **Statistically significant treatment favoring treatment**
McGill Pain Questionnaire (number of words checked; 0-15) 6 months												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	25	10	-	MD **1.39 higher** (1.85 lower to 4.63 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant treatment
Function: HAQ Disability Index 6 months												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	25	10	-	MD **0.94 lower** (8.16 lower to 6.28 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant treatment
Pain VAS (0-10) 4 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	25	10	-	MD **0.57 lower** (2.77 lower to 1.63 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant treatment
Pain at rest VAS (0-10) 4 weeks												
1	randomised trials	serious[g]	not serious	not serious	very serious[b,c]	none	13	13	-	MD **0.8 higher** (0.29 higher to 1.31 higher)	⨁◯◯◯ Very low	CRITICAL **Statistically significant treatment favoring treatment**
ADL Pain VAS (0-10) 4 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Laser	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[g]	not serious	not serious	very serious[b,c]	none	13	13	-	MD **0.3 higher** (0.12 lower to 0.72 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant treatment

Night Pain VAS (0-10) 4 weeks

1	randomised trials	serious[g]	not serious	not serious	very serious[b,c]	none	13	13	-	MD **0.3 higher** (0.05 lower to 0.65 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant treatment

Function as inferred from walking speed over 20 meters (seconds) 6 months

1	randomised trials	serious[a]	not serious	serious[h]	very serious[b,c]	none	25	10	-	MD **1.91 lower** (11.47 lower to 7.65 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant treatment

Function as inferred from morning stiffness (hours) 6 months

1	randomised trials	serious[a]	not serious	serious[h]	very serious[b,c]	none	25	10	-	MD **0.04 lower** (1.04 lower to 0.96 higher)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant treatment favoring treatment**

Function as inferred from morning stiffness (min) 4 weeks

1	randomised trials	serious[g]	not serious	serious[h]	very serious[b,i]	none	13	13	-	MD **22.3 higher** (6.71 higher to 37.89 higher)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant treatment favoring treatment**

Ritchie index 4 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Laser	Placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[g]	not serious	not serious	very serious[b,i]	none	13	13	-	MD **3.3 higher** (0.79 higher to 5.81 higher)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant treatment favoring treatment**

MCP Swelling 4 weeks

1	randomised trials	serious[g]	not serious	not serious	very serious[b,c]	none	13	13	-	MD **0.4 lower** (0.63 lower to 0.17 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant treatment favoring treatment**

PIP Swelling 4 weeks

1	randomised trials	serious[g]	not serious	not serious	very serious[b,c]	none	13	13	-	MD **0.9 higher** (0.13 higher to 1.67 higher)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant treatment favoring treatment**

CI: confidence interval; **MD:** mean difference

Explanations

a. Outcome assessor blinding not reported

b. Small sample size

c. No statistically significant difference

d. Randomization methods not reported

e. Blinding required patients to close eyes so that they could not see red light from experimental vs placebo laser.

f. Only pain assessed and laser therapy administered to peripheral nerves as well as joints, thus unclear if improvement of pain related to improved RA disease activity or alternative analgesia.

g. 14 out of 40 patients (35%) lost at 3 month follow up

h. Surrogate measure of functional status

i. Treatment favored placebo

Table 2: Ultrasonic hand treatment compared to none for arthritis

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrasonic hand treatment	none	Relative (95% CI)	Absolute (95% CI)		
Pain (VAS 0-100) 14 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	23	-	MD **1 lower** (12.83 lower to 10.83 higher)	⨁⨁⨁◯ Moderate	CRITICAL No statistically significant difference
Number of painful articulations 14 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	23	-	MD **1.1 lower** (3.01 lower to 0.81 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No statistically significant difference
Functional status (HAQ) 14 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	23	-	MD **0.23 lower** (0.65 lower to 0.19 higher)	⨁⨁⨁◯ Moderate	CRITICAL No statistically significant difference
Number of painful articulations 3 weeks												
1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	25	-	MD **1.2 higher** (0.45 higher to 1.95 higher)	⨁⨁⨁◯ Moderate	CRITICAL **Statistically significant difference favoring treatment**
Function as inferred from morning stiffness (minutes) 14 weeks												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrasonic hand treatment	none	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	25	23	-	MD **1.02 higher** (14.74 lower to 16.78 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference

DAS28 (VAS; 0-100 mm) 14 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	23	-	MD **0.49 lower** (1.05 lower to 0.07 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No statistically significant difference

