Appendix contents list

Appendix H: GRADE profiles, subgroup and sensitivity analyses

Appendix A: Search strategies

Appendix A.1. Search strategy: Pharmacological interventions for children with chronic pain (April 2020)

Database searched	Date searched
CENTRAL (*The Cochrane Library)* Issue 4 of 12, 2020	9/4/20
MEDLINE & MEDLINE in Process (OVID) 1946 to April 6th 2020	8/4/20
EMBASE (OVID) 1980 to 2020 week 14	8/4/20

CENTRAL (*The Cochrane Library*)

#1 MeSH descriptor: [Child] explode all trees
#2 MeSH descriptor: [Infant] this term only
#3 MeSH descriptor: [Adolescent] this term only
#4 ((child* or infant* or baby or babies or preschooler* or pre-schooler* or toddler* or schoolchild* or girl* or boy* or adolescen* or teen*)):ti,ab,kw (Word variations have been searched)
#5 #1 or #2 or #3 or #4
#6 MeSH descriptor: [Analgesics, Opioid] explode all trees
#7 ((morphine or buprenorphine or codeine or dextromoramide or diphenoxylate or dipipanone or dextropropoxyphene or propoxyphene or diamorphine or dihydrocodeine or alfentanil or fentanyl or remifentanil or meptazinol or methadone or nalbuphine or oxycodone or papaveretum or pentazocine or meperidine or pethidine or phenazocine or hydrocodone or hydromorphone or levorphanol or oxymorphone or butorphanol or dezocine or sufentanil or ketobemidone)):ti,ab,kw (Word variations have been searched)
#8 MeSH descriptor: [Ketamine] this term only
#9 ((ketamine or ketalar or calipsol or ketanest)):ti,ab,kw (Word variations have been searched)
#10 ((ketaset or calypsol or kalipsol)):ti,ab,kw (Word variations have been searched)
#11 (ci-581):ti,ab,kw (Word variations have been searched)
#12 MeSH descriptor: [Anti-Inflammatory Agents, Non-Steroidal] explode all trees
#13 (NSAID*):ti,ab,kw (Word variations have been searched)
#14 ((ibuprofen or aspirin or naproxen or fenoprofen or ketoprofen or tiaprofenic acid or diclofenac or aceclofenac or etodolac or indometacin or mefenamic acid or meloxicam or nabumeton or phenylbutazone or piroxicam or sulindac or tenoxicam or tolfenamic acid or ketorolac or parecoxib or celecoxib or etoricoxib)):ti,ab,kw (Word variations have been searched)
#15 MeSH descriptor: [Acetaminophen] this term only
#16 ((acetaminophen or paracetamol or Panadol or Tylenol)):ti,ab,kw (Word variations have been searched)
#17 MeSH descriptor: [Antidepressive Agents] explode all trees
#18 ((amitriptyline or clomipramine or doxepin or imipramine or nortriptyline or trimipramine or mianserin or trazadone or citalopram or fluoxetine or fluvoxamine or sertraline)):ti,ab,kw (Word variations have been searched)
#19 MeSH descriptor: [Anticonvulsants] explode all trees

#20 (non-steroidal anti-inflammatory drug):ti,ab,kw (Word variations have been searched)
#21 (non-steroidal anti-inflammatory agent):ti,ab,kw (Word variations have been searched)
#22 MeSH descriptor: [Acetaminophen] this term only
#23 ((acetaminophen or paracetamol or Panadol or Tylenol)):ti,ab,kw (Word variations have been searched)
#24 MeSH descriptor: [Antidepressive Agents] explode all trees
#25 ((amitriptyline or clomipramine or doxepin or imipramine or nortriptyline or trimipramine or mianserin or trazadone or citalopram or fluoxetine or fluvoxamine or sertraline)):ti,ab,kw (Word variations have been searched)
#26 MeSH descriptor: [Anticonvulsants] explode all trees
#27 ((carbamazepine or clobazam or clonazepam or ethosuximide or gabapentin or lacosamide or lamotrigine or levetiracetam or oxcarbazepine or phenytoin or pregabalin or rufinamide or topiramate or valproate or vigabatrin or zonisamide)):ti,ab,kw (Word variations have been searched)
#28 MeSH descriptor: [Duloxetine Hydrochloride] this term only
#29 (duloxetine):ti,ab,kw (Word variations have been searched)
#30 (serotonin norepinephrine reuptake inhibitor):ti,ab,kw (Word variations have been searched)
#31 (SNRI):ti,ab,kw (Word variations have been searched)
#32 MeSH descriptor: [Milnacipran] this term only
#33 (milnacipran):ti,ab,kw (Word variations have been searched)
#34 (Flupirtine):ti,ab,kw (Word variations have been searched)
#35 (gabapentinoid*):ti,ab,kw (Word variations have been searched)
#36 (Indomethacin):ti,ab,kw (Word variations have been searched)
#37 MeSH descriptor: [Venlafaxine Hydrochloride] this term only
#38 (Venlafaxine):ti,ab,kw (Word variations have been searched)
#39 MeSH descriptor: [Desipramine] this term only
#40 (Desipramine):ti,ab,kw (Word variations have been searched)
#41 MeSH descriptor: [Tramadol] this term only
#42 (tramadol):ti,ab,kw (Word variations have been searched)
#43 MeSH descriptor: [Nefopam] explode all trees
#44 (Nefopam):ti,ab,kw (Word variations have been searched)
#45 #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44
#46 MeSH descriptor: [Fibromyalgia] this term only
#47 ((headache* or migraine* or fibromyalgia* or neuralgia*)):ti,ab,kw (Word variations have been searched)
#48 (pain):ti,ab,kw (Word variations have been searched)
#49 MeSH descriptor: [Pain] this term only
#50 MeSH descriptor: [Abdominal Pain] explode all trees
#51 MeSH descriptor: [Arthralgia] explode all trees
#52 MeSH descriptor: [Back Pain] explode all trees
#53 MeSH descriptor: [Breakthrough Pain] explode all trees
#54 MeSH descriptor: [Cancer Pain] this term only
#55 MeSH descriptor: [Chest Pain] explode all trees
#56 MeSH descriptor: [Chronic Pain] this term only

#57 MeSH descriptor: [Earache] this term only
#58 MeSH descriptor: [Eye Pain] this term only
#59 MeSH descriptor: [Facial Pain] explode all trees
#60 MeSH descriptor: [Flank Pain] this term only
#61 MeSH descriptor: [Glossalgia] this term only
#62 MeSH descriptor: [Headache] explode all trees
#63 MeSH descriptor: [Mastodynia] this term only
#64 MeSH descriptor: [Metatarsalgia] this term only
#65 MeSH descriptor: [Musculoskeletal Pain] explode all trees
#66 MeSH descriptor: [Neck Pain] explode all trees
#67 MeSH descriptor: [Neuralgia] explode all trees
#68 MeSH descriptor: [Nociceptive Pain] explode all trees
#69 MeSH descriptor: [Pelvic Pain] explode all trees
#70 MeSH descriptor: [Renal Colic] this term only
#71 MeSH descriptor: [Pain, Referred] this term only
#72 MeSH descriptor: [Pain, Intractable] this term only
#73 #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57 or #58 or #59 or #60 or #61 or #62 or #63 or #64 or #65 or #66 or #67 or #68 or #69 or #70 or #71 or #72
#74 #5 and #45 and #73

MEDLINE (OVID)
1. *child/ or *child, preschool/
2. *Infant/
3. *Adolescent/
4. (child* or infant* or baby or babies or preschooler* or pre-schooler* or toddler* or schoolchild* or girl* or boy* or adolescen* or teen*).tw.
5. 1 or 2 or 3 or 4
6. exp Analgesics, Opioid/
7. (morphine or buprenorphine or codeine or dextromoramide or diphenoxylate or dipipanone or dextropropoxyphene or propoxyphene or diamorphine or dihydrocodeine or alfentanil or fentanyl or remifentanil or meptazinol or methadone or nalbuphine or oxycodone or papaveretum or pentazocine or meperidine or pethidine or phenazocine or hydrocodone or hydromorphone or levorphanol or oxymorphone or butorphanol or dezocine or sufentanil or ketobemidone).ti,ab.
8. Ketamine/
9. (ketamine or ketalar or calipsol or ketanest).ti,ab.
10. (ketaset or calypsol or kalipsol).ti,ab.
11. ci-581.tw.
12. exp Anti-Inflammatory Agents, Non-Steroidal/
13. NSAID*.tw.
14. "non-steroidal anti-inflammatory drug*".tw.
15. (ibuprofen or aspirin or naproxen or fenoprofen or ketoprofen or tiaprofenic acid or diclofenac or aceclofenac or etodolac or indometacin or mefenamic acid or meloxicam or nabumeton or phenylbutazone or piroxicam or sulindac or tenoxicam or tolfenamic acid or ketorolac or parecoxib or celecoxib or etoricoxib).tw.
16. "non-steroidal anti-inflammatory agent*".tw.
17. Acetaminophen/
18. (acetaminophen or paracetamol or Panadol or Tylenol).tw.

19. exp Antidepressive Agents/
20. (amitriptyline or clomipramine or doxepin or imipramine or nortriptyline or trimipramine or mianserin or trazadone or citalopram or fluoxetine or fluvoxamine or sertraline).tw.
21. exp Anticonvulsants/
22. (carbamazepine or clobazam or clonazepam or ethosuximide or gabapentin or lacosamide or lamotrigine or levetiracetam or oxcarbazepine or phenytoin or pregabalin or rufinamide or topiramate or valproate or vigabatrin or zonisamide).tw.
23. Duloxetine Hydrochloride/
24. duloxetine.tw.
25. serotonin norepinephrine reuptake inhibitor.mp.
26. SNRI.tw.
27. Milnacipran/
28. milnacipran.tw.
29. Flupirtine.tw.
30. gabapentinoid*.tw.
31. Indomethacin.tw.
32. Venlafaxine Hydrochloride/
33. Venlafaxine.tw.
34. Desipramine/
35. Desipramine.tw.
36. Tramadol/
37. tramadol.tw.
38. Nefopam/
39. Nefopam.tw.
40. or/6-39
41. randomized controlled trial.pt.
42. controlled clinical trial.pt.
43. randomized.ab.
44. placebo.ab.
45. drug therapy.fs.
46. randomly.ab.
47. trial.ab.
48. exp Clinical Trials as topic/
49. Cross-Over Studies/
50. PLACEBOS/
51. Research Design/
52. latin square.tw.
53. Comparative Study/
54. Evaluation Studies/
55. or/41-54
56. exp animals/ not humans.sh.
57. 55 not 56
58. pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or breakthrough pain/ or cancer pain/ or exp chest pain/ or chronic pain/ or earache/ or eye pain/ or facial pain/ or flank pain/ or glossalgia/ or exp headache/ or mastodynia/ or metatarsalgia/ or exp musculoskeletal pain/ or exp neck pain/ or neuralgia/ or exp nociceptive pain/ or pain, intractable/ or pain, referred/ or exp pelvic pain/ or renal colic/
59. pain.tw.

60. (headache* or migraine* or fibromyalgia* or neuralgia*).tw.
61. Fibromyalgia/
62. 58 or 59 or 60 or 61
63. 5 and 40 and 57 and 62

EMBASE (OVID)
1. *child/ or *child, preschool
2. *Infant/
3. *Adolescent/
4. (child* or infant* or baby or babies or preschooler* or pre-schooler* or toddler* or schoolchild* or girl* or boy* or adolescen* or teen*).tw.
5. 1 or 2 or 3 or 4
6. exp Analgesics, Opioid/
7. (morphine or buprenorphine or codeine or dextromoramide or diphenoxylate or dipipanone or dextropropoxyphene or propoxyphene or diamorphine or dihydrocodeine or alfentanil or fentanyl or remifentanil or meptazinol or methadone or nalbuphine or oxycodone or papaveretum or pentazocine or meperidine or pethidine or phenazocine or hydrocodone or hydromorphone or levorphanol or oxymorphone or butorphanol or dezocine or sufentanil or ketobemidone).ti,ab.
8. *Ketamine/
9. (ketamine or ketalar or calipsol or ketanest).ti,ab.
10. (ketaset or calypsol or kalipsol).ti,ab.
11. ci-581.tw.
12. exp Anti-Inflammatory Agents, Non-Steroidal/
13. NSAID*.tw.
14. "non-steroidal anti-inflammatory drug*".tw.
15. (ibuprofen or aspirin or naproxen or fenoprofen or ketoprofen or tiaprofenic acid or diclofenac or aceclofenac or etodolac or indometacin or mefenamic acid or meloxicam or nabumeton or phenylbutazone or piroxicam or sulindac or tenoxicam or tolfenamic acid or ketorolac or parecoxib or celecoxib or etoricoxib).tw.
16. "non-steroidal anti-inflammatory agent*".tw.
17. *Acetaminophen/
18. (acetaminophen or paracetamol or Panadol or Tylenol).tw.
19. exp Antidepressive Agents/
20. (amitriptyline or clomipramine or doxepin or imipramine or nortriptyline or trimipramine or mianserin or trazadone or citalopram or fluoxetine or fluvoxamine or sertraline).tw.
21. exp Anticonvulsants/
22. (carbamazepine or clobazam or clonazepam or ethosuximide or gabapentin or lacosamide or lamotrigine or levetiracetam or oxcarbazepine or phenytoin or pregabalin or rufinamide or topiramate or valproate or vigabatrin or zonisamide).tw.
23. *Duloxetine Hydrochloride/
24. duloxetine.tw.
25. serotonin norepinephrine reuptake inhibitor.mp.
26. SNRI.tw.
27. *Milnacipran/
28. milnacipran.tw.
29. Flupirtine.tw.
30. gabapentinoid*.tw.
31. Indomethacin.tw.

