Appendix F. Risk of bias assessments

All studies in this review were individually randomised, hence all assessments use the ROB 2 tools for trials with a parallel design. Assessments are presented in alphabetical order by study ID.

For each study, an assessment was done for each outcome and comparison contributing to the MA (or where results could not be included in the MA but were tabulated).

For each study we report

- the comparison for the assessment,
- the outcome domain for the assessment,
- other outcomes included in MAs for the study (noting if the assessment was the same for these or other comparisons),
- the study design (parallel trial)

Where the RoB assessment was the same for all outcomes, comparisons or both, only one assessment is reported.

The assessment includes

- The overall risk of bias judgement
- The judgement for each domain, with an explanation provided for each signalling questions for which the response could lead to a judgement of high risk of bias or some concerns
- The response to each signalling question (numbers, the questions are reported in full below)

We did not assess studies that were counted as 'missing results' (i.e. those studies where the result was judged to be uninterpretable or where there were major concerns about the integrity of the data such that it would be misleading to report the results). In such cases, concerns about bias leading to an under- or over-estimate of effect are inconsequential compared to the impact of major errors in reported data or the interpretation of that data.

Box F1. Signalling questions from the revised Cochrane risk of bias (ROB 2) tool for randomised trials (parallel design)

Parallel (individually randomised)
Domain 1. Bias arising from the randomisation process
1.1 Was the allocation sequence random?
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?
1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?
Domain 2. Bias due to deviations from intended interventions
2.1 Were participants aware of their assigned intervention during the trial?
2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?
2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?
2.4 If Y/PY to 2.3 Were these deviations likely to have affected the outcome?
2.5 If Y/PY to 2/4: Were these deviations from intended intervention balanced between groups?
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?
Domain 3. Bias due to missing outcome data
3.1 Were data for this outcome available for all, or nearly all, participants randomized?
3.2 If N/PN/NI to 3.1a or 3.1b: Is there evidence that the result was not biased by missing data?
3.3 If N/PN to 3.2 Could missingness in the outcome depend on its true value?
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?
Domain 4. Bias in the measurement of the outcome
4.1 Was the method of measuring the outcome inappropriate?
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?
Domain 5. Bias from selection of the reported result

Parallel (individually randomised)
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?
5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?
5.3 ... multiple eligible analyses of the data?

Appendix F. Risk of bias assessments – parallel randomised trials

Study ID. **Eardley 2013**	**Outcome domain.** Pain **Assessments**. Pain		**Comparison.** C1 inactive - sham **Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
1. Bias arising from the randomisation process	Some concerns	Volunteers were randomised to treatment group by a research assistant from a stack of coded sealed envelopes which had been prepared by the trial statistician and blocked in units of 9. If randomisation completed in blocks of 9 then for 70 people randomised the difference between groups shouldn't be 24, 21, 25. The M/F proportion was very different between groups (15%, 30% and 50%).	PY	PY	Y				
2. Bias due to deviations from the intended intervention	Low	McNemar's tests determined no significant difference between patient guess for treatment allocation at last treatment, week 5 (p = 0.17) indicating that blinding was secure. Modified intention-to-treat (mITT) analysis (excluding participants with missing outcome data)	PN	Y	PN	NA	NA	Y	NA
3. Bias due missing outcome data	Low	I: 20/24 (16% missing) C: 20/21 (4% missing) Reasons for loss to follow-up reported and unlikely due to true value of outcome.	N	N	PN	NA			
4. Bias in the measurement of the outcome	Low	Measure was self-report. McNemar's tests determined no significant difference between patient guess for treatment allocation at last treatment, week 5 (p = 0.17) indicating that blinding was secure.	N	PN	N	NA	NA		
5. Bias in the selection of the reported results	Low	No protocol or analysis plan, however the registry record shows pre-specified outcomes, measures that are fully reported in the study report for our selected timepoint (5 weeks). There is only one possible way in which the outcome can be measured (and at a single timepoint). Results are reported for multiple ways of analysing/handling the VAS, and it is unlikely that these were selected from other analyses.	PY	PN	PN				
OVERALL risk of bias	**Some concerns**								