12.50CRP (mg/l) 14 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	23	-	MD **0.12 lower** (4.63 lower to 4.39 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No statistically significant difference

SDESR (mm/h) 14 weeks

1	randomised trials	not serious	not serious	not serious	serious[a]	none	25	23	-	MD **3.71 lower** (10.87 lower to 3.45 higher)	⨁⨁⨁◯ Moderate	IMPORTANT No statistically significant difference

Function as inferred from number of swollen articulations 14 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	25	23	-	MD **0.19 lower** (1.05 lower to 0.67 higher)	⨁⨁◯◯ Low	IMPORTANT No statistically significant difference

Function as inferred from morning stiffness (minutes) 3 weeks

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrasonic hand treatment	none	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	25	25	-	MD **28.54 higher** (0.18 higher to 56.9 higher)	⨁⨁◯◯ Low	IMPORTANT **Statistically significant difference favoring treatment**

Function as inferred from number of swollen articulations 3 weeks

1	randomised trials	not serious	not serious	serious[b]	serious[a]	none	25	25	-	MD **1.02 higher** (0.45 higher to 1.59 higher)	⨁⨁◯◯ Low	IMPORTANT **Statistically significant difference favoring treatment**

CI: confidence interval; **MD:** mean difference

Explanations

a. Small sample size

b. Surrogate measure

Table 3: Dry heat treatment compared to None for RA

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Dry heat treatment	None	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS (0–100 mm) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	40	40	-	MD **0** (117.92 lower to 117.92 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Functional: HAQ (0-5) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	40	40	-	MD **0.2 higher** (0.04 lower to 0.44 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Duruoz Hand Index (0-90, higher score means greater difficulty performing tasks) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	40	40	-	MD **4.5 higher** (5.97 lower to 14.97 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Function as inferred from stiffness (VAS; 0–100 mm) 12 weeks												
1	randomised trials	serious[a]	not serious	Serious[d]	very serious[b,c]	none	40	40	-	MD **0** (39.59 lower to 39.59 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference
Disease Activity (DAS-28) (VAS; 0-100 mm)12 weeks												
1	randomised trials	serious[a]	not serious	not serious	very serious[b,c]	none	40	40	-	MD **0.37 higher** (0.04 lower to 0.78 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. Unclear outcome assessor blinding

b. Small sample size

c. Wide confidence interval

d. Surrogate measure

Table 4: Cryotherapy compared to Rehabilitation for arthritis (Randomized Controlled Trial)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cryotherapy	Rehabilitation	Relative (95% CI)	Absolute (95% CI)		
Pain: 0-10 (higher scores indicate more pain) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	31	25	-	MD 1.31 lower (2.09 lower to 0.53 lower)	⊕⊕◯◯ Low	CRITICAL **Statistically significant difference favoring treatment**
Function: HAQ disability index (0-3, higher sscore indicate higher disability) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	31	25	-	MD 0.21 lower (0.35 lower to 0.07 lower)	⊕⊕◯◯ Low	CRITICAL **Statistically significant difference favoring treatment**
DAS28: 0-10 (higher scores indicate higher disease activity) 12 weeks												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	31	25	-	MD 0.67 lower (1.31 lower to 0.02 lower)	⊕⊕◯◯ Low	IMPORTANT **Statistically significant difference favoring treatment**

a: High attrition bias

b. Small sample size

Table 5: Cryotherapy compared to Rehabilitation for arthritis (Nonrandomized Controlled Trial)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cryotherapy	Rehabilitation	Relative (95% CI)	Absolute (95% CI)		
Morning Pain VAS (0-10) 20 days												
1	observational studies	very serious[a]	not serious	not serious	very serious[b,c]	none	25	25	-	MD **0.1 lower** (1.24 lower to 1.04 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
ADL Pain VAS (0-10) 20 days												
1	observational studies	very serious[a]	not serious	not serious	very serious[b,c]	none	25	25	-	MD **0.5 lower** (1.67 lower to 0.67 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Night Pain VAS (0-10) 20 days												
1	observational studies	very serious[a]	not serious	not serious	very serious[b,c]	none	25	25	-	MD **0.8 lower** (1.97 lower to 0.37 higher)	⨁◯◯◯ Very low	CRITICAL No statistically significant difference
Disease activity: DAS28 (VAS; 0-100 mm) 20 days												
1	observational studies	very serious[a,c]	not serious	not serious	serious[b]	none	25	25	-	MD **0.5 lower** (0.97 lower to 0.03 lower)	⨁◯◯◯ Very low	IMPORTANT **Statistically significant difference favoring treatment**
ESR 20 days												
1	observational studies	very serious[a]	not serious	not serious	very serious[b,c]	none	25	25	-	MD **2.2 lower** (9.68 lower to 5.28 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference
CRP 20 days												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cryotherapy	Rehabilitation	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious[a]	not serious	not serious	very serious[b,c]	none	25	25	-	MD **1.9 lower** (7.28 lower to 3.48 higher)	⨁◯◯◯ Very low	IMPORTANT No statistically significant difference