32. *Venlafaxine Hydrochloride/
33. Venlafaxine.tw.
34. *Desipramine/
35. Desipramine.tw.
36. *Tramadol/
37. tramadol.tw.
38. *Nefopam/
39. Nefopam.tw.
40. or/6-39
41. *pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or *breakthrough pain/ or *cancer pain/ or exp chest pain/ or *chronic pain/ or *earache/ or *eye pain/ or *facial pain/ or *flank pain/ or *glossalgia/ or exp headache/ or *mastodynia/ or *metatarsalgia/ or exp musculoskeletal pain/ or exp neck pain/ or *neuralgia/ or exp nociceptive pain/ or *pain, intractable/ or exp pain, postoperative/ or *pain, referred/ or exp pelvic pain/ or *renal colic/
42. pain.tw.
43. (headache* or migraine* or fibromyalgia* or neuralgia*).tw.
44. Fibromyalgia/
45. 41 or 42 or 43 or 44
46. 5 and 40 and 45
47. random$.tw.
48. factorial$.tw.
49. crossover$.tw.
50. cross over$.tw.
51. cross-over$.tw.
52. placebo$.tw.
53. (doubl$ adj blind$).tw.
54. (singl$ adj blind$).tw.
55. assign$.tw.
56. allocat$.tw.
57. volunteer$.tw.
58. Crossover Procedure/
59. double-blind procedure.tw.
60. Randomized Controlled Trial/
61. Single Blind Procedure/
62. placebo/
63. methodology/
64. latin square.tw.
65. comparative study/
66. evaluation study/
67. or/47-66
68. (animal/ or nonhuman/) not human/
69. 67 not 68
70. 46 and 69

Appendix A.2. Search Overview Physical interventions for children with chronic pain (April 2020)

Database searched	Date searched
CENTRAL (*The Cochrane Library)* Issue 4 of 12 2020	14/4/20
MEDLINE & MEDLINE in Process (OVID) 1946 to April 13 2020	14/4/20
EMBASE (OVID) 1980 to 2020 week 15 2020	14/4/20

CENTRAL (*The Cochrane Library*)

#1 MeSH descriptor: [Physical Education and Training] this term only
#2 MeSH descriptor: [Exercise Therapy] explode all trees
#3 ((exercise* or physical activit*)):ti,ab,kw (Word variations have been searched)
#4 MeSH descriptor: [Exercise Therapy] explode all trees
#5 MeSH descriptor: [Muscle Stretching Exercises] this term only
#6 #4 not #5
#7 MeSH descriptor: [Physical Therapy Modalities] explode all trees
#8 (physiotherap*):ti,ab,kw (Word variations have been searched)
#9 (physical therap*):ti,ab,kw (Word variations have been searched)
#10 (manipulative therapy):ti,ab,kw (Word variations have been searched)
#11 (((therapeutic or therapy) Near/2 exercise)):ti,ab,kw (Word variations have been searched)
#12 ("graded motor imagery"):ti,ab,kw (Word variations have been searched)
#13 (mirror therapy):ti,ab,kw (Word variations have been searched)
#14 MeSH descriptor: [Musculoskeletal Manipulations] explode all trees
#15 (hydrotherapy):ti,ab,kw (Word variations have been searched)
#16 ((pain Near/3 (advice or education))):ti,ab,kw (Word variations have been searched)
#17 ((flexibility Near/2 exercise*)):ti,ab,kw (Word variations have been searched)
#18 MeSH descriptor: [Yoga] this term only
#19 (yoga):ti,ab,kw (Word variations have been searched)
#20 MeSH descriptor: [Tai Ji] this term only
#21 ((tai chi or tai ji)):ti,ab,kw (Word variations have been searched)
#22 MeSH descriptor: [Qigong] this term only
#23 (Qigong):ti,ab,kw (Word variations have been searched)
#24 (ch'i kung):ti,ab,kw (Word variations have been searched)
#25 MeSH descriptor: [Mind-Body Therapies] this term only
#26 #1 or #2 or #3 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25
#27 MeSH descriptor: [Child] explode all trees
#28 MeSH descriptor: [Infant] this term only
#29 MeSH descriptor: [Adolescent] this term only
#30 ((child* or infant* or baby or babies or preschooler* or pre-schooler* or toddler* or schoolchild* or girl* or boy* or adolescen* or teen*)):ti,ab,kw (Word variations have been searched)
#31 #27 or #28 or #29 or #30
#32 (pain):ti,ab,kw (Word variations have been searched)

#33 ((headache* or migraine* or fibromyalgia* or neuralgia*)):ti,ab,kw (Word variations have been searched)
#34 MeSH descriptor: [Fibromyalgia] this term only
#35 MeSH descriptor: [Pain] this term only
#36 MeSH descriptor: [Abdominal Pain] explode all trees
#37 MeSH descriptor: [Arthralgia] explode all trees
#38 MeSH descriptor: [Back Pain] explode all trees
#39 MeSH descriptor: [Breakthrough Pain] this term only
#40 MeSH descriptor: [Cancer Pain] this term only
#41 MeSH descriptor: [Chest Pain] explode all trees
#42 MeSH descriptor: [Chronic Pain] this term only
#43 MeSH descriptor: [Earache] this term only
#44 MeSH descriptor: [Eye Pain] this term only
#45 MeSH descriptor: [Facial Pain] this term only
#46 MeSH descriptor: [Facial Pain] this term only
#47 MeSH descriptor: [Glossalgia] this term only
#48 MeSH descriptor: [Headache] explode all trees
#49 MeSH descriptor: [Mastodynia] this term only
#50 MeSH descriptor: [Metatarsalgia] this term only
#51 MeSH descriptor: [Musculoskeletal Pain] explode all trees
#52 MeSH descriptor: [Neck Pain] explode all trees
#53 MeSH descriptor: [Sciatic Neuropathy] this term only
#54 MeSH descriptor: [Pain, Intractable] this term only
#55 MeSH descriptor: [Pain, Referred] this term only
#56 MeSH descriptor: [Pelvic Pain] explode all trees
#57 MeSH descriptor: [Renal Colic] this term only
#58 #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57
#59 #26 and #31 and #58

MEDLINE (OVID)
1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. exp Clinical Trials as topic/
9. Cross-Over Studies/
10. PLACEBOS/
11. Research Design/
12. latin square.tw.
13. Comparative Study/
14. Evaluation Studies/
15. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14
16. exp animals/ not humans.sh.
17. 15 not 16

18. "physical education and training"/
19. exp Exercise Therapy/
20. (exercise* or physical activit*).tw.
21. exp Exercise/ not Muscle Stretching Exercises/
22. exp Physical Therapy Modalities/
23. physiotherap*.tw.
24. "physical therap*".tw.
25. manipulative therapy.tw.
26. ((therapeutic or therapy) adj2 exercise).tw.
27. "graded motor imagery".tw.
28. mirror therapy.tw.
29. exp Musculoskeletal Manipulations/
30. hydrotherapy.tw.
31. (pain adj3 (advice or education)).tw.
32. (flexibility adj2 exercise*).tw.
33. (mobility adj2 exercise*).tw.
34. Yoga/
35. yoga.tw.
36. Tai Ji/
37. (tai chi or tai ji).tw.
38. Qigong/
39. Qigong.tw.
40. ch'i kung.tw.
41. Mind-Body Therapies/
42. *child/ or *child, preschool/
43. *Infant/
44. *Adolescent/
45. (child* or infant* or baby or babies or preschooler* or pre-schooler* or toddler* or schoolchild* or girl* or boy* or adolescen* or teen*).tw.
46. 42 or 43 or 44 or 45
47. pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or breakthrough pain/ or cancer pain/ or exp chest pain/ or chronic pain/ or earache/ or eye pain/ or facial pain/ or flank pain/ or glossalgia/ or exp headache/ or mastodynia/ or metatarsalgia/ or exp musculoskeletal pain/ or exp neck pain/ or neuralgia/ or exp nociceptive pain/ or pain, intractable/ or pain, referred/ or exp pelvic pain/ or renal colic/
48. pain.tw.
49. (headache* or migraine* or fibromyalgia* or neuralgia*).tw.
50. Fibromyalgia/
51. 47 or 48 or 49 or 50
52. or/18-41
53. 17 and 46 and 51 and 52

EMBASE (OVID)

1. "physical education and training"/
2. exp Exercise Therapy/
3. (exercise* or physical activit*).tw.
4. exp Exercise/ not Muscle Stretching Exercises/
5. exp Physical Therapy Modalities/
6. physiotherap*.tw.

7. "physical therap*".tw.
8. manipulative therapy.tw.
9. ((therapeutic or therapy) adj2 exercise).tw.
10. "graded motor imagery".tw.
11. mirror therapy.tw.
12. exp Musculoskeletal Manipulations/
13. hydrotherapy.tw.
14. (pain adj3 (advice or education)).tw.
15. (flexibility adj2 exercise*).tw.
16. (mobility adj2 exercise*).tw.
17. Yoga/
18. yoga.tw.
19. Tai Ji/
20. (tai chi or tai ji).tw.
21. Qigong/
22. Qigong.tw.
23. ch'i kung.tw.
24. Mind-Body Therapies/
25. *child/ or *child, preschool/
26. *Infant/
27. *Adolescent/
28. (child* or infant* or baby or babies or preschooler* or pre-schooler* or toddler* or schoolchild* or girl* or boy* or adolescen* or teen*).tw.
29. 25 or 26 or 27 or 28
30. pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or breakthrough pain/ or cancer pain/ or exp chest pain/ or chronic pain/ or earache/ or eye pain/ or facial pain/ or flank pain/ or glossalgia/ or exp headache/ or mastodynia/ or metatarsalgia/ or exp musculoskeletal pain/ or exp neck pain/ or neuralgia/ or exp nociceptive pain/ or pain, intractable/ or pain, referred/ or exp pelvic pain/ or renal colic/
31. pain.tw.
32. (headache* or migraine* or fibromyalgia* or neuralgia*).tw.
33. Fibromyalgia/
34. 30 or 31 or 32 or 33
35. random$.tw.
36. factorial$.tw.
37. crossover$.tw.
38. cross over$.tw.
39. cross-over$.tw.
40. placebo$.tw.
41. (doubl$ adj blind$).tw.
42. (singl$ adj blind$).tw.
43. assign$.tw.
44. allocat$.tw.
45. volunteer$.tw.
46. Crossover Procedure/
47. double-blind procedure.tw.
48. Randomized Controlled Trial/
49. Single Blind Procedure/
50. placebo/

51. methodology/
52. latin square.tw.
53. comparative study/
54. evaluation study/
55. or/35-54
56. (animal/ or nonhuman/) not human/
57. 55 not 56
58. or/1-24
59. 29 and 34 and 57 and 58

Appendix A.3. Search Overview Psychological interventions for children with chronic pain (April 2020)

Database searched	Date searched
CENTRAL (*The Cochrane Library)* Issue 3 of 12, 2020	16/3/20
MEDLINE & MEDLINE in Process (OVID) 1946 to March 2020	16/3/20
EMBASE (OVID) 1980 to March 2020	16/3/20
PsycINFO (EBSCO) to March 2020	16/3/20

For the psychological therapies search, we updated the following Cochrane systematic reviews:

1. **Fisher, E.,** Law, E., Dudeney, J., Palermo, T.M., & Eccleston, C. (2019). Psychological therapies (remotely delivered) for the management of chronic and recurrent pain in children and adolescents. *Cochrane Database of Systematic Reviews*, Issue 4. Art. No: CD011118. DOI: 10.1002/14651858.CD011118.pub3.
2. **Fisher, E.,** & Law, E., Dudeney, J., Palermo, T.M., Steward, G., & Eccleston, C. (2018). Psychological therapies for the management of chronic and recurrent pain in children and adolescents. *Cochrane Database of Systematic Reviews, 9*, CD003968. DOI: 10.1002/14651858.
3. Law, E. F., **Fisher, E.,** Eccleston, C., & Palermo, T. M. (2019). Psychological therapies for parents of children and adolescents with chronic illness. *Cochrane Database of Systematic Reviews,* Issue 3. Art. No: CD009660. DOI: 10.1002/14651858.CD009660.pub4.

We also ran a search for children and cancer-related pain from inception.

Appendix C. Additional results

Pharmacological trials

Risk of bias judgements can be found in Appendix D. Risk of bias figures for RCTs can be found in Figures 4 & 7. ROBINS-I judgements can be found in Appendix D.

Random sequence generation (checking for possible selection bias): We rated seven studies as low risk of random sequence generation. The remaining studies did not describe their method of randomisation and we judged these as unclear.

Allocation concealment (checking for possible selection bias): Four studies described their allocation concealment method, and we judged these studies as low risk of bias. The remaining studies did not describe allocation concealment and so we judged these studies as unclear risk of bias.

Blinding of personnel and participants (checking for possible detection bias): We rated 13 studies as low risk of bias for blinding of personnel and participants. These studies provided a convincing methodology of blinding personnel and participants to the assigned group. We judged 15 studies as unclear risk of bias; these studies did not provide a convincing methodology. One study was rated as high risk of bias, as the drugs differed in appearance.

Blinding of outcome assessment (checking for possible detection bias). We rated 11 studies as low risk of bias for blinding outcome assessment. We rated 17 studies as unclear, and one study as high risk of bias which did not mention blinding and delivered in number and appearance of drugs.

Incomplete outcome data (checking for possible attrition bias due to the amount, nature and handling of incomplete outcome data): We rated 11 studies as low risk of bias; these studies reported less than 10% attrition. We rated 13 studies as unclear risk of bias; these studies reported more than 10% attrition and used last observation carried forward or were unclear about their data imputation. The remaining five studies reported more the 10% attrition and used completer analyses, so were rated as high risk of bias.

Selective reporting (checking for reporting bias): We rated four studies as low risk of bias; these studies reported a trial registration and reported all outcomes. Five studies were judged to be high risk of bias, where the manuscript differed to the protocol. The remaining 20 studies did not report a protocol or trial registration, or it was not available.

For non-randomised comparative studies with a control group, we used the Cochrane Risk of Bias in Non-randomized Studies - of Interventions (ROBINS-I). For non-randomised comparative studies where we considered the bias due to confounding to be “serious” or “critical”, the overall risk of bias for the study was also considered “serious” or “critical” and other domains were not assessed. We considered the most important confounders to be age, sex, baseline pain intensity and co-interventions. We rated five studies using the ROBINS criteria (Appendix D). We found all five studies were critical risk of bias and confounding variables were not controlled for in the analyses.

Physical therapies

We did not judge physical therapy interventions for blinding of participants and personnel. Risk of bias judgements for the 24 RCTs can be found in Appendix C and D. Risk of bias figures can be found in Figures 5 & 8. One trial registry (non-

comparative study) was not rated for risk of bias as there was insufficient evidence to rate it from the trial registration.

Random sequence generation (checking for possible selection bias): We rated 18 studies as low risk of bias for random sequence generation. These studies provided a convincing method of randomisation. The remaining six studies did not provide a clear description, and we rated these studies as unclear risk of bias.

Allocation concealment (checking for possible selection bias): Similar to randomisation, we found 13 studies that provided a convincing method of allocation concealment, and we rated these studies as low risk of bias. Eleven studies did not provide adequate detail, and we rated these studies as unclear risk of bias.

Blinding of outcome assessment (checking for possible detection bias): We rated two studies as low risk of bias for blinding outcome assessment. We rated 22 studies as unclear.

Incomplete outcome data (checking for possible attrition bias due to the amount, nature and handling of incomplete outcome data): We rated eight studies as low risk of bias; these studies reported less than 10 % attrition or used baseline observation carried forwards. Eight studies were rated as unclear risk of bias; these studies reported more than 10% attrition or used last observation carried forward. The remaining eight studies used completer analyses and reported more than 10% attrition, as so we rated these studies as high risk of bias.

Selective reporting (checking for reporting bias): We rated seven studies as low risk of bias; these studies reported a prospective trial registration and reported outcomes in the manuscript. Eleven studies did not report a trial registration or protocol, or one was not available, and we rated these studies as unclear risk of bias. We rated six studies as high risk of bias. These studies incompletely reported data or the outcomes did not match the trial registration.