Study ID. Eardley 2013	**Outcome domain.** Physical function (disability)		**Comparison.** C1 inactive - sham						
	Assessments. Physical function (disability)		**Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
1. Bias arising from the randomisation process	Some concerns	Volunteers were randomised to treatment group by a research assistant from a stack of coded sealed envelopes which had been prepared by the trial statistician and blocked in units of 9. If randomisation completed in blocks of 9 then for 70 people randomised the difference between groups shouldn't be 24, 21, 25. The M/F proportion was very different between groups (15%, 30% and 50%).	PY	PY	Y				
2. Bias due to deviations from the intended intervention	Low	McNemar's tests determined no significant difference between patient guess for treatment allocation at last treatment, week 5 (p = 0.17) indicating that blinding was secure. Modified intention-to-treat (mITT) analysis (excluding participants with missing outcome data)	PN	Y	PN	NA	NA	Y	NA
3. Bias due missing outcome data	Low	I: 20/24 (16% missing) C: 20/21 (4% missing) Reasons for loss to follow-up reported and unlikely due to true value of outcome.	N	N	PN	NA			
4. Bias in the measurement of the outcome	Low	Measure was self-report. McNemar's tests determined no significant difference between patient guess for treatment allocation at last treatment, week 5 (p = 0.17) indicating that blinding was secure.	N	PN	N	NA	NA		
5. Bias in the selection of the reported results	Low	No protocol or analysis plan, however the registry record shows pre-specified outcomes, measures that are fully reported in the study report for our selected timepoint (5 weeks). There is only one possible way in which the outcome can be measured (and at a single timepoint). Results are reported for multiple ways of analysing/handling the RMDQ data, and it is unlikely that these were selected from other analyses.	PY	PN	PN				
OVERALL risk of bias	**Some concerns**								

Study ID. Eardley 2013	**Outcome domain.** HR-QoL		**Comparison.** C1 inactive - sham						
	Assessments. HR-QoL, EFMH		**Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
1. Bias arising from the randomisation process	Some concerns	Volunteers were randomised to treatment group by a research assistant	PY	PY	Y				

Study ID. Eardley 2013	**Outcome domain.** HR-QoL **Assessments**. HR-QoL, EFMH		**Comparison.** C1 inactive - sham **Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
		from a stack of coded sealed envelopes which had been prepared by the trial statistician and blocked in units of 9. If randomisation completed in blocks of 9 then for 70 people randomised the difference between groups shouldn't be 24, 21, 25. The M/F proportion was very different between groups (15%, 30% and 50%).							
2. Bias due to deviations from the intended intervention	Low	McNemar's tests determined no significant difference between patient guess for treatment allocation at last treatment, week 5 (p = 0.17) indicating that blinding was secure. Modified intention-to-treat (mITT) analysis (excluding participants with missing outcome data)	PN	Y	PN	NA	NA	Y	NA
3. Bias due missing outcome data	Low	I: 20/24 (16% missing) C: 20/21 (4% missing) Reasons for loss to follow-up reported and unlikely due to true value of outcome.	N	N	PN	NA			
4. Bias in the measurement of the outcome	Low	Measure was self-report. McNemar's tests determined no significant difference between patient guess for treatment allocation at last treatment, week 5 (p = 0.17) indicating that blinding was secure.	N	PN	N	NA	NA		
5. Bias in the selection of the reported results	Low	No protocol or analysis plan, however the registry record shows pre-specified outcomes, measures that are fully reported in the study report for our selected timepoint (5 weeks). The triallists did not specify how the SF-36 would be reported, however it is a common approach to report physical and mental dimensions only, and not subdomains. Results are reported for multiple ways of analysing/handling the SF-36 data, and it is unlikely that these were selected from other analyses.	PY	PN	PN				
OVERALL risk of bias	**Some concerns**								

Study ID. Eardley 2013	**Outcome domain.** Pain **Assessments**. Pain		**Comparison.** C2 inactive - waitlist control **Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
1. Bias arising from the randomisation process	Some concerns	Volunteers were randomised to treatment group by a research assistant	PY	PY	Y				