CI: confidence interval; **MD:** mean difference

Explanations

a. No randomization, no allocation concealment, outcome assessors not blinded

b. Small sample size

c. Wide confidence interval

Table 6: Additional data from RCT and observational studies

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
8514, Harris et al., 1955	A randomized clinical trial	3 weeks, 6 weeks	RA patients = 90 Group 1 (no treatment) n = 23 Age, mean: 46 Male: 4 Female: 19 Group 2 (3 weeks treatment) n = 25 Age, mean: 50 Male: 9 Female: 16 Group 3 (six weeks treatment) n = 23 Age, mean: 48 Male: 4 Female: 19	Group I received no local treatment to the hands, Group II had wax baths daily for 3 weeks, and Group III had wax baths daily for 6 weeks	At the end of the study, group 1 had a mean pain score (0-3) of 0.3, while group 2 had a 1.6 and group 3 had a 0.9. At the end of the study, group 1 had a mean tenderness score of 5.8, compared to 10.4 in group 2 and 7.2 in group 3. At the end of the study, group 1 had a mean swelling score of 3.3, compared to 4.7 in group 2 and 3.9 in group 3. At the end of the study, group 1 had a mean grip strength (mm Hg) score of 128, compared to 116 in group 2 and 91 in group 3. At the end of the study, group 1 had a mean dexterity score of 35, compared to 33 in group 2 and 31 in group 3.

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
8515, Huessler et al., 1993	A double-blind randomized trial	5 weeks	RA patients = 25 Hands receiving low level laser treatment n = 25 Age, mean: 64.8 (43-77) Female: 25 Hands receiving sham laser n = 25 Age, mean: 62.5 (40-80) Female: 25	25 hands treated with active laser and 25 hands treated with sham laser. All patients were right hand dominant. A course of 12 laser treatments was given over a four week period. The active and sham laser probes were identical in external appearance.	Eighteen of 25 patients (72%) reported improvement in pain, but pain (as measured by the visual analogue scale) was reduced in both treated and sham treated hands after laser intervention ($p<0.001$). To assess the patients' ability to discriminate between active laser and placebo they were asked to identify which, if either, hand they felt had improved with treatment. Only five of 25 (20%) identified the treated hand as the one that had improved, whereas five thought that the sham treated hand had improved, eight thought both hands had improved equally, six noted no change in either hand, and one reported that both hands had worsened. There were no significant differences between the treated hand and the sham treated hand over the period of the trial in the duration of early morning stiffness, total swollen joint count, the joint circumferences index, range of motion, pulp tip to distal palmar crease distance, grip strength measures, or the Jebsen Activities Index.

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td>8516, Johannsen, 1994</td><td>RCT</td><td>1 month</td><td>22 patients with active RA</td><td>Low energy laser therapy directed to MCPs, vs placebo</td><td>Median (IQR) at 1 month
<table>
<tr><th>Outcome</th><th>Low energy laser (N=10)</th><th>Control (N=12)</th></tr>
<tr><td>Pain</td><td>7 (2.8-10.3)</td><td>5.5 (3-8.8)</td></tr>
<tr><td>Grip strength</td><td>6.5 (1.5-11.8)</td><td>5.5 (3.3-10.3)</td></tr>
<tr><td>CRP (nmol)</td><td>96 (30-630)</td><td>216 (122-470)</td></tr>
<tr><td>ESR</td><td>12 (5-45)</td><td>32 (14-95)</td></tr>
</table>
*No statistically significant between group differences found

30% (3/10) intervention patients compared to 8% (1/12) control patients reported “improved morning stiffness”</td></tr>
</table>