Psychological trials

We did not judge psychological trials for blinding of participants and personnel. Risk of bias judgements can be found in Appendix D. Risk of bias figures can be found in Figures 6 & 8.

Random sequence generation (checking for possible selection bias): We judged 28 studies as low risk of bias, which provided a convincing method of randomisation. We judged the remaining 35 studies as unclear risk of bias

Allocation concealment (checking for possible selection bias): We judged 23 studies as low risk of bias as they provided a convincing method of allocation concealment. Two studies were rated as high risk of bias, and we judged the remaining 38 studies as unclear risk of bias, as they did not describe how participants were allocated to the trial arms.

Blinding of outcome assessment (checking for possible detection bias): We rated 22 studies as low risk of bias for blinding of outcome assessors and 41 studies as unclear risk of bias. No studies were rated as high risk of bias. Most studies did not have a description of how outcomes were assessed.

Incomplete outcome data (checking for possible attrition bias due to the amount, nature and handling of incomplete outcome data): We rated 20 studies as low risk of bias for incomplete outcome data. These studies had a low attrition rate (<10%) or used baseline observation carried forwards. We rated 24 studies as unclear risk of bias, as they did not describe how they imputed missing data or used last observation carried forwards. We rated 19 studies as high risk of bias, most of which used completer analyses despite >10% dropout.

Selective reporting (checking for reporting bias): We rated seven studies as low risk of bias. These studies had a pre-registered protocol and reported all outcomes from the pre-registration in their manuscripts. We rated 38 studies as unclear risk of bias, these studies did not report a protocol or trial registration. We rated 18 studies as high risk of bias; these studies did not report all outcomes from their trial registration in the manuscript.

Description of superiority analyses; cross-over and non-randomised studies

Critical outcomes

Pain intensity: Two parallel RCTs did not report data, but narratively reported that sumatriptan reduced migraines compared to placebo[1] and a further study reported no difference between amitriptyline compared to placebo.[2]

One cross-over trial[3] reported that zolmitriptan and ibuprofen showed similar pain reduction for children with migraine, and that this was superior compared to placebo post-treatment and at follow-up. Another crossover trial,[4] reported that participants in the acetylsalicylic acid group reported significantly greater pain reduction compared to the control group. A third cross-over trial[5] reported reduced migraines in the zolmitriptan group compared to the placebo group. Four remaining cross-over trials did not provide data. One reported no difference between fluoxetine and placebo for children with chronic headache[6] and another showed no difference between montelukast vs placebo for menstrual symptoms.[7] Two trials reported sumatriptan naproxen[8] or progestin[9] improved symptoms in migraine and abdominal pain respectively, compared to placebo.

Finally, two non-randomised studies with no data reported no differences for reducing pain between amitriptyline and relaxation[10] or between mefenamic acid and fennel extract.[11]

In one study[12], we found no beneficial effect of 30% pain reduction for psychological therapies post-treatment or at follow-up (RR 1.13, 95% CI 0.64 to 2.02; RR 1.07, 95% CI 0.77 to 1.49, respectively). We rated both outcomes as very low-certainty, downgraded twice for very serious limitations to study design and imprecision, and once for indirectness.

HRQOL: One pharmacological study with 33 participants reported the amitriptyline treatment group were more likely to improve quality of life from baseline to post-treatment and follow-up, compared to placebo.[13] We rated this outcome as very low-certainty at post-treatment and follow-up, downgraded once for serious limitation to study design and indirectness, and twice for very serious imprecision.

Functional disability: one pharmacological study comparing antidepressants (duloxetine) to placebo post-treatment for functional disability showed no beneficial effect. We rated this outcome as very low-certainty, downgraded for serious limitations to study design, indirectness, and very serious imprecision. No studies reported functional disability at follow-up. No studies reported functional disability at follow-up.

Role functioning: One pharmacological cross-over trial (29 participants[4]) reported fewer school absences compared to baseline in the treatment group compared to the control group (very low-certainty evidence, downgraded for serious limitations to study design, indirectness, and very serious imprecision). One physical therapy study reported the number of absences from school post-treatment[14] and another study with (43 participants) reported participation in school.[15] There were no differences reported between groups post-treatment for either study. We rated this outcome as very low-certainty, downgraded twice for limitations for study design and imprecision, and once for indirectness. At follow-up, one study reported no differences between groups[15] on role/social physical functioning. Again, we judged this as very low-certainty, downgraded twice for limitations for study design and imprecision.

Sleep: One pharmacological study (104 participants) comparing anticonvulsants (pregabalin) to placebo for sleep outcomes post-treatment did not find a beneficial effect (SMD -0.09, 95% CI -0.47 to 0.30). We rated this outcome as very low, downgraded once for indirectness and twice for very serious imprecision. No pharmacological studies reported data at follow-up, and there were no data we could analyse for physical therapy studies at either time-point.

Adverse events: One cross-over trial comparing zolmitriptan, ibuprofen and placebo reported no SAEs in the trial.[3]

There were four cross-over studies that reported AEs. The first reported significantly more AEs in the zolmitriptan compared to placebo, but ibuprofen did not differ to placebo.[3] A second study reported 9/31 participants in the fluoxetine and 3/29 participants in the placebo group reported AEs.[6] Four participants in the fluoxetine group stopped receiving the drug. Two other studies did not report any AEs in either group (dydrogesterone or montelukast vs placebo).[7,9]

Four physical therapy studies with 161 participants reported adverse events in their trials. Two studies[16,17] reported no AEs during testing or training sessions. A separate study[18] reported two children had accidental injuries, five reported joint pain, two reported somatic symptoms, and one reported another illness. The authors did not report which group these AEs occurred, but these AEs were not associated with participation in the study. However, children participating in physical exercise in the treatment group, reported muscle soreness associated with learning new exercises, which was typically resolved within a couple of days. One study also reported one AE in the treatment group (n = 18) and none in the control group (n = 14).[19] We rated AEs as very low-certainty, downgraded due to serious imprecision and indirectness. No treatment-related SAEs or other AEs were reported across physical therapy trials.

For psychological trials, AEs, SAEs and other AEs were poorly reported. We found 5 studies reported no adverse events (SAEs or treatment-related AEs) in any condition. One study reported more AEs in the control group (education + amitriptyline) compared to control, and most were attributed to amitriptyline. A final study reported mild headache in the treatment group when listening to CDs. We rated this certainty of evidence as low-certainty, downgraded once each for indirectness and imprecision.

Activity participation: One non-randomised pharmacological study[20] reported activity participation post-treatment and reported no differences between children receiving

citalopram and placebo post-treatment (very low-certainty). No studies reported activity participation at follow-up. One physical therapy study (63 participants)[14] reported fewer absences from physical activity in the exercise group compared to the control group (very low-certainty). No psychological studies assessed activity participation post-treatment, and one study (44 participants) reported a beneficial effect of psychological therapies at improving activity participation at follow-up. We rated the post-treatment outcome as very low, downgraded once for limitations to study design, indirectness, and twice for imprecision.

Global satisfaction with treatment: Across pharmacological trials, one study (490 participants) reported a higher percentage of subjects treated with sumatriptan and naproxen versus placebo reported being satisfied/very satisfied for 'how effective the medication is overall' and 'overall satisfaction with medication' at 2- and 24-hours post-dose ($p = 0.014$). Two further studies (205 participants[20,21]) comparing antidepressants (citalopram and amitriptyline) to placebo did not find any differences between groups in the intention-to-treat analyses (low-certainty). At follow-up, one study comparing citalopram to placebo with 115 participants[20] did not find any differences between groups ($p = 0.491$; very low-certainty). We rated post-treatment as low-certainty of evidence; we downgraded once for inconsistency and indirectness. At follow-up, we rated the certainty of evidence as very low; downgraded once for indirectness and twice for imprecision. No physical therapy studies reported global satisfaction with treatment at post-treatment or follow-up.

Patient global impression of change: We found one study with 104 participants reported participants in the pregabalin groups reported significant improvement compared to placebo ($p = 0.013$) with 53.1% of subjects much improved or very much improved at endpoint (very low-certainty, downgraded once for serious limitations to study design, indirectness, and twice for imprecision).[22] No studies reported the outcome at follow-up. One physical therapy trial (42 participants) reported that 18/21 participants in the treatment group reported a 'slight but noticeable change', and 10/21 reported a 'definite improvement that has made a real and worthwhile difference'. In the waitlist control group, only one participant agreed with either of the categories, which we rated as very low-certainty.[16] We downgraded this outcome twice for imprecision and once for indirectness. One psychological therapy trial (143 participants) assessed patient global impression of change post-treatment and at follow-up. The study reported participants in the psychological therapy reported a greater global impression of change at both time points[23] (very low-certainty, downgraded twice for imprecision and once for indirectness).

Effects of pharmacological therapies; subgroup analysis of two pharmacological interventions

Forest plots for the below analyses are shown in Appendix G.1. For comprehensiveness, we have included pharmacological intervention vs. non-pharmacological control and pharmacological intervention vs another pharmacological intervention in the forest plots. The individual drugs, pain condition, and age of participants are included in Table 1.

Pain intensity
We found two studies that compared two NSAIDs post-treatment (529 participants, rofecoxib vs naproxen and meloxicam vs. naproxen), and no beneficial effect was found (SMD -0.10, 95% CI -0.38 to 0.18, I^2 58%, very low-certainty). At post-treatment, one study with 34 participants showed no difference between anticonvulsants (gabapentin) and antidepressants (amitriptyline; SMD -0.17, 95% CI -0.85 to 0.50, very low-certainty). One study (300 participants) comparing mefenamic acid plus vitamin E to mefenamic acid showed beneficial effects for the former group (SMD -2.55, 95% CI -2.85 to -2.24, very low-certainty). At follow-up, one study comparing antidepressants vs. anticonvulsants (57 participants; SMD 2.96, 95% CI 2.19 to 3.72, very low-certainty) and one study comparing two NSAIDs (meloxicam vs. naproxen, 225 participants; SMD 0.03, 95% CI -0.25 to 0.30, very low-certainty). Neither study showed a superior beneficial effect.

No studies compared two pharmacological interventions and assessed 30% or 50% pain reduction, at either time-point.

Health-related quality of life
We found one study with 303 participants that could be entered into a health-related quality of life, post-treatment analysis comparing two types of NSAIDs (rofecoxib vs naproxen). This analysis did not show a beneficial effect of either treatment for improving health-related quality of life, very low-certainty evidence. No studies reported follow-up data that could be entered into a meta-analysis.

A second study [24] compared naproxen and two doses of Celecoxib (3mg/kg; 6mg/kg) and reported improvements in all groups, but no significant differences.
A non-randomised study [25] showed greater health-related quality of life benefits for participants in the steroid group compared to participants in the NSAID and Methotrexate group.

Functional disability
Three studies comparing two NSAIDs (770 participants, celecoxib, rofecoxib and meloxicam vs. naproxen) reported on functional disability, post-treatment, but no beneficial effect was found (SMD -0.08, 95% CI -0.07 to 0.23, I^2 0%, very low-certainty). At follow-up, one study (225 participants, meloxicam vs. naproxen) comparing two NSAIDs did not find any beneficial effects, very low-certainty.

Role functioning
No studies reported role functioning as a separate outcome, that was not included in overall health-related quality of life assessments.

Emotional functioning
We analysed studies assessing changes in depression and anxiety across trials. For depression, we found one study (225 participants) compared two NSAIDs which did not show a beneficial effect on reducing depression (meloxicam vs. naproxen, SMD 0.00, 95% CI -0.27 to 0.28, very low-certainty). At follow-up, the same study (225 participants) comprising two NSAIDs also failed to show any beneficial effect (SMD -0.05, 95% CI -0.33 to 0.22, very low-certainty)

We found no studies comparing two pharmacological interventions for anxiety at either time point.

Treatment-related serious adverse events
We could conduct one subgroup analyses investigating SAEs; NSAID vs. NSAID (two studies comparing rofecoxib or meloxicam vs. naproxen, 535 participants) and there were no differences between groups (RD 0.02, 95% CI -0.05 to 0.10, very low-certainty).

Treatment-related adverse events
We could conduct two subgroup analyses; anticonvulsants vs antidepressants (two studies comparing amitriptyline to gabapentin or topiramate, 91 participants) and NSAID vs. NSAID (five studies, 801 participants comparing celebcoxib, piroxicam, rofecoxib, and meloxicam to naproxen, and aspirin vs. ibuprofen). Neither analysis showed a higher number to adverse events (anticonvulsants vs. antidepressants RD -0.04, 95% CI -0.16 to 0.07, very low-certainty; NSAIDs vs NSAIDs -0.04, 95% CI -0.12 to 0.03, I^2 55%, very low-certainty).

Moran et al., (1979) [26] reported 1/23 participants in the naproxen group compared to 6//23 participants in the aspirin group reported an AE. Leak et al., (1988) [27] reported AEs in the tolmetin group (3/29), naproxen group (3/29) and diclofenac group (6/29). Price et al., (1985) [28] reported seven patient reported gastrointestinal symptoms during the study, but it was not clear which drug they were associated with (indoprofen, aloxiprin). One patient withdrew from the study but the authors did not report if this was linked to the drugs. Soriani et al., (2001) [29] reported no significant difference in AEs between groups taking acetaminophen and nimesulide.

Other adverse events
One study comparing two NSAIDs (225 participants, meloxicam vs. naproxen) did not find any difference between groups when assessing other types of AEs (RD 0.91, 95% CI 0.80 to 1.03, very low-certainty).

Sleep
A study comparing two anticonvulsants found no differences between gabapentin and amitriptyline [30].

Secondary outcomes
Activity participation
No studies compared two pharmacological interventions to each other and assessed activity participation.

Global judgement of satisfaction with treatment

One study with 28 participants compared naproxen, tolmetin and diclofenac on global judgement of satisfaction. The study reported 9 participants preferred naproxen, 8 preferred tolmetin and 6 preferred diclofenac. Five participants had not preferences [27].

Patient global impression of change
One study with 46 participants compared ibuprofen to aspirin and found 22/26 in the ibuprofen group and 18/20 in the aspirin group rated themselves as improved. There were no differences between groups.

Fatigue
No studies reported data on fatigue at post-treatment or at follow-up.

Physical therapies vs. other physical therapies

Of the 13 studies that compared two physical therapy arms, there were eight studies that we could enter into a meta-analysis investigating how physical therapies compare to each other. We included TOAT video based games vs TOAT daily living conditions;[31] Resistive underwater exercises and interferential current vs standard physical therapy;[32] land physiotherapy vs combined hydrotherapy and land physiotherapy;[33] targeted exercise group vs generalised physiotherapy;[34] Pilates vs. conventional exercise;[17] hypermobile range group vs neural control group (included similar exercises;[35] aerobics vs Qigong;[36] and unsupervised vs. supervised home exercise programme.[37] NCT03046472[38] was a non-randomised trial report with results, but with no attached peer-reviewed publication. This trial compared physical therapy for postural behaviour and daily home exercise vs physical therapy for postural behaviour and daily home exercise plus a group exercise class once per week. We describe this study separately and did not include it in the meta-analyses.