Study ID. Eardley 2013	**Outcome domain.** Pain **Assessments**. Pain		**Comparison.** C2 inactive - waitlist control **Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
		from a stack of coded sealed envelopes which had been prepared by the trial statistician and blocked in units of 9. If randomisation completed in blocks of 9 then for 70 people randomised the difference between groups shouldn't be 24, 21, 25. The M/F proportion was very different between groups (15%, 30% and 50%).							
2. Bias due to deviations from the intended intervention	Low	Intervention group received kinesiology and comparator no intervention [wait list] (i.e. not a sham/placebo or 'active' standard care), so it is likely that participants were aware of their assigned intervention. The same poeple were involved in care for both arms and it is likely that they were aware of the participants' assigned intervention. Modified intention-to-treat (mITT) analysis (excluding participants with missing outcome data)	Y	Y	PN	NA	NA	Y	NA
3. Bias due missing outcome data	Low	I: 20/24 (17% missing) C: 17/25 (32% missing) A greater proportion of participants were missing from the wait list group and more withdrawals were due to participants discontinuing and not responding to requests for data. This could be because of pain worsening; however it is more likely due to lack of engagement because the wait list group did not receive any intervention.	PN	N	PN	NA			
4. Bias in the measurement of the outcome	High	Participants (i.e. the outcome assessors) were aware that they had received kinesiology or not intervention (wait list). Participants' knowledge of the intervention they received could have influenced their response. Participants may have had a prior belief about the benefits of kinesiology compared to no treatment that were likely to influence the outcome.	N	PN	Y	PY	PY		
5. Bias in the selection of the reported results	Low	No protocol or analysis plan, however the registry record shows pre-specified outcomes, measures that are fully reported in the study report for our selected timepoint (5 weeks). There is only one possible way in which the outcome can be measured (and at a single timepoint). Results are reported for multiple ways of analysing/handling the VAS data, and it is unlikely that these were selected from other analyses.	PY	PN	PN				
OVERALL risk of bias	**High**								

Study ID. **Eardley 2013**	**Outcome domain.** Physical function (disability) **Assessments**. Physical function (disability)		**Comparison.** C2 inactive - waitlist control **Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
1. Bias arising from the randomisation process	Some concerns	Volunteers were randomised to treatment group by a research assistant from a stack of coded sealed envelopes which had been prepared by the trial statistician and blocked in units of 9. If randomisation completed in blocks of 9 then for 70 people randomised the difference between groups shouldn't be 24, 21, 25. The M/F proportion was very different between groups (15%, 30% and 50%).	PY	PY	Y				
2. Bias due to deviations from the intended intervention	Low	Intervention group received kinesiology and comparator no intervention [wait list] (i.e. not a sham/placebo or 'active' standard care), so it is likely that participants were aware of their assigned intervention. The same poeple were involved in care for both arms and it is likely that they were aware of the participants' assigned intervention. Modified intention-to-treat (mITT) analysis (excluding participants with missing outcome data)	Y	Y	PN	NA	NA	Y	NA
3. Bias due missing outcome data	Low	I: 20/24 (17% missing) C: 17/25 (32% missing) A greater proportion of participants were missing from the wait list group and more withdrawals were due to participants discontinuing and not responding to requests for data. This could be because of disability worsening; however it is more likely due to lack of engagement because the wait list group did not receive any intervention.	PN	N	PN	NA			
4. Bias in the measurement of the outcome	High	Participants (i.e. the outcome assessors) were aware that they had received kinesiology or not intervention (wait list). Participants' knowledge of the intervention they received could have influenced their response. Participants may have had a prior belief about the benefits of kinesiology compared to no treatment that were likely to influence the outcome.	N	PN	Y	PY	PY		
5. Bias in the selection of the reported results	Low	No protocol or analysis plan, however the registry record shows pre-specified outcomes, measures that are fully reported in the study report for our selected timepoint (5 weeks). There is only one possible way in which the outcome can be measured (and at a single timepoint). Results are reported for multiple ways of analysing/handling the RMDQ, and it is	PY	PN	PN				

Study ID. **Eardley 2013**	**Outcome domain.** Physical function (disability) **Assessments**. Physical function (disability)		**Comparison.** C2 inactive - waitlist control **Design.** parallel (individually randomised)						
Domain	**Judgment**	**Explanation** (for concerns that lead to high or some concerns about RoB)	**Response to signalling questions**						
			SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7
		unlikely that these were selected from other analyses.							
OVERALL risk of bias	**High**								