<table>
<tr><th>Ref ID, Author, year</th><th>Study type</th><th>Duration</th><th>Population Description</th><th>Treatment given to relevant population</th><th>Results</th></tr>
<tr><td>8522, Palmgren, 1989</td><td>RCT</td><td>4 weeks</td><td>35 patients with RA</td><td>Low energy laser to MCPs and PIPs vs placebo</td><td>
Median (95% CI) at 1 month
<table>
<tr><th>Outcome</th><th>Low energy laser (N=19)</th><th>Control (N=16)</th></tr>
<tr><td>ESR</td><td>19 (10-40)</td><td>25.5 (16-39)</td></tr>
<tr><td>Grip strength</td><td>24.8 (14.9-39.6)</td><td>15.3 (8.4-33.8)</td></tr>
<tr><td>AM stiffness (hr)</td><td>0.49 (0-0.98)</td><td>0.79 (0-1.5)</td></tr>
</table>
Within group changes
Change scores not reported
<table>
<tr><th>Outcome</th><th>Low energy laser (N=19)</th><th>Control (N=16)</th></tr>
<tr><td>Pain</td><td>Significant (p<0.001) decrease</td><td>Significant (p<0.001) decrease</td></tr>
<tr><td>Grip strength (kPa)</td><td>Pre: 19
Post: 25
p<0.001</td><td>Pre: 17
Post: 15.5
No sig change</td></tr>
<tr><td>AM stiffness (hr)</td><td>Pre: 1.25
Post: 0.5
p<0.01</td><td>Pre: 1.0
Post: 0.8
No sig change</td></tr>
<tr><td>ESR</td><td>No sig change</td><td>No sig change</td></tr>
</table>
</td></tr>
</table>

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
8510 Bliddal 1987	Double-Blinded, Randomized Controlled Study.	4 weeks	17 patients with symmetrical involvement of the metacarpophalangeal joint of the index.	Nine treatments with a He-Ne laser, 6 J/cm*, were given on the one hand (**Laser**) with a sham irradiation of the other (**Placebo**). The study was double-blind. The patients were randomized to therapy on the MP joint of the right and left index with laser or placebo. Therapy was given on 3 alternate days a week for 3 consecutive weeks, followed by an observation time of 4 weeks.	Each day before, during, and 4 weeks after therapy the patients estimated for right and leftindex separately: I) pain by a visual analogue scale (VAS), 2) duration of morning stiffness (MS), and3) other effects of the treatment NO SE or 95%CI or SD were reported. At 4 weeks after therapy: **VAS Score:** laser better than placebo: 5 Placebo better than laser: 0 **MS Score:** laser better than placebo: 4 Placebo better than laser: 0 Detailed numbers were not provided: 1. The joint ability score showed a tendency to amelioration in both laser- and placebo treated joints, although this difference did not reach statistical significance 2. No changes in laboratory tests (sedimentation rate, hemoglobin, leukocyte and platelet counts) 3. **Adverse effects were noted in 3 patients**, who complained of a burning sensation in the irradiated joint-all on the **laser**-treated side but none withdrew from study

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1193 Hamilton 1959	Nonrandomized comparison study	20 weeks	18 patients with RA of the knee 33 patients with RA of the hand	1. Short wave diathermy 2. Infrared treatment 3. Hot wax for RA of the hand 4. Faradism of the quadriceps for RA of the knee (included in PICO 26) 5. Sham diathermy (control group)	Walk time at conclusion of 4 week of wax -2.34 standard errors of the mean when compared to the mean improvement seen across all treatment groups. All other measures (stair time, power, range of motion) had a difference of less than 2 SEM for cold diathermy, short-wave diathermy, infra-red therapy, and paraffin-wax baths and were statistically nonsignificant.

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1728 Dellhag 1992	Randomized controlled trial	4 weeks	52 Patients with RA, younger than age 70, with impairment of hand function	Wax bath, 20 minutes, 3 times a week for 4 weeks	Mean pain: wax group 1.6, control 1.5 Grip function: wax group 75.0, control 75.0 Pinch function: wax group 28.3, control 29.2 All p values n.s., no standard deviations are reported.

References:

1. Dellhag B, Wollersjo I, Bjelle A. Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. *Arthritis Care Res*. 1992;5(2):87-92.

2. Kiraly M, Varga Z, Szanyo F, Kiss R, Hodosi K, Bender T. Effects of underwater ultrasound therapy on pain, inflammation, hand function and quality of life in patients with rheumatoid arthritis - a randomized controlled trial. *Braz J Phys The*. 2017;21(3):199-205. doi.org/10.1016/j.bjpt.2017.04.002

3. Erdinc Gunduz N, Erdem D, Kizil R, et al. Is dry heat treatment (fluidotherapy) effective in improving hand function in patients with rheumatoid arthritis? A randomized controlled trial. *Clin Rehabil*. 2019;33(3):485-493. doi.org/10.1177/0269215518810778

4. Hall J, Clarke AK, Elvins DM, Ring EF. Low level laser therapy is ineffective in the management of rheumatoid arthritic finger joints. *BR J Rheumatol*. 1994;33(2):142-147.