Please note, the first mentioned intervention above was entered as the 'experimental' condition in the meta-analysis and the second intervention was entered as the 'control' intervention in the analyses. We did not conduct certainty of evidence assessments on these outcomes. Forest plots for the below analyses are shown in Appendix G.2.

Pain intensity

Of the eight studies and 305 participants we could enter into a forest plot on pain intensity post-treatment, we found four studies showed a beneficial effect [17,32,36,37] and four did not.[31,34,35,39] At follow-up, three studies with 94 participants showed only one study with a beneficial effect.[32]

A further, non-randomised study[38] reported lower back pain in the group that received physical therapy for postural behaviour and daily home exercise plus a group exercise class once per week vs those who only received physical therapy for postural behaviour and daily home exercise, post-treatment.

Health-related quality of life

Three studies including 154 participants reported on health-related quality of life and could be entered into a forest plot. One showed a significant effect post-treatment,[17] and the remaining two studies did not.[33,36] No studies reported at follow-up which could be entered into a meta-analysis.

Functional disability

Five studies with 180 participants reported functional disability post-treatment, and two studies with 64 participants reported at follow-up, which we entered into a forest plot. We found two studies were beneficial at reducing functional disability post-treatment,[17,36] two studies were not beneficial [34,35] and one study favoured the "control".[37] At follow-up, neither study that reported follow-up data showed beneficial effects.[34,37]

Role functioning (e.g., school attendance)

One study with 25 participants reported participation in school and activities between groups having physical therapy three times/week vs once/week. No differences between groups were found at post-treatment.

Emotional functioning
Only one study reported depression outcomes post-treatment[36] and this did not show a beneficial effect of either treatment.

Sleep
We found no studies that reported sleep outcomes at post-treatment or follow-up.
Treatment-related adverse events, serious adverse events, and other adverse events
Two studies with 96 participants delivering different forms of physical exercise reported not adverse events during testing or training sessions.[17,40] NCT03046472[38] also reported no adverse events from treatments.

Secondary outcomes
Activity participation
One study[36] with 30 participants reported a higher number of hours involved in activity in participants in the aerobics group compared to the Qigong group. The authors reported no differences between baseline and post-treatment, but differences between groups were not reported.

Global judgement of satisfaction with treatment
No studies reported global satisfaction with treatment.

Patient global impression of change
Two studies reported patient global impression of change at post-treatment.[35,37] Neither study found differences between groups in relation to global impression of change.

Fatigue
No studies assessed fatigue separate from health-related quality of life.

Psychological therapies: Subgroup analyses

Following our protocol, we conducted subgroup analyses on outcomes that included more than 10 studies. We had initially planned to conduct a subgroup analysis by age of participants; however, this was not possible. All studies included children and adolescents and did not present data separately. The average age was 12.8 years. We also made several post-hoc decisions in order to help aid the GDG decision-making. We initially planned only to investigate route of intervention in pharmacological studies. As there were a number of remotely delivered psychological trials, we also conducted a subgroup analysis by route of intervention. It also seemed pertinent to analyse studies by therapy type, to help the GDG in recommending any specific types of therapies in their recommendations. Finally, there were no studies that we could include in our planned sensitivity analysis that included more than 200 participants/arm. Therefore, we ran a subgroup analysis of trials with less or more than 20 participants/arm. In summary, we conducted the following subgroup analyses:

- Control type (active; placebo)
- Chronic pain condition following the ICD-11 classification
- Dose/duration of treatment
- Route of delivery
- Therapy type
- Size

We conducted subgroup analyses on 10 outcomes that included more than 10 studies; pain intensity post-treatment and follow-up, 50% reduction in pain post-treatment, functional disability post-treatment and follow-up, health-related quality of life post-treatment, and emotional functioning (depression and anxiety) post-treatment and follow-up. The remaining outcomes did not include 10 or more studies and therefore were not included in any subgroup analysis. GRADE profiles are provided for each subgroup analysis. Reasons for downgrading are included in the GRADE profiles, but are not included here, in the interests to brevity.

Subgroup analysis: by control type

We analysed studies by active or standard care control and waitlist control. There were 52 active control arms and 17 waitlist control across the included studies. The GRADE evidence profile is shown in Appendix H and forest plots in Appendix G.3.

Overall, we found beneficial effects for the same outcomes in both active and waitlist control, as we found in the main analyses. However, we found that active control analyses were more similar to the overall effect size, most likely because they included the majority of studies. Therefore, certainty ratings were similar to the certainty ratings of the overall effect. We found the waitlist control subgroup analyses were rated mostly as very low-certainty, mainly because they included few studies and had serious limitations to study designs.

We found beneficial effects of psychological therapies versus active control for the outcomes of pain intensity post-treatment (low-certainty), 50% reduction in pain post-treatment (low-certainty), and functional disability post-treatment (moderate-certainty) and at follow-up (high-certainty). No other beneficial effects were found for pain at follow-up (low-certainty), health-related quality of life post-treatment (moderate-certainty), or emotional distress at either time point (depression was rated high-

certainty at both time points; anxiety was rated moderate-certainty post-treatment and high-certainty at follow-up).

For waitlist control, we found the same pattern of results. Psychological therapies were beneficial compared to waitlist control for the outcomes of pain intensity post-treatment (low-certainty), 50% reduction in pain post-treatment (very low-certainty), and functional disability post-treatment and at follow-up (very low-certainty). No other beneficial effects were found for pain at follow-up (very low-certainty), health-related quality of life post-treatment (very low-certainty), or emotional distress (very low-certainty post-treatment, no studies available at follow-up).

Subgroup analysis: by pain condition

We categorised trials by the pain condition, according to the ICD-11 classification. There were three subgroup analyses that included most studies and therefore we can draw conclusions from: chronic primary visceral pain, mixed pain conditions, and non-chronic headache. Non-chronic headache included studies of children with headache, but did not meet the IHS criteria for a chronic headache condition. Beneficial effects followed the same pattern as the main findings in subgroups with sufficient data. We could only include a limited number of studies for most subgroups, and therefore it is not possible to conclude if psychological therapies are more beneficial for any particular pain condition compared to another. The GRADE profile is shown in Appendix H and forest plots in Appendix G.4. We did not include analyses with two or less studies in the GRADE profile due to length, but all analyses were rated as very low, downgraded due to very serious imprecision.

Psychological therapies were beneficial for children with chronic primary visceral pain (10 studies, 844 participants, very low-certainty), mixed pain conditions (12 studies, 968 participants, moderate-certainty), but not for children with non-chronic headache (10 studies, 574 participants, low-certainty) at reducing pain intensity. Analyses investigating children with chronic secondary visceral pain and chronic widespread pain showed beneficial effects, but included a maximum of two studies (both very low-certainty). Analyses of children with secondary musculoskeletal pain and chronic secondary headache or orofacial pain also only included a maximum of two studies, but did not show a beneficial effect (very low-certainty). Lack of effect is most likely to be due to lack of data. No subgroup analyses showed a beneficial effect at follow-up for pain intensity.

Chronic primary headache or orofacial pain (low-certainty), mixed pain conditions (moderate-certainty), and children with non-chronic headache (low-certainty) provided data that could be included in the analysis assessing 50% pain reduction. All three subgroups showed a beneficial effect of psychological interventions compared to control.

Beneficial effects for disability were found in subgroup analyses of participants with chronic primary visceral pain (low-certainty) and mixed chronic pain conditions (moderate-certainty), post-treatment. No other subgroups showed beneficial effects post-treatment or at follow-up.

Beneficial effects were found in the primary chronic visceral pain subgroup for health-related quality of life (very low-certainty). Psychological therapies were not beneficial for other pain conditions.

There were no other beneficial effects of subgroup analyses for the other outcomes showing that psychological therapies do not work more effectively for one particular pain group compared to another.

Subgroup analysis: by treatment duration (dose)
For duration of treatment, we calculated the median hours of treatment delivered. Of the 63 full texts included, 46 reported treatment duration. The remaining 26 studies either reported sessions with no duration or did not report duration of treatment. The median duration of treatment from the studies reported was 4 hours. Therefore, treatments 4 hours and less were analysed separately to 5 hours and more. Studies where we could not calculate a duration are grouped in an 'unknown' category, which we did not conduct GRADE certainty of evidence ratings for or report here. The GRADE profile is shown in Appendix H and forest plots in Appendix G.5.

Overall, we did not find conclusive results that shorter or longer treatment duration was more favourable across multiple outcome domains. We found longer duration of treatment showed benefits for reducing pain intensity post-treatment (low-certainty) and achieving 50% reduction in pain post-treatment (very low-certainty). No other outcomes showed beneficial effects for longer treatment duration.

Shorter treatment duration showed benefits for reducing 50% pain reduction (very low-certainty), functional disability post-treatment and at follow-up (moderate-certainty). The remaining outcomes on pain intensity post-treatment and follow-up, health-related quality of life post-treatment, and emotional functioning at either time-point did not show beneficial effects for shorter treatment duration.

Certainty of evidence ranged from very low to high, following a similar pattern to the certainty of ratings of the main analyses.

Subgroup analysis: by delivery mode (route)
We analysed studies by whether they delivered treatment face-to-face or remote from the therapist. Remotely delivered therapies (14 studies) are most often delivered via the internet or smartphone, but have also been delivered via CD ROM or manuals. It is important to recognise that the evidence regarding remote therapies is smaller, and all subgroup analyses including remote therapies included fewer studies (although not always fewer participants) compared to remote therapist. Another theme that emerged is that subgroups of remotely delivered treatments were rated either the same or higher certainty of evidence, compared to face-to-face therapies. We believe that remote therapies could be utilised and improve symptoms in children. The GRADE profile is shown in Appendix H and forest plots in Appendix G.6.

We found that face-to-face therapies were beneficial at reducing pain intensity post-treatment (low-certainty), reducing pain by 50% or more post-treatment (very low-certainty), and reducing disability post-treatment (low-certainty) and at follow-up (moderate certainty). There were no beneficial findings for reducing emotional distress post-treatment or at follow-up. Conversely, remote therapies were also beneficial at

reducing pain intensity post-treatment (moderate-certainty) and 50% pain reduction (low-certainty) but no other beneficial effects were found.

Subgroup and sensitivity analysis: by therapy classification
We originally planned to analyse studies by individual therapy types, using classifications of cognitive behavioural therapy, acceptance commitment therapy, hypnosis, and relaxation. These results are summarised Appendix G.7 and Appendix H. The WHO GDG requested a sensitivity analysis of the combined effects of CBT, relaxation, BT and ACT for outcomes, which we performed and can be found in Appendix G.8. The WHO GDG also requested a sensitivity analysis of the combined effects of CBT, relaxation, BT and ACT by route (face-to-face vs. remote) for outcomes, which we performed and can be found in Appendix G.9.

We found small beneficial effects for combined CBT on the following outcomes; pain intensity post-treatment (low-certainty), 50% pain reduction post-treatment and follow-up (low-certainty and very low-certainty, respectively), functional disability post-treatment (low-certainty) and at follow-up (moderate-certainty). We did not find beneficial effects of CBT for pain at follow-up (low-certainty) and emotional functioning (depression: moderate-certainty post-treatment, high-certainty follow-up; anxiety: moderate-certainty post-treatment, high-certainty follow-up). For important outcomes, we found few studies could be included in analyses. In one study, there was a beneficial effect for activity participation at follow-up (very low-certainty), global satisfaction with treatment post-treatment (6 studies, moderate-certainty) and at follow-up (1 study, very low-certainty), patient global impression of change post-treatment and follow-up (1 study, very low-certainty). No data was reported for fatigue outcomes and no other beneficial effects were found.

The WHO GDG also requested analyses combined CBT separated by remote or face-to-face delivery. These findings were very similar to those presented for the subgroup analysis on delivery. Face-to-face delivery was beneficial at reducing pain intensity post-treatment (low-certainty), reducing 50% reduction pain intensity post-treatment and follow-up (very low-certainty), functional disability post-treatment and follow-up (low-certainty and moderate-certainty, respectively), and global satisfaction with treatment post-treatment (very low-certainty). No benefit of combined CBT therapies were found for reducing pain intensity at follow-up (low-certainty), 30% reduction in pain intensity post-treatment and follow-up (very low-certainty), health-related quality of life post-treatment and follow-up (very low-certainty), role functioning (very low-certainty), emotional functioning (moderate to high-certainty), or sleep quality (very low-certainty). No data were available for other outcomes.

For remote therapies, we found beneficial effects for reducing pain intensity post-treatment (moderate-certainty), 50% reduction in pain post-treatment (very low-certainty), global satisfaction with treatment post-treatment and at follow-up (low and very low-certainty), patient global impression of change post-treatment and at follow-up (very low-certainty). Please note only one study could be included in the analyses of global satisfaction with treatment and patient global impression of change. No beneficial effect was found for pain intensity (low-certainty) and 50% reduction in pain at follow-up (very low-certainty), health-related quality of life post-treatment and follow-up (moderate-certainty), and functional disability (moderate-certainty), role functioning (moderate-certainty post-treatment, very low-certainty follow-up), emotional

functioning (moderate to high-certainty) sleep quality (low-certainty), at either time-point. There were no data available for analysis for the remaining outcomes.

Subgroup analysis: by size

There were no studies that included more than 200 participants per arm at post-treatment, meaning we could not conduct our a-priori size sensitivity analysis. The largest study at post-treatment included 265 participants in total [41]. Therefore, we made a post-hoc decision to run a subgroup analysis of studies with more or less than 20 participants per arm. Both treatment arms had to include more than 20 participants per arm to be included as a 'larger study' in the analysis. There were 23 studies with at least one arm including less than 20 participants. The GRADE profile is shown in Appendix H and forest plots in Appendix G.10.

Overall, larger studies were consistently produced smaller effects and were rated as higher certainty evidence compared to smaller studies. We found smaller studies consistently had poorer study quality, larger confidence intervals and smaller number of participants, reducing our overall confidence in the estimates of effect.

For larger studies, we found small beneficial effects for pain intensity post-treatment (low-certainty), 50% reduction in pain intensity (very low-certainty), and small beneficial effects for reducing functional disability post-treatment (moderate-certainty) and at follow-up (high-certainty). We did not find any other beneficial effects for reducing pain intensity at follow-up (low-certainty), health-related quality of life post-treatment (high-certainty), or emotional distress at either time point (all high-certainty). We found beneficial effects for smaller studies; we found moderate beneficial effects for pain intensity post-treatment (very low-certainty), 50% reduction in pain intensity (very low-certainty), moderate beneficial effects for reducing functional disability post-treatment and large effect at follow-up (both very low-certainty). No beneficial effects were found for health-related quality of life or emotional distress at either timepoint (all very low-certainty).

Psychological therapies: Sensitivity analyses

We planned to conduct a sensitivity analysis only including low risk of bias studies. However, we rated only two studies as low risk of bias across all domains [42,43]. Therefore, we did not complete subgroup analyses on these two trials alone.