5. Konrad K. Randomized, double blind, placebo-controlled study of ultrasonic treatment of the hands of rheumatoid arthritis patients. *Eur J Phys Rehabil Med*. 1994;4(5):155-157.

6. Bliddal H, Hellesen C, Ditlevsen P, et al. Soft-laser therapy of rheumatoid arthritis. *Scand J Rheumatol*. 1987(16):225–228.

7. Goats GC, Flett E, Hunter JA, Stirling A. Low-intensity laser and phototherapy for rheumatoid arthritis. *Physiotherapy*. 1996(82):311–320.

8. Harris R, Millard JB. Paraffin wax baths in the treatment of Rheumatoid Arthritis. *Ann Rheum*. 1955(14):278-282.

9. Heussler JK, Hinchey G, Margiotta E, et al. A double blind randomized trial of low-power laser treatment in rheumatoid arthritis. *Ann Rheum Dis*. 1993(52):703–706.

10. Johannsen F, Hauschild B, Remvig L, et al. Low-energy laser therapy in rheumatoid arthritis. *Scand J Rheumatol*. 1994(23):145–147.

11. Palmgren N, Jensen GF, Kamma K, et al. Low-power laser therapy in rheumatoid arthritis. 1989;4:193-196.

12. Walker JB, Akhanjee LK, Cooney MM, et al. Laser therapy for pain of rheumatoid arthritis. *Clin J Pain*. 1987;3:54-59.

13. Hamilton DE, Bywaters EG, Please NW. A controlled trial of various forms of physiotherapy in arthritis. *Physiotherapy*. 1959;45(6):139-142.

14. Sadura-Sieklucka T, Soltysiuk B, Karlicka A, Sokolowska B, Kontny E, Ksiezopolska-Orlowska K. Effects of whole body cryotherapy in patients with rheumatoid arthritis considering immune parameters. *Reumatologia*. 2019;57(6):320-325. doi.org/10.5114/reum.2019.90825

15. Klemm P, Hoffmann J, Asendorf T, et al. Whole-body cryotherapy for the treatment of rheumatoid arthritis: a monocentric, single-blinded, randomised controlled trial. *Clin Exp Rheum*. 2022:Online ahead of print. 10.55563/clinexprheumatol/lrff6k

PICO 26: Should patients with RA receive electrotherapy?

Summary**:** Literature searches identified three randomized controlled trial (RCT)[1, 3, 4] and one non-randomized comparison study [2] addressing this question. These studies investigated four modalities:

- Neuromuscular Electrical Stimulation (NMES) [1]. This is similar to electric muscle stimulation but geared towards rehabilitation use, and is known for muscle strengthening and activation
- Galvanic stimulation [3]. This is long duration interrupted direct current and is known to penetrate deep into tissue to reduce pain and improve circulation (e.g., motor point stimulation).
- Faradic stimulation [2]. This is short duration interrupted direct current and is similar to galvanic stimulation, except faradic currents are used for innervated muscle and galvanic currents are used for denervated muscle.
- Transcutaneous electrical nerve stimulation (TENS) [4] uses alternating current and may reduce acute and chronic pain

Below we discuss separately the evidence on these four modalities. The only modality for which effect sizes were computable was NMES (Table 1); the other studies did not report enough information to determine effect sizes, so their data appear in a separate table (Table 2)

NMES

Piva et al [1] included 59 participants (31 in NMES and 28 in the volitional training groups) aged 21 years or older diagnosed with RA for >5 years by a rheumatologist as per the American College of Rheumatology criteria, fluent in English, and able to ambulate independently without an assistive device. However, only 50 participants completed 4-month follow-up (24 in NMES and 26 in the volitional training groups)

- Patients were randomized to either NES-Training or volitional training.
- NMES Training was administered using an Infinity Plus portable NMES unit. Stimulus parameters were pulse rate of 75 pulses/second and pulse duration of 450 microseconds. Stimulus on/off time was 12-sec on (3-sec ramp up, 6-sec full contraction, 3-sec ramp down), and 48-sec off to minimize muscle fatigue (1-min cycle)
- Volitional training was done with exercise equipment and based on best evidence. The exercises targeted mainly the quadriceps muscles and included leg extension and leg press exercises on the respective machines.
- Both groups received 36 sessions over 16 weeks by physical therapists who were masked to participants' performance on outcome measures

Primary outcomes included changes in muscle function and performance-based physical function and patient-reported outcomes from baseline to 4 months (post intervention). Both groups experienced significant improvements in function from baseline to 4 months, but there were no statistically significant between group differences.

Galvanic stimulation

An RCT by Dulgeroglu et al [3] reported no significant between group differences in any outcomes at 2 weeks for 16 participants who received galvanic therapy and conservative hand exercises compared to 14 participants who received conservative hand exercises only. The study reported medians for outcome data, therefore the data do not appear in the GRADE table below.

Faradic stimulation

Hamilton et al [2] was a nonrandomized comparative study published in 1959 that found no statistically significant difference between any of the other modes of physiotherapy (diathermy, infrared radiation, hot wax and sham diathermy). However, they found improved walk times for patients with RA who received faradism to the quadriceps. The age of this study makes it less relevant.

TENS

An RCT by Abelson et al [4] reported that the 16 participants who received 1x/week TENS over 3 weeks experienced improvements for all 4 outcomes (lower resting pain, grip pain and higher grip strength power score and work score) by the end of 3 sessions of TENS (i.e., at 3 weeks) compared to 16 participants who received placebo (i.e., no stimulation). The study did not report dispersion (e.g., SDs) of outcome data, therefore the data do not appear in the GRADE table.

Two outcomes were reported by multiple RCTs.

- No statistically significant between-group differences (with very similar estimates of the minimal between-group difference) for HAQ scores
    - for NMES vs. Volitional training [1]
    - for galvanic stim + exs vs. exs only [3]
- No statistically significant between-group differences (with very similar estimates of the minimal between-group difference) for HAQ scores
    - For galvanic stim + exs vs. exs only [3]
    - For TENS vs. placebo [4]

Quality of evidence across critical outcomes: Low for NMES

Table 1: Data from Randomized Controlled Trials [1]

Question: NEMS compared to Volitional training for patients with RA

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NEMS	Volitional training	Relative (95% CI)	Absolute (95% CI)		
Function as inferred from Change in Stair climbing test (sec) from baseline to 4 months (lower values are better)												
1	randomised trials	not serious	not serious	serious	very serious[a]	none	24	26	-	MD **0.1 lower** (0.99 lower to 0.79 higher)	⨁◯◯◯ Very Low	Critical NS
Function as inferred from Change in Timed chair stand (sec) from baseline to 4 months (lower values are better)												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	24	26	-	MD **0.2 lower** (2.51 lower to 2.11 higher)	⨁◯◯◯ Very low	Critical NS
Change in Lower Extremity Functional Scale from baseline to 4 months (higher values are better)												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	24	26	-	MD **5 higher** (0.47 higher to 9.53 higher)	⨁⨁◯◯ Low	Critical NS
Change in HAQ from baseline to 4 months (lower values are better)												
1	randomised trials	not serious	not serious	not serious	very serious[a]	none	24	26	-	MD **0** (0.18 lower to 0.18 higher)	⨁⨁◯◯ Low	Critical NS
Function as inferred from Change in Gait speed (m/sec) from baseline to 4 months (higher values are better)												
1	randomised trials	not serious	not serious	serious	very serious[a]	none	24	26	-	MD **0.01 lower** (0.12 lower to 0.1 higher)	⨁◯◯◯ Very Low	Critical NS
Function as inferred from Change in Right-Single leg stance (sec) from baseline to 4 months (higher values are better)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NEMS	Volitional training	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious	very serious[a]	none	24	26	-	MD **1 lower** (3.63 lower to 1.63 higher)	⊕◯◯◯ Very Low	Critical NS
Function as inferred from Change in Left-Single leg stance (sec) from baseline to 4 months (higher values are better)												
1	randomised trials	not serious	not serious	serious	very serious[a]	none	24	26	-	MD **2.6 lower** (6.26 lower to 1.06 higher)	⊕◯◯◯ Very Low	Critical NS
Function as inferred from Change in Right-MVIC (Nm) from baseline to 4 months (higher values are better)												
1	randomised trials	not serious	not serious	serious	very serious[a]	none	24	26	-	MD **0** (9.64 lower to 9.64 higher)	⊕◯◯◯ Very Low	Critical NS
Function as inferred from Change in Left-MVIC (Nm) from baseline to 4 months (higher values are better)												
1	randomised trials	not serious	not serious	serious	very serious[a]	none	24	26	-	MD **0** (10.83 lower to 10.83 higher)	⊕◯◯◯ Very Low	Critical NS