Sensitivity analysis excluding headache and migraine studies

We ran a sensitivity analysis excluding studies of children with migraine, tension-type headache, or migraine. We continued to include children with chronic headaches and migraines.

We found beneficial effects for the same outcomes as analyses run with the headache and migraine studies included. There were no notable differences between analyses including non-chronic headache studies and those that did not, and there were no notable differences in the certainty ratings for outcomes. The results are summarised in Appendix H and forest plots in Appendix G.11.

References

1. Derosier FJ, Lewis D, Hershey AD, et al. Randomized trial of sumatriptan and naproxen sodium combination in adolescent migraine. *Pediatrics* 2012; **129**(6): e1411-20.
2. Saps MYNMANSHPCJDLC. Multicenter, randomized, placebo-controlled trial of amitriptyline in children with functional gastrointestinal disorders. *Gastroenterology* 2009; **137**(4 XCP1 - PAIN CNO - CN-00719092): 1261-9.
3. Evers S, Rahmann A, Kraemer C, et al. Treatment of childhood migraine attacks with oral zolmitriptan and ibuprofen. *Neurology* 2006; **67**(3 XCP1 - PAIN/HEADACHE CNO - CN-00565864): 497-9.
4. Klein JR, Litt IF, Rosenberg A, Udall L. The effect of aspirin on dysmenorrhea in adolescents. *Journal of pediatrics* 1981; **98**(6 CNO - CN-00025016): 987-90.
5. Lewis DW, Winner P, Hershey AD, Wasiewski WW, Adolescent Migraine Steering C. Efficacy of zolmitriptan nasal spray in adolescent migraine. *Pediatrics* 2007; **120**(2): 390-6.
6. Gherpelli JL, Esposito SB. A prospective randomized double blind placebo controlled crossover study of fluoxetine efficacy in the prophylaxis of chronic daily headache in children and adolescents. *Arquivos de neuro-psiquiatria* 2005; **63**(3A XCP1 - PAIN/HEADACHE CNO - CN-00552272): 559-63.
7. Harel Z, Riggs S, Vaz R, Flanagan P, Harel D. The use of the leukotriene receptor antagonist montelukast (Singulair) in the management of dysmenorrhea in adolescents. *Journal of pediatric and adolescent gynecology* 2004; **17**(3 XCP1 - PAIN CNO - CN-00481787): 183-6.
8. Winner P, Pearlman EM, Linder SL, et al. Topiramate for migraine prevention in children: a randomized, double-blind, placebo-controlled trial. *Headache* 2005; **45**(10): 1304-12.
9. Gould CH. Dydrogesterone in teenage dysmenorrhoea. A multicentre trial in general practice. *Practitioner* 1979; **222**(1331 XCP1 - PAIN CNO - CN-00222958): 718-23.
10. Andrasik F, Grazzi L, Usai S, Bussone G. Pharmacological treatment compared to behavioural treatment for juvenile tension-type headache: results at two-year follow-up. *Neurological sciences* 2007; **28 Suppl 2**: S235-8.
11. Modaress Nejad V, Asadipour M. Comparison of the effectiveness of fennel and mefenamic acid on pain intensity in dysmenorrhoea. *La revue de sante de la Mediterranee orientale / al-Majallah al-sihhiyah li-sharq al-mutawassit [Eastern Mediterranean health journal]* 2006; **12**(3-4 XCP1 - PAIN CNO - CN-00572794): 423-7.
12. Van Der Veek SMC, Derkx HHF, Benninga MA, Boer F, De Haan E. Cognitive behavioural therapy for pediatric functional abdominal pain: a randomized controlled trial. *Pediatrics* 2013; **132**(5): e1163-e72.
13. Bahar RJ, Collins BS, Steinmetz B, Ament ME. Double-blind Placebo-Controlled Trial of Amitriptyline for the Treatment of Irritable Bowel Syndrome in Adolescents. *Journal of Pediatrics* 2008; **152**(5): 685-9.
14. Jones M, Stratton G, Reilly T, Unnithan V. The efficacy of exercise as an intervention to treat recurrent nonspecific low back pain in adolescents. *Pediatric exercise science* 2007; **19 RTY - Journal article**(3): 349-59.
15. Sandstedt E, Fasth A, Eek MN, Beckung E. Muscle strength, physical fitness and well-being in children and adolescents with juvenile idiopathic arthritis and the effect of an exercise programme: a randomized controlled trial. *Pediatric rheumatology online journal* 2013; **11 RTY - Journal article**(1): 7.

16. Andias R, Neto M, Silva AG. The effects of pain neuroscience education and exercise on pain, muscle endurance, catastrophizing and anxiety in adolescents with chronic idiopathic neck pain: a school-based pilot, randomized and controlled study. *Physiotherapy theory and practice* 2018; **34 RTY - Journal article**(9): 682-91.
17. Mendonca TM, Terreri MT, Silva CH, et al. Effects of Pilates exercises on health-related quality of life in individuals with juvenile idiopathic arthritis. *Archives of physical medicine and rehabilitation* 2013; **94 RTY - Journal article**(11): 2093-102.
18. Kashikar-Zuck S, Black WR, Pfeiffer M, et al. Pilot Randomized Trial of Integrated Cognitive-Behavioral Therapy and Neuromuscular Training for Juvenile Fibromyalgia: the FIT Teens Program. *Journal of pain* 2018; **19 RTY - Journal article**(9): 1049-62.
19. Evans S, Lung KC, Seidman LC, Sternlieb B, Zeltzer LK, Tsao JC. Iyengar yoga for adolescents and young adults with irritable bowel syndrome. *J Pediatr Gastroenterol Nutr* 2014; **59**(2): 244-53.
20. Roohafza H, Pourmoghaddas Z, Saneian H, Gholamrezaei A. Citalopram for pediatric functional abdominal pain: a randomized, placebo-controlled trial. *Neurogastroenterol Motil* 2014; **26**(11): 1642-50.
21. Saps M, Youssef N, Miranda A, et al. Multicenter, randomized, placebo-controlled trial of amitriptyline in children with functional gastrointestinal disorders. *Gastroenterology* 2009; **137**(4): 1261-9.
22. Arnold LM, Schikler KN, Bateman L, et al. Safety and efficacy of pregabalin in adolescents with fibromyalgia: a randomized, double-blind, placebo-controlled trial and a 6-month open-label extension study. *Pediatr Rheumatol Online J* 2016; **14**(1): 46.
23. Palermo TM, Vega R, Murray C, Law E, Zhou C. A digital health psychological intervention (WebMAP Mobile) for children and adolescents with chronic pain: results of a hybrid effectiveness-implementation stepped wedge cluster randomized trial. *Pain* 2020.
24. Foeldvari I, Szer IS, Zemel LS, et al. A prospective study comparing celecoxib with naproxen in children with juvenile rheumatoid arthritis. *Journal of rheumatology* 2009; **36**(1 CNO - CN-00687185): 174-82.
25. Riddle R, Ryser CN, Morton AA, et al. The impact on health-related quality of life from non-steroidal anti-inflammatory drugs, methotrexate, or steroids in treatment for juvenile idiopathic arthritis. *Journal of pediatric psychology* 2006; **31**(3): 262-71.
26. Moran H, Hanna DB, Ansell BM, Hall M, Engler C. Naproxen in juvenile chronic polyarthritis. *Annals of the rheumatic diseases* 1979; **38**(2).
27. Leak AM, Richter MR, Clemens LE, Hall MA, Ansell BM. A crossover study of naproxen, diclofenac and tolmetin in seronegative juvenile chronic arthritis. *Clinical and experimental rheumatology* 1988; **6**(2): 157-60.
28. Price T, Venning H, Ansell BM. Indoprofen in juvenile chronic arthritis. *Clinical and experimental rheumatology* 1985; **3**(1): 59-62.
29. Soriani S, Battistella PA, Naccarella C, Tozzi E, Fiumana E, Fanaro S. Nimesulide and acetaminophen for the treatment of juvenile migraine: A study for comparison of efficacy, safety, and tolerability. *Headache Quarterly* 2001; **12**(4): 233-6.
30. Brown S, Johnston B, Amaria K, et al. A randomized controlled trial of amitriptyline versus gabapentin for complex regional pain syndrome type I and neuropathic pain in children. *Scandinavian journal of pain* 2016; **13**: 156-63.
31. Arman N, Tarakci E, Tarakci D, Kasapcopur O. Effects of Video Games-Based Task-Oriented Activity Training (Xbox 360 Kinect) on Activity Performance

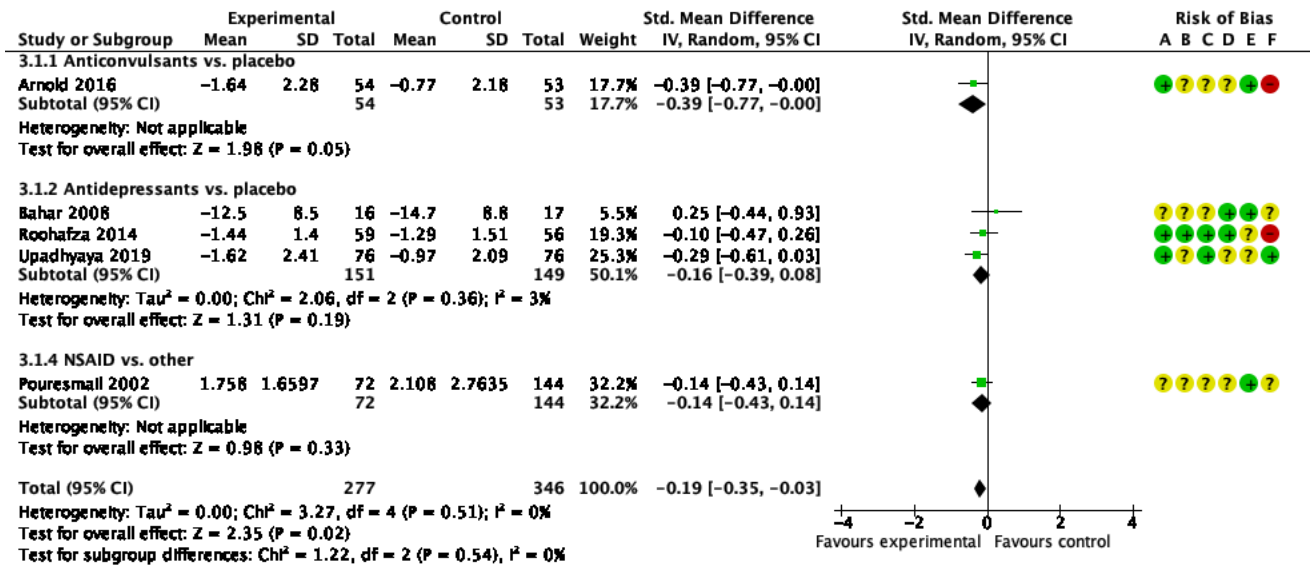
and Participation in Patients With Juvenile Idiopathic Arthritis: a Randomized Clinical Trial. *American journal of physical medicine & rehabilitation* 2019; **98 RTY - Journal article**(3): 174-81.
32. Elnaggar RK, Elshafey MA. Effects of Combined Resistive Underwater Exercises and Interferential Current Therapy in Patients with Juvenile Idiopathic Arthritis: a Randomized Controlled Trial. *American journal of physical medicine & rehabilitation* 2016; **95 RTY - Journal article**(2): 96-102.
33. Epps H, Ginnelly L, Utley M, et al. Is hydrotherapy cost-effective? A randomised controlled trial of combined hydrotherapy programmes compared with physiotherapy land techniques in children with juvenile idiopathic arthritis. *Health technology assessment (Winchester, England)* 2005; **9 RTY - Journal article**(39): iii-iv, ix-x, 1-59.
34. Kemp SRIGCWSDJEBEMCAGMLJBMW. A randomized comparative trial of generalized vs targeted physiotherapy in the management of childhood hypermobility. *Rheumatology (Oxford, England)* 2010; **49 RTY - Journal article**(2): 315-25.
35. Pacey V, Tofts L, Adams RD, Munns CF, Nicholson LL. Exercise in children with joint hypermobility syndrome and knee pain: a randomised controlled trial comparing exercise into hypermobile versus neutral knee extension. *Pediatric rheumatology online journal* 2013; **11 RTY - Journal article**(1): 30.
36. Stephens S, Feldman BM, Bradley N, et al. Feasibility and effectiveness of an aerobic exercise program in children with fibromyalgia: results of a randomized controlled pilot trial. *Arthritis and rheumatism* 2008; **59 RTY - Journal article**(10): 1399-406.
37. Zapata KA, Wang-Price SS, Sucato DJ, Thompson M, Trudelle-Jackson E, Lovelace-Chandler V. Spinal Stabilization Exercise Effectiveness for Low Back Pain in Adolescent Idiopathic Scoliosis: a Randomized Trial. *Pediatric physical therapy* 2015; **27 RTY - Journal article**(4): 396-402.
38. Physical Therapy Treatment Once a Month Versus Once a Week for Posture Improvement in Children and Adolescents. *https://clinicaltrialsgov/show/NCT03046472* 2017.
39. Ahlqwist A, Hagman M, Kjellby-Wendt G, Beckung E. Physical therapy treatment of back complaints on children and adolescents. *Spine* 2008; **33 RTY - Journal article**(20): E721-7.
40. Perez Ramirez NE, Cares PN, Penailillo PSM. Effectiveness of Watsu therapy in patients with juvenile idiopathic arthritis. A parallel, randomized, controlled and single-blind clinical trial. *Revista chilena de pediatria* 2019; **90 RTY - Journal article**(3): 282-92.
41. Connelly M, Schanberg LE, Ardoin S, et al. Multisite Randomized Clinical Trial Evaluating an Online Self-Management Program for Adolescents With Juvenile Idiopathic Arthritis. *J Pediatr Psychol* 2019; **44**(3): 363-74.
42. Bonnert M, Olen O, Lalouni M, et al. Internet-delivered exposure-based cognitive behaviour therapy for adolescents with functional abdominal pain or functional dyspepsia: A feasibility study. *Journal of pediatric gastroenterology and nutrition* 2017; **64 (Supplement 1)**: 134-5.
43. Palermo TM, Law EF, Fales J, Bromberg MH, Jessen-Fiddick T, Tai G. Internet-delivered cognitive-behavioral treatment for adolescents with chronic pain and their parents: A randomized controlled multicenter trial. *Pain* 2016; **157**(1): 174.

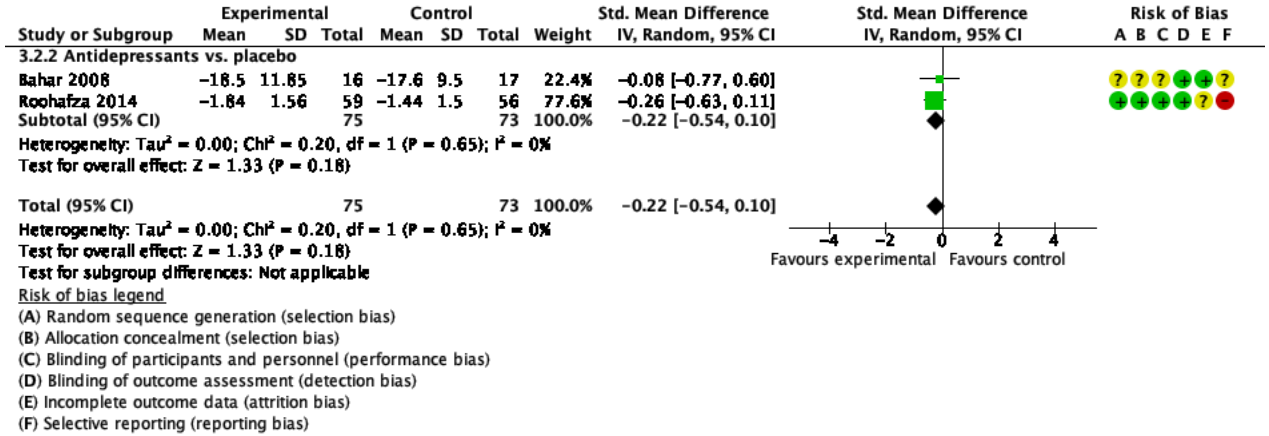
Appendix E

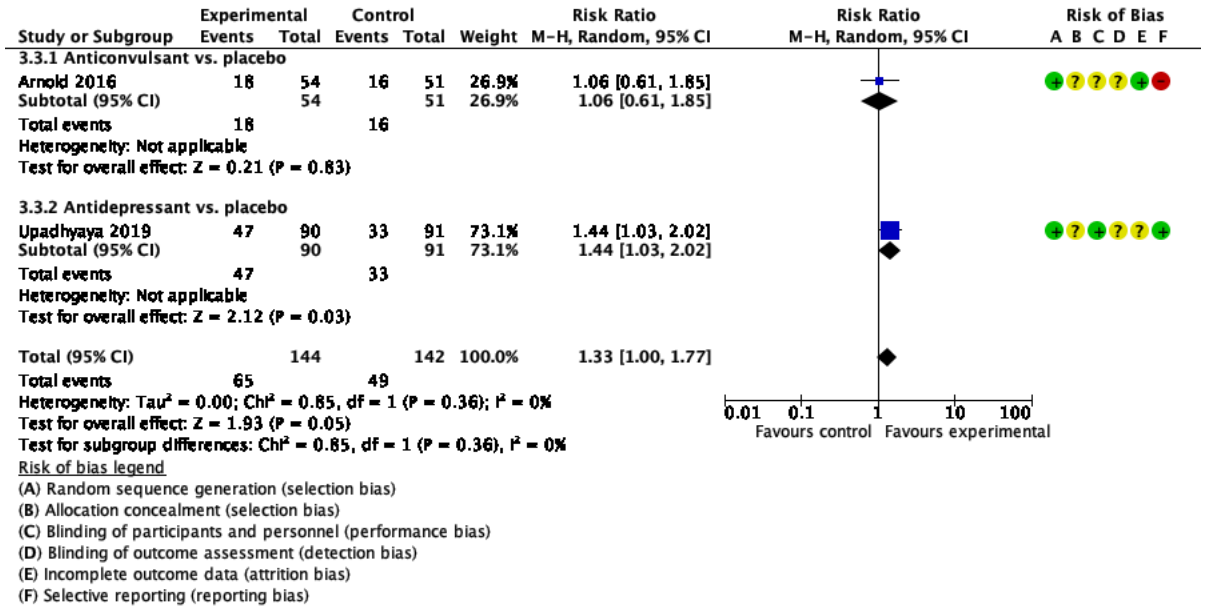
Appendix E.1. WHO review: Pharmacological interventions for children with chronic pain

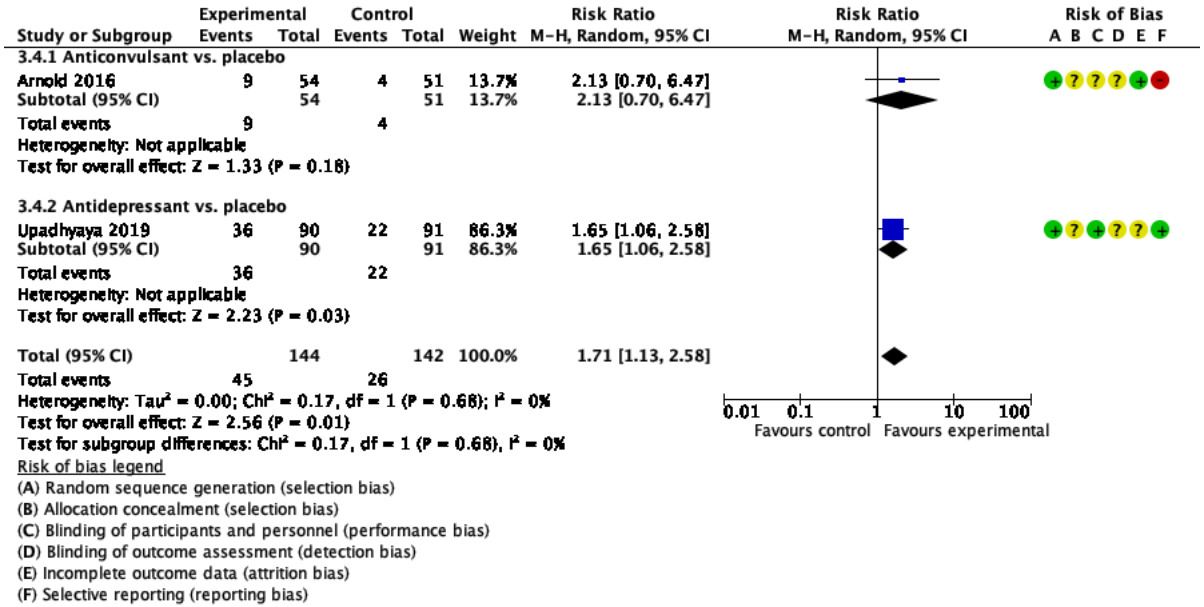
Comparison: Pharmacological therapies versus placebo, non-pharmacological or waitlist control
Population: Children and adolescents with chronic pain
Setting: Any setting
Studies: Randomised controlled trials
Please note, pharmacological interventions compared to other pharmacological interventions are not included in these analyses.

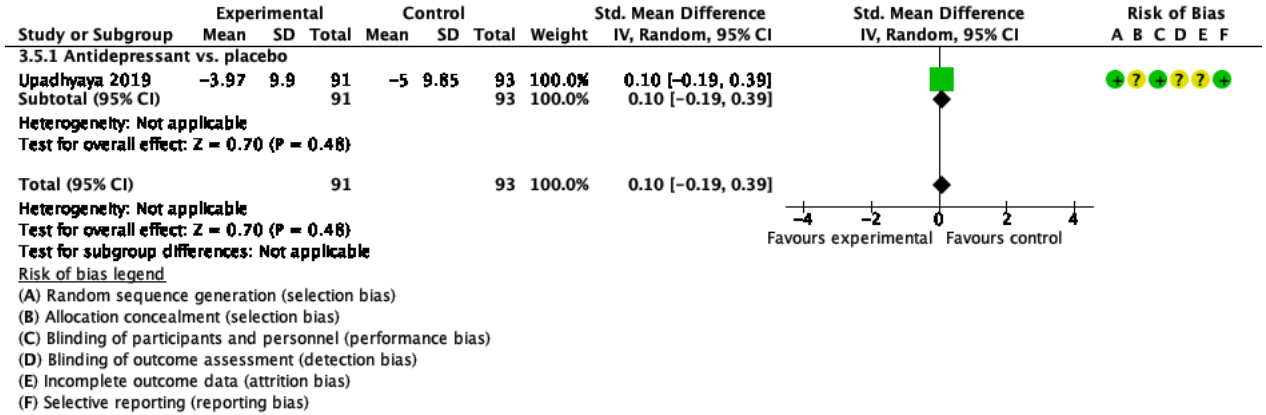
Risk of bias legend
(**A**) Random sequence generation (selection bias)
(**B**) Allocation concealment (selection bias)
(**C**) Blinding of outcome assessment (detection bias)
(**D**) Incomplete outcome data (attrition bias)
(**E**) Selective reporting (reporting bias)

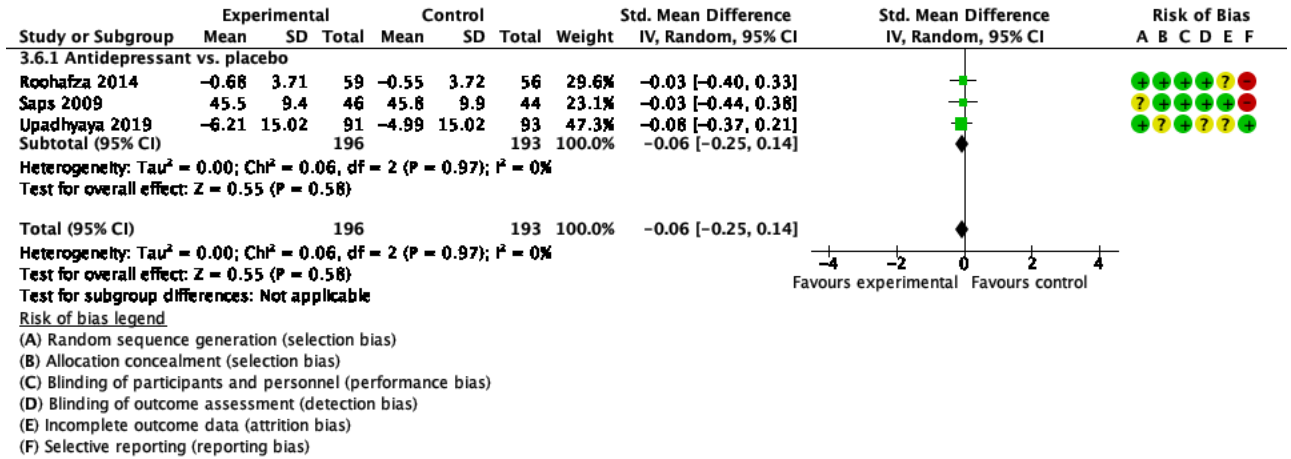
Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Bahar 2008: Amitriptyline vs. placebo; Chronic secondary visceral pain (IBS), 14 years • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years • Pouresmall 2002: Ibuprofen vs. acupressure or sham acupressure; Chronic primary visceral pain (Dysmenorrhea), 14-18 years	Overall: ⊕⊕⊕◯ MODERATE Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕⊕◯◯ LOW NSAID vs. other: ⊕◯◯◯ VERY LOW

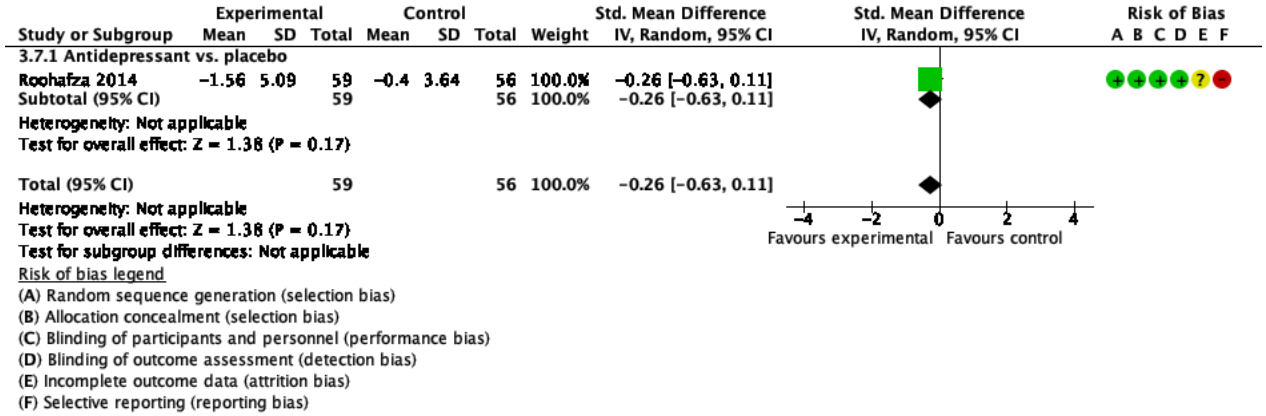
Pain intensity, follow-up *Higher scores indicate higher pain intensity*	Pain intensity, follow-up  • Bahar 2008: Amitriptyline vs. placebo; Chronic secondary visceral pain (IBS), 14 years • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years	Overall: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

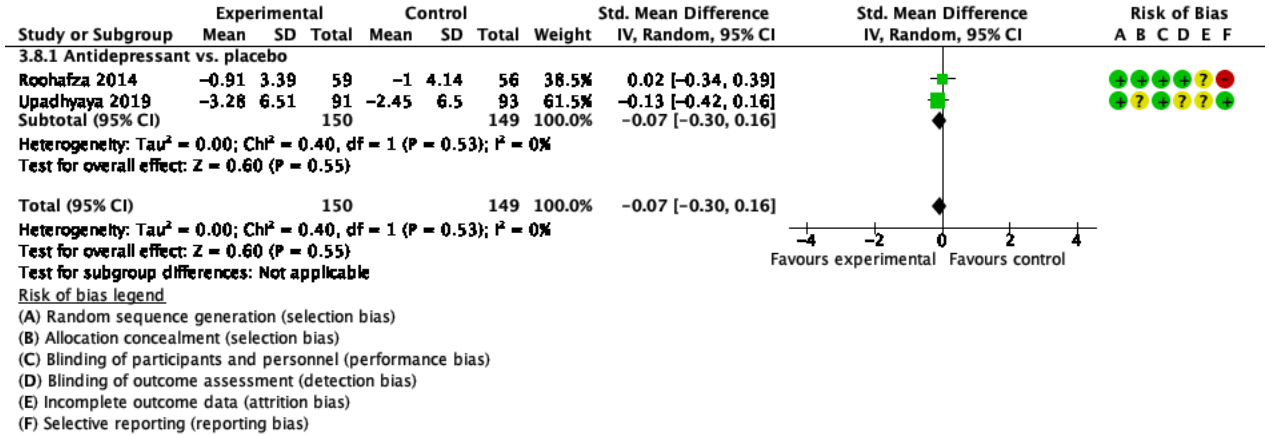
30% reduction, post-treatment	30% reduction, post-treatment  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Overall: ⊕◯◯◯ VERY LOW Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