Explanations

a – very small study resulting in wide confidence intervals
CI: confidence interval; **MD:** mean difference

Risk of bias was deemed not serious since most information is from studies at low or unclear risk of bias and potential limitations are unlikely to lower confidence in the estimate of effect. E.g. the study used a statistician generated the randomization sequence and the research coordinator, not involved with testing/treatment, randomized the participants through a web-based computer system after the baseline visit. Completion rate at 4-month follow-up in intervention group vs. volitional training group was 77% vs. 93%. The testers were masked to group assignment. While participants could not be masked to treatment assignment, they were instructed not to discuss treatment with the testers. Physical therapists masked to participants' performance on outcome measures delivered the interventions.

Table 2. Additional Data from RCT [3,4] and non-RCT [2]

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
1902, Dulgeroglu, 2016 [3]	RCT	5 weeks	Total n = 30 patients w RA 100% female; age 54 y+/- 11.2 (range 50-75 y)	16 participants were in intervention group: received Galvanic electrotherapy + conservative hand exercises 14 participants were in Control group: received only conservative hand exercises	**Timepoint is 2 weeks** **Data presented as Median (Min, Max) Change from Baseline** **<u>Changes scores = Scores at baseline – 2weeks</u>** *Negative means lower scores are better; positive means higher scores are better.* **Tender Joint Count (n) (positive) p=0.140** Intervention: 1.5 (-2, 9) Control: 0 (-3, 5) **Swollen Joint Count (n) (positive) p=0.823** Intervention: 0 (-7, 4) Control: 0 (-1, 5) **Patient Global Assessment (VAS 0-100) (mm) (positive) p=0.966** Intervention: 5 (-20, 70) Control: -2.5 (-20, 40) **Hand of pain (VAS 0-100) (mm) (positive) p=0.190** Intervention: 12.5 (-20, 60) Control: 0 (-20, 25) **Health Assessment Questionnaire (positive) p=0.601** Intervention: 0.20 (-0.13, 1.50) Control: 0.17 (0, 1.30) **Duruöz Hand Index (positive) p=0.692** Intervention: 7 (-2, 19) Control: 9 (-16, 26) **Deficit (cm) (positive)** **Flexion (R Hand) p=0.874** Intervention: 0 (0, 2.5) Control: 0 (-0.5, 1.6) **Flexion (L Hand) p=0.906** Intervention: 0 (0, 1.5) Control: 0 (0, 1) **Extension (R Hand) p=0.487**

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					Intervention: 0 (0, 3) Control: 0 (0, 3.1) **Extension (L Hand) p=0.457** Intervention: 0 (0, 3) Control: 0 (0, 2.3) **Opposition (R Hand) p=0.094** Intervention: 0 (0, 1.7) Control: 0 (0, 0) **Opposition (L Hand) p=0.094** Intervention: 0 (0, 1.2) Control: 0 (0, 0) **Range of Motion (degrees) (Negative)** **Wrist Palmar Flexion (R Hand) p=0.982** Intervention: -10 (-52, 10) Control: -10 (-34, 11) **Wrist Palmar Flexion (L Hand) p=0.287** Intervention: -10 (-50, 10) Control: -10 (-30, 20) **Wrist Dorsal Flexion (R Hand) p=0.502** Intervention: -7.5 (-30, 10) Control: -6.5 (-33, 5) **Wrist Dorsal Flexion (L Hand) p=0.966** Intervention: -10 (-45, 10) Control: -11 (-49, 15) **Grip Strength (kg) (Negative)** **R Hand p=0.307** Intervention: -0.42 (-5.64, 5.0) Control: 0 (-5.9, 6.34) **L Hand p=0.505** Intervention: -1.0 (-7.97, 3.66) Control: -1.82 (-5.3, 2.67) **Pinch Strength (kg) (Positive)** **R Hand p=0.429** Intervention: -0.2 (-3, 1.33) Control: -0.58 (-2.3, 0.36)

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					L Hand p=0.917 Intervention: -0.48 (-4.5, 1.16) Control: -0.3 (-2.16, 0.83) **9 hole peg test (sec) (Negative)** **R Hand p=0.308** Intervention: 0.9 (-1, 5.5) Control: 1.3 (-2.7, 5.39) **L Hand p=0.422** Intervention: -1.3 (-1, 5.5) Control: -0.91 (-1.8, 4.41)
6358, Abelson, 1983 [4]	RCT	3 weeks	Total n= 32 patients w RA with wrist involvement Intervention: 13/16 female; mean age: 55 y (range 35-68); mean disease duration 13 y (range 1-27) Control: 13/16 female; mean age 57 y (range 40-72); mean disease duration 12 y (1-32)	16 participants in Intervention group: 1x/week transcutaneous electrical nerve stimulation; 3 sessions total 16 participants in Control group: placebo	**Data presented as Mean Change from Baseline to 3 weeks** *Negative means lower scores are better; positive means higher scores are better.* **Resting Pain (mm) (negative)**; 0=severe pain and 100= no pain) Data reported: Mean Change from Baseline (Baseline data —> Intervention: 60.5 ± 24.6 mm; Control: 75.0 ± 24.7 mm) Summary: After 3 sessions of electrical nerve stimulation, the intervention group experienced a statistically significant reduction in pain (mean change from baseline of +41.67 mm) while the control group did not (mean change from baseline of -3.60 mm). **Grip Pain (mm) (negative)**; 0=severe pain and 100= no pain) Data reported: Mean Change from Baseline (Baseline data —> Intervention: 56.0 ± 24.5 mm; Control: 61.0 ± 27.1 mm) Summary: After 3 sessions of electrical nerve stimulation, the intervention group experienced a statistically significant reduction in pain (mean change from baseline of +31.37 mm]; while the control group also improved their pain scores (mean change from baseline of +23.64 mm), it was not statistically significant.

Ref ID, Author, year	Study type	Duration	Population Description	Treatment given to relevant population	Results
					Grip Strength Power Score (Watts) (positive) Data reported: Mean change from Baseline (Baseline data —> Intervention: 1.64 ± 1.50 Watts; Control: 1.91 ± 1.49 Watts) Summary: After 3 sessions of electrical nerve stimulation, the intervention group experienced a large improvement (p val not reported) in Grip Strength Power Score (mean change from baseline of +.74 Watts) whereas the control group showed a decline (mean change from baseline of -.25 Watts). Note, though the intervention group showed improvements at the end of each session, in between sessions their scores dropped close to baseline values. **Grip Strength Work Score (Joules) (positive)** Data reported: Mean change from Baseline (Baseline data —> Intervention: .82 ± 1.23 Joules; .69 ± .64 Joules) Summary: After 3 sessions of electrical nerve stimulation, the intervention group experienced great improvements (no p val reported) in Grip Strength Work Scores (mean change from baseline of +.14 Joules), whereas the control group declined (mean change from baseline of-.02 Joules). Note, though the intervention group showed improvements at the end of each session, in between sessions their scores dropped to baseline values.
1193 Hamilton 1959 [2]	Nonrandomized comparison study	20 weeks	18 patients with RA of the knee 33 patients with RA of the hand	6. Sham diathermy 7. Infrared treatment 8. Hot wax for RA of the hand 9. Faradism of the quadriceps for RA of the knee	Walk time at conclusion of 4 week faradism -2.34 standard errors of the mean when compared to the mean improvement seen across all treatment groups. All other measures had a difference of less than 2 SEM. NO no-treatment GROUP. Participants received 4 different types of treatments over time.

References:

1. **4051** Piva SR, Khoja SS, Toledo FGS, et al. Neuromuscular Electrical Stimulation Compared to Volitional Exercise for Improving Muscle Function in Rheumatoid Arthritis: A Randomized Pilot Study. Arthritis Care Res (Hoboken). 2019;71(3):352-361.
2. **1193** Hamilton DE, Bywaters EG, Please NW. A controlled trial of various forms of physiotherapy in arthritis. Physiotherapy. 1959;45(6):139-142.
3. **1902** Dülgeroğlu D, Bal A, Karaahmet Ö, Umay E, Noyan S, Çakcı A. The effectiveness of galvanic electrotherapy and a conservative hand exercise program in a rheumatoid hand: a randomized controlled trial. Turkish Journal of Physical Medicine & Rehabilitation/Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi. 2016 Jun 1;62(2).
4. **6358** Abelson, K., Langley, G.B., Sheppeard, H., Vlieg, M. and Wigley, R.D., 1983. Transcutaneous electrical nerve stimulation in rheumatoid arthritis. The New Zealand Medical Journal, 96(727), pp.156-158.

PICO 27. Should patients with RA receive chiropractic therapy?

No studies met inclusion criteria for this question.

PICO 28. Should patients with RA who are current smokers engage in a smoking cessation program?

No studies met inclusion criteria for this question.