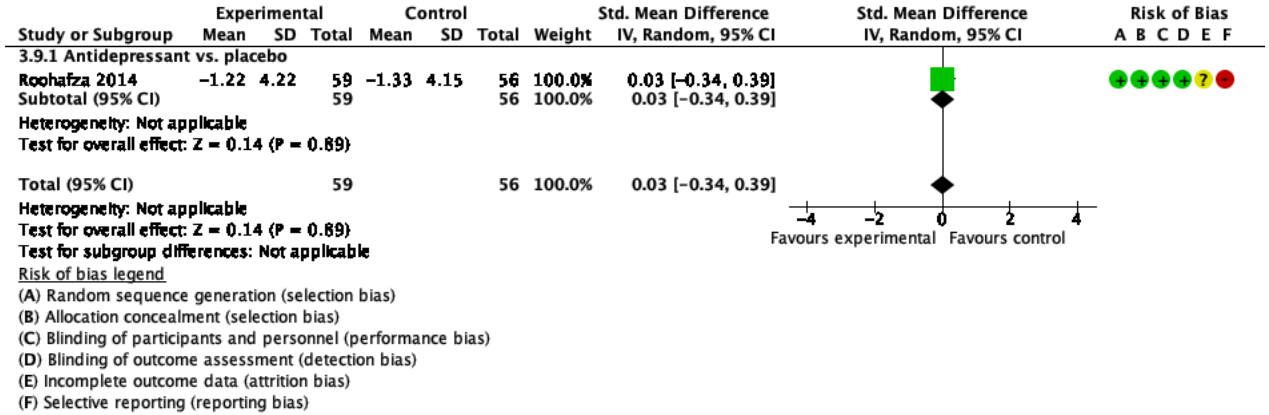
50% pain reduction, post-treatment	50% pain reduction, post-treatment  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Overall: ⊕◯◯◯ VERY LOW Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

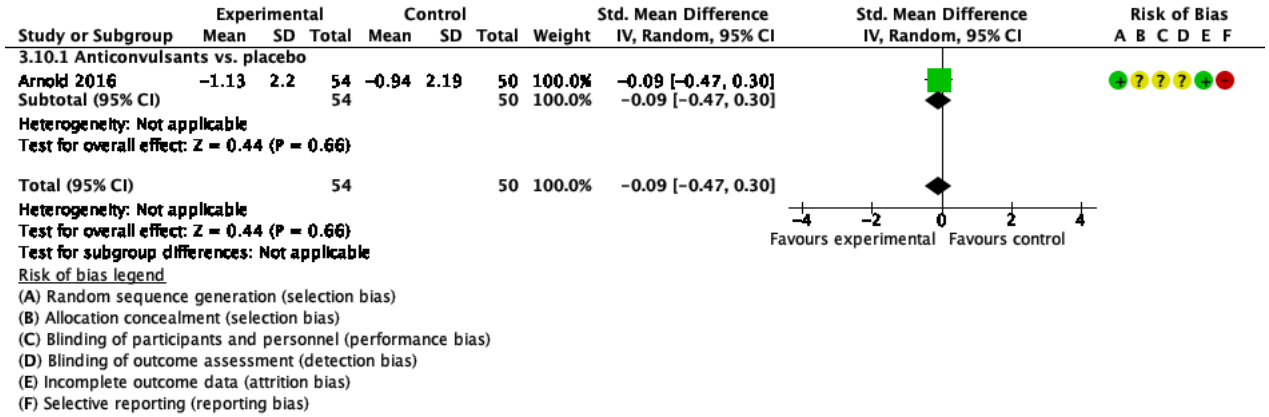
Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment  • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Overall: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

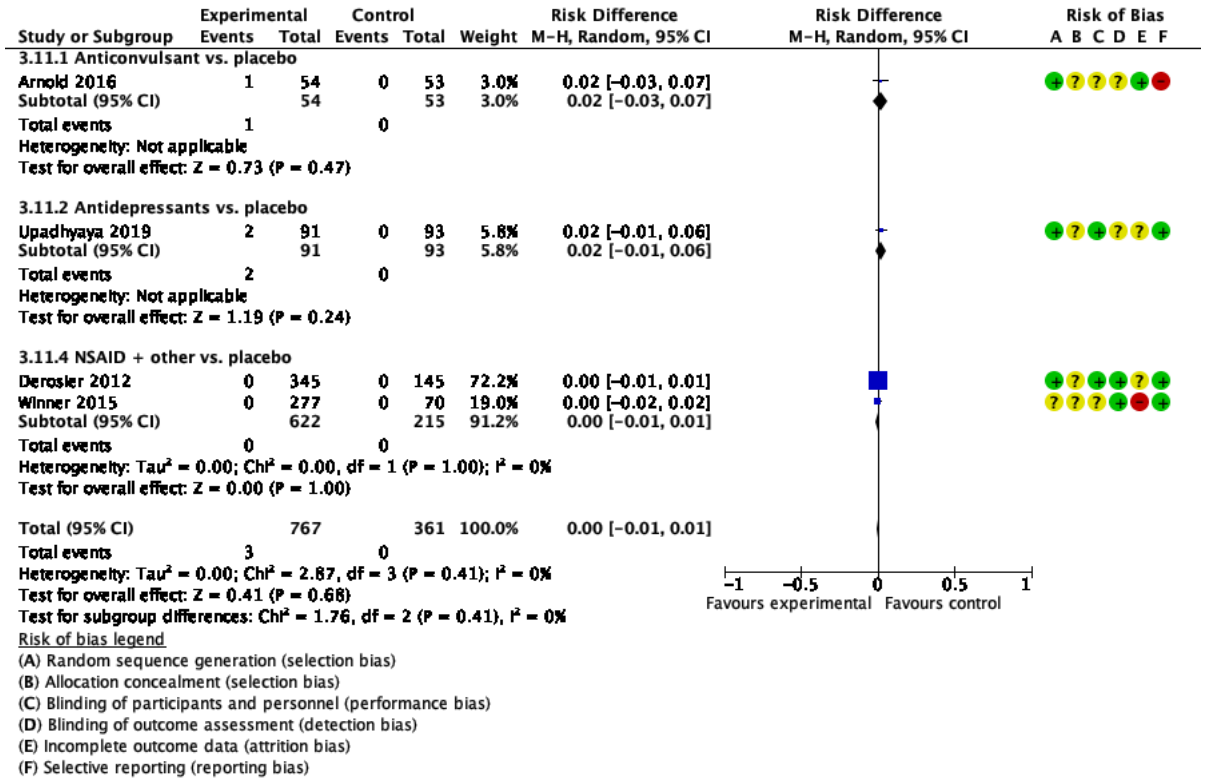
Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, post-treatment  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Saps 2009: Amitriptyline vs. placebo; Mixed pain (FAP, Functional dyspepsia, IBS), 12 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Overall: ⊕⊕◯◯ LOW Antidepressants vs. placebo: ⊕⊕◯◯ LOW

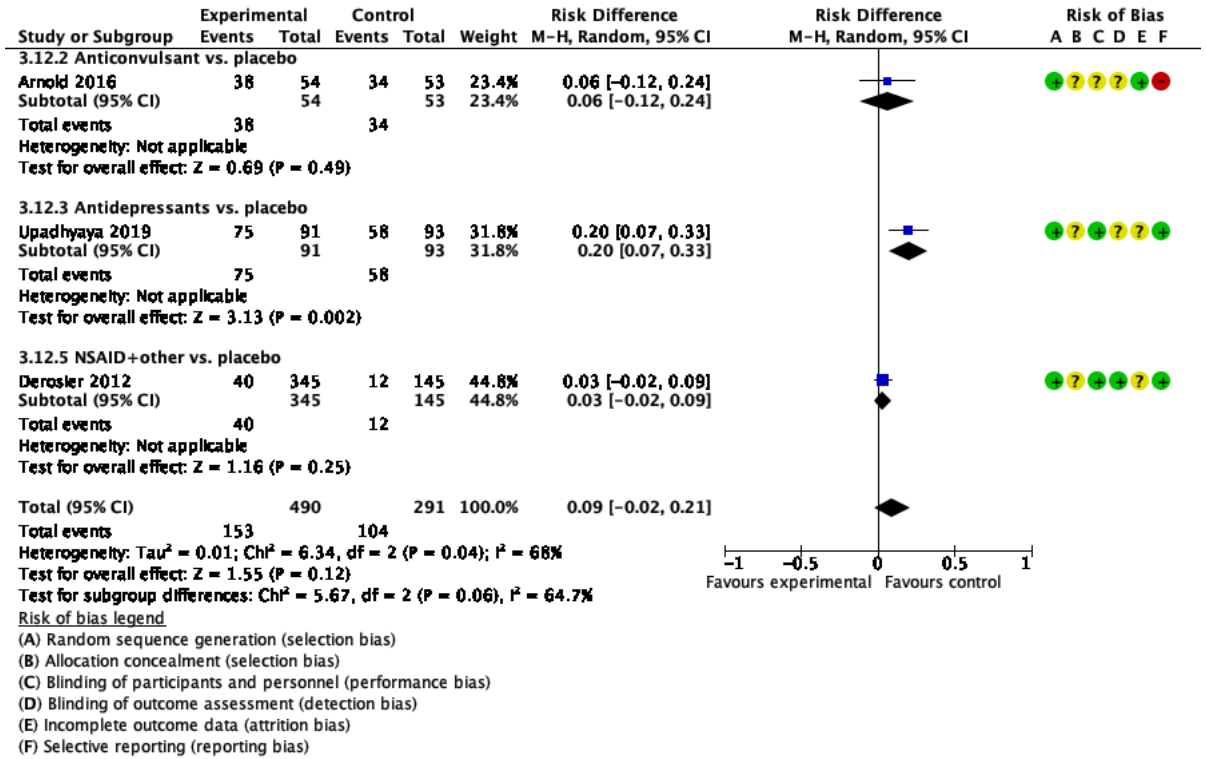
Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow up  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years	Overall: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher anxious symptomology*	Emotional functioning: Anxiety, post-treatment  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Overall: ⊕⊕◯◯ LOW Antidepressants vs. placebo: ⊕⊕◯◯ LOW

Emotional functioning: Anxiety, follow-up *Higher scores indicate higher anxious symptomology*	Emotional functioning: Anxiety, follow-up  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years	Overall: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

Sleep, post-treatment *Lower scores indicate worse sleep quality*	Sleep, post-treatment  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years	Overall: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

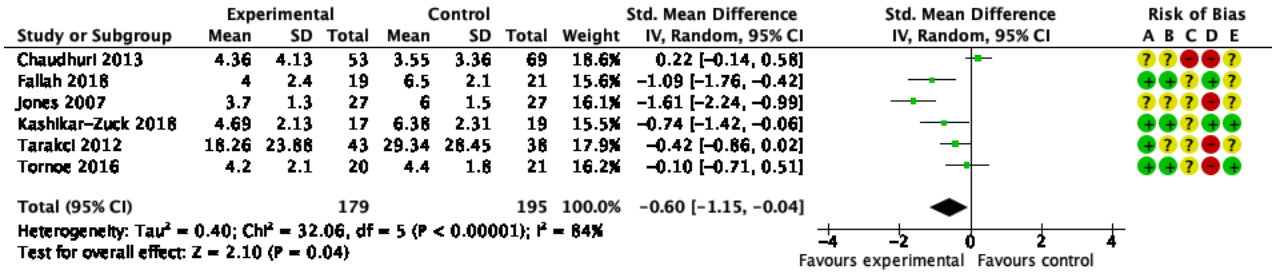
Treatment-related serious adverse events	Treatment-related serious adverse events  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years • Derosier 2012: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine) • Winner 2015: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine), 15 years, 15 years	Overall: ⊕◯◯◯ VERY LOW Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW NSAID + other vs. placebo: ⊕◯◯◯ VERY LOW

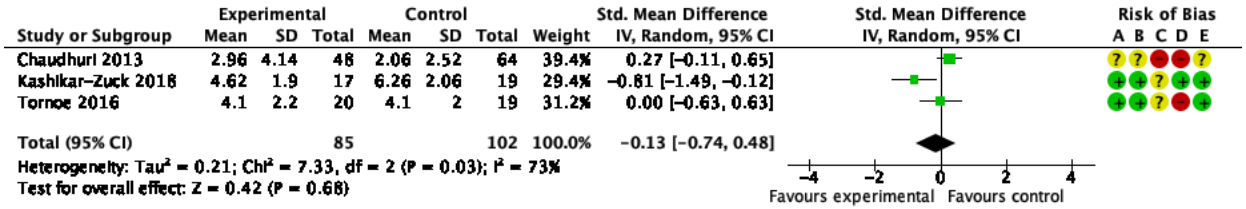
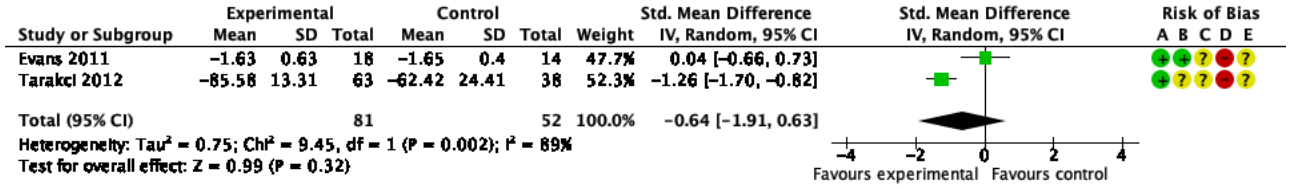
Treatment-related adverse events	Treatment-related adverse events  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Derosier 2012: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine), 15 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Overall: ⊕◯◯◯ VERY LOW Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW NSAID + other vs. placebo: ⊕◯◯◯ VERY LOW

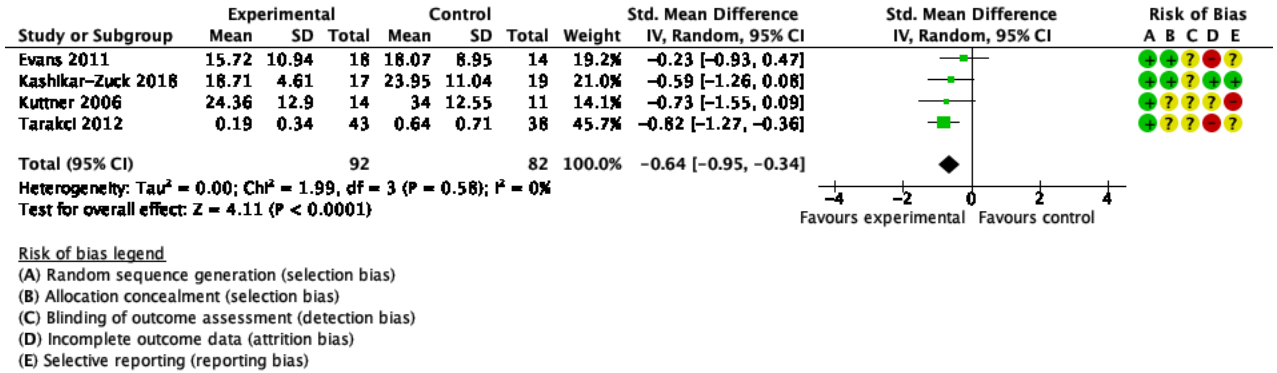
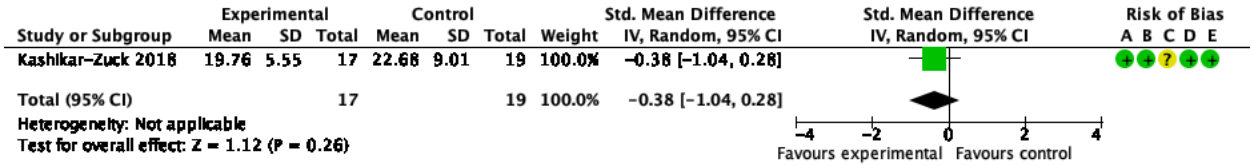
Appendix E.2. WHO review: Physical interventions for children with chronic pain

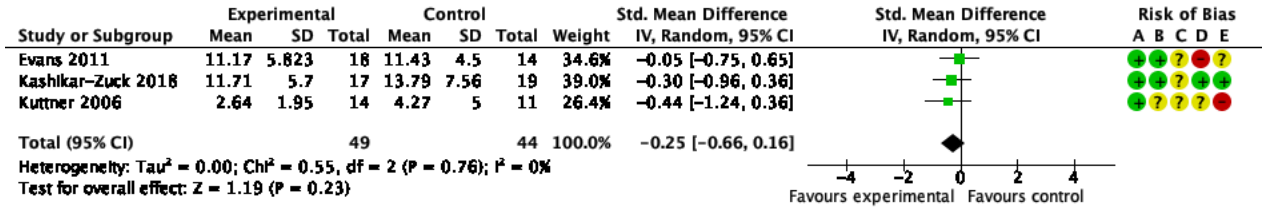
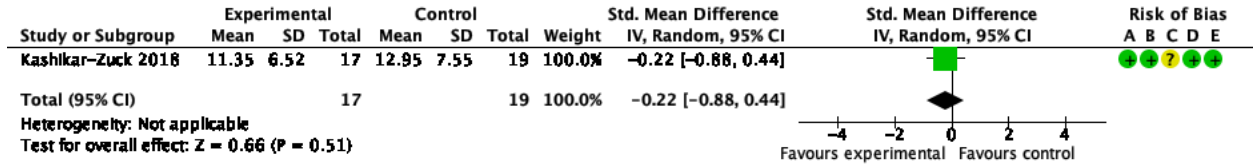
Comparison: Physical therapies versus treatment as usual, waitlist control, or non-physical therapy control
Population: Children with any chronic pain
Setting: Any setting
Studies: Randomised controlled trials

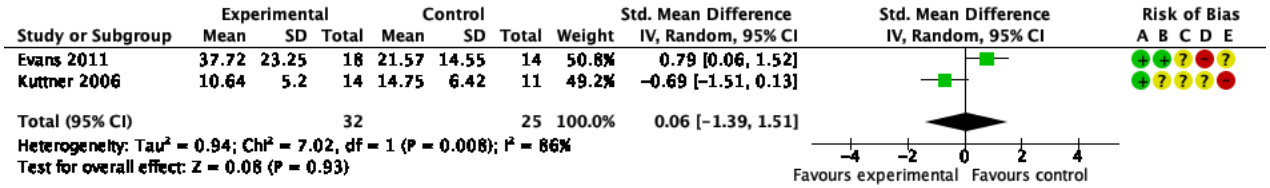
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

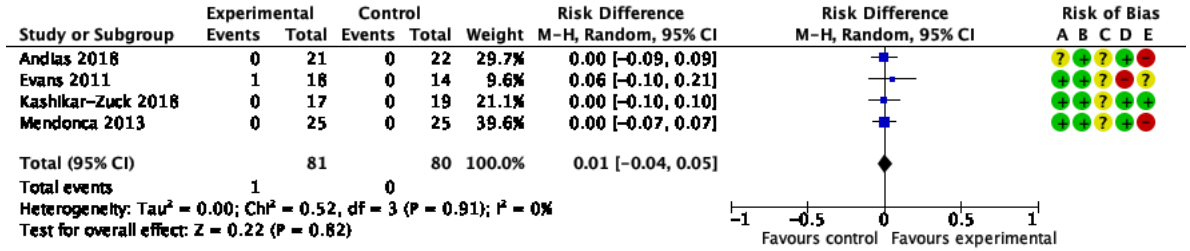
Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment  	⊕◯◯◯ VERY LOW

Pain intensity, follow-up *Higher scores indicate higher pain intensity*	Pain intensity, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW
Health-related quality of life, post-treatment *Lower scores indicate better quality of life*	Health-related quality of life, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW

Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW
Functional disability, follow-up *Higher scores indicate lower disability*	Functional disability, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW
Emotional functioning: Depression, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW

<table>
<tr>
<td>Emotional functioning: Anxiety, post-treatment
Higher scores indicate higher anxious symptomology</td>
<td>
Emotional functioning: Anxiety, post-treatment


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
</td>
<td>⊕◯◯◯
VERY LOW</td>
</tr>
<tr>
<td>Treatment-related adverse events, post-treatment</td>
<td>
Treatment-related adverse events


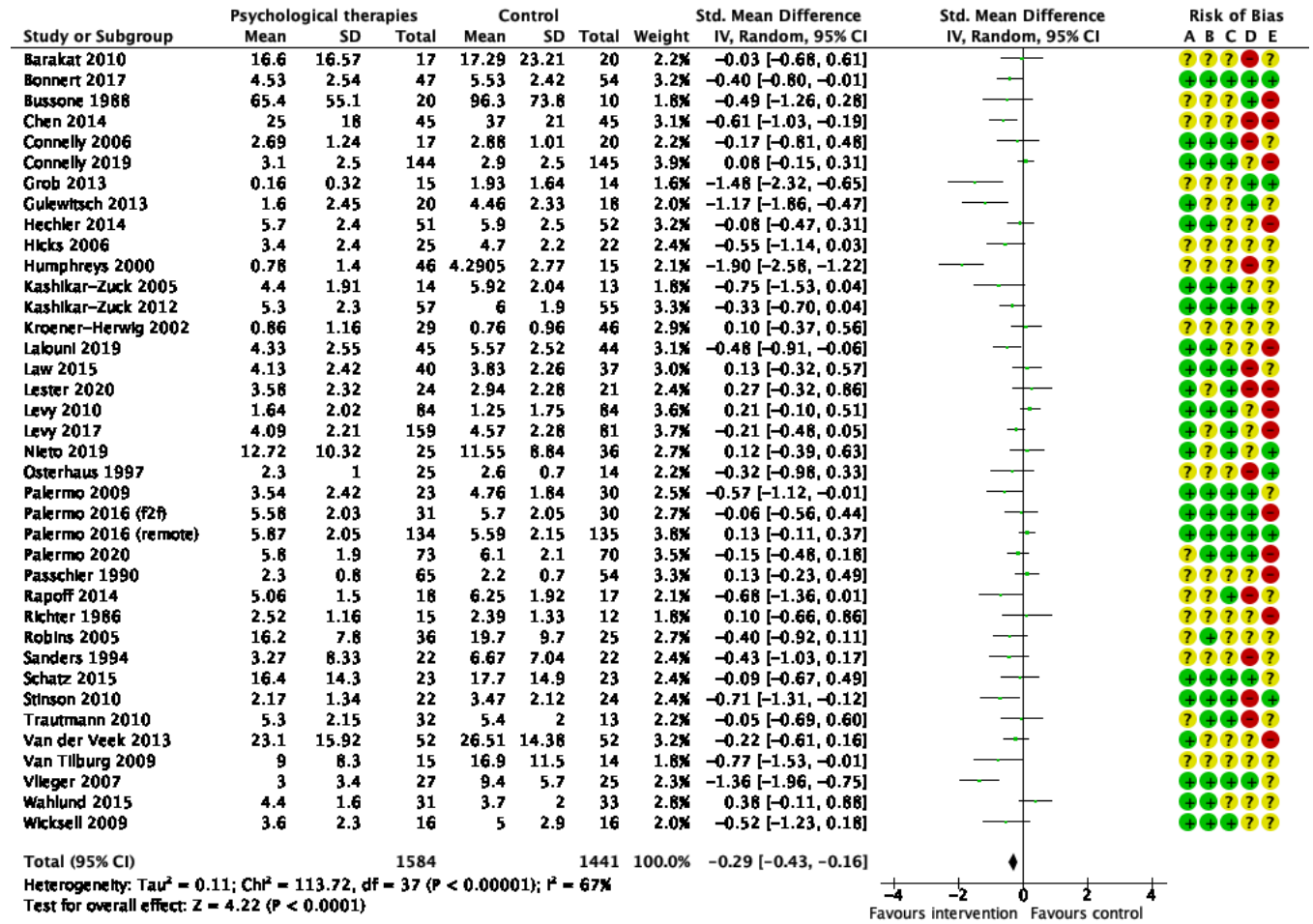
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
</td>
<td>⊕◯◯◯
VERY LOW</td>
</tr>
</table>

Appendix E.3. WHO review: Psychological interventions for children with chronic pain

Comparison: Psychological therapies versus any control (standard care, waitlist control, active (non-psychological therapy) control)
Population: children with any chronic pain
Setting: Any setting
Studies: Randomised controlled trials

Risk of bias legend
(**A**) Random sequence generation (selection bias)
(**B**) Allocation concealment (selection bias)
(**C**) Blinding of outcome assessment (detection bias)
(**D**) Incomplete outcome data (attrition bias)
(**E**) Selective reporting (reporting bias)

Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment	⊕⊕◯◯ LOW



Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Barakat 2010	16.6	16.57	17	17.29	23.21	20	2.2%	−0.03 [−0.68, 0.61]
Bonnert 2017	4.53	2.54	47	5.53	2.42	54	3.2%	−0.40 [−0.80, −0.01]
Bussone 1988	65.4	55.1	20	96.3	73.8	10	1.8%	−0.49 [−1.26, 0.28]
Chen 2014	25	18	45	37	21	45	3.1%	−0.61 [−1.03, −0.19]
Connelly 2006	2.69	1.24	17	2.88	1.01	20	2.2%	−0.17 [−0.81, 0.48]
Connelly 2019	3.1	2.5	144	2.9	2.5	145	3.9%	0.08 [−0.15, 0.31]
Grob 2013	0.16	0.32	15	1.93	1.64	14	1.6%	−1.48 [−2.32, −0.65]
Gulewitsch 2013	1.6	2.45	20	4.46	2.33	18	2.0%	−1.17 [−1.86, −0.47]
Hechler 2014	5.7	2.4	51	5.9	2.5	52	3.2%	−0.08 [−0.47, 0.31]
Hicks 2006	3.4	2.4	25	4.7	2.2	22	2.4%	−0.55 [−1.14, 0.03]
Humphreys 2000	0.78	1.4	46	4.2905	2.77	15	2.1%	−1.90 [−2.58, −1.22]
Kashikar-Zuck 2005	4.4	1.91	14	5.92	2.04	13	1.8%	−0.75 [−1.53, 0.04]
Kashikar-Zuck 2012	5.3	2.3	57	6	1.9	55	3.3%	−0.33 [−0.70, 0.04]
Kroener-Herwig 2002	0.86	1.16	29	0.76	0.96	46	2.9%	0.10 [−0.37, 0.56]
Lalouni 2019	4.33	2.55	45	5.57	2.52	44	3.1%	−0.48 [−0.91, −0.06]
Law 2015	4.13	2.42	40	3.83	2.26	37	3.0%	0.13 [−0.32, 0.57]
Lester 2020	3.58	2.32	24	2.94	2.28	21	2.4%	0.27 [−0.32, 0.86]
Levy 2010	1.64	2.02	84	1.25	1.75	84	3.6%	0.21 [−0.10, 0.51]
Levy 2017	4.09	2.21	159	4.57	2.28	81	3.7%	−0.21 [−0.48, 0.05]
Nieto 2019	12.72	10.32	25	11.55	8.84	36	2.7%	0.12 [−0.39, 0.63]
Osterhaus 1997	2.3	1	25	2.6	0.7	14	2.2%	−0.32 [−0.98, 0.33]
Palermo 2009	3.54	2.42	23	4.76	1.84	30	2.5%	−0.57 [−1.12, −0.01]
Palermo 2016 (f2f)	5.58	2.03	31	5.7	2.05	30	2.7%	−0.06 [−0.56, 0.44]
Palermo 2016 (remote)	5.87	2.05	134	5.59	2.15	135	3.8%	0.13 [−0.11, 0.37]
Palermo 2020	5.8	1.9	73	6.1	2.1	70	3.5%	−0.15 [−0.48, 0.18]
Passchier 1990	2.3	0.8	65	2.2	0.7	54	3.3%	0.13 [−0.23, 0.49]
Rapoff 2014	5.06	1.5	18	6.25	1.92	17	2.1%	−0.68 [−1.36, 0.01]
Richter 1986	2.52	1.16	15	2.39	1.33	12	1.8%	0.10 [−0.66, 0.86]
Robins 2005	16.2	7.8	36	19.7	9.7	25	2.7%	−0.40 [−0.92, 0.11]
Sanders 1994	3.27	8.33	22	6.67	7.04	22	2.4%	−0.43 [−1.03, 0.17]
Schatz 2015	16.4	14.3	23	17.7	14.9	23	2.4%	−0.09 [−0.67, 0.49]
Stinson 2010	2.17	1.34	22	3.47	2.12	24	2.4%	−0.71 [−1.31, −0.12]
Trautmann 2010	5.3	2.15	32	5.4	2	13	2.2%	−0.05 [−0.69, 0.60]
Van der Veek 2013	23.1	15.92	52	26.51	14.38	52	3.2%	−0.22 [−0.61, 0.16]
Van Tilburg 2009	9	8.3	15	16.9	11.5	14	1.8%	−0.77 [−1.53, −0.01]
Vlieger 2007	3	3.4	27	9.4	5.7	25	2.3%	−1.36 [−1.96, −0.75]
Wahlund 2015	4.4	1.6	31	3.7	2	33	2.8%	0.38 [−0.11, 0.88]
Wicksell 2009	3.6	2.3	16	5	2.9	16	2.0%	−0.52 [−1.23, 0.18]
Total (95% CI)			1584			1441	100.0%	−0.29 [−0.43, −0.16]

Heterogeneity: Tau² = 0.11; Chi² = 113.72, df = 37 (P < 0.00001); I² = 67%
Test for overall effect: Z = 4.22 (P < 0.0001)

Std. Mean Difference IV, Random, 95% CI

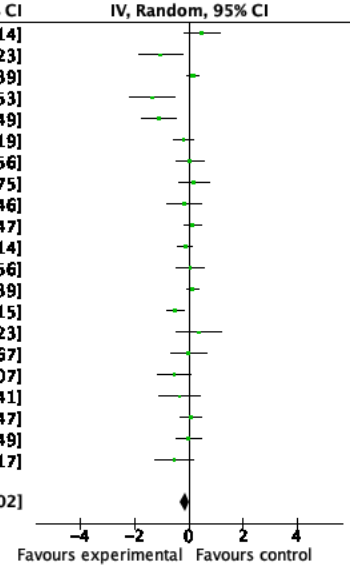
−4 −2 0 2 4
Favours intervention Favours control

Risk of Bias A B C D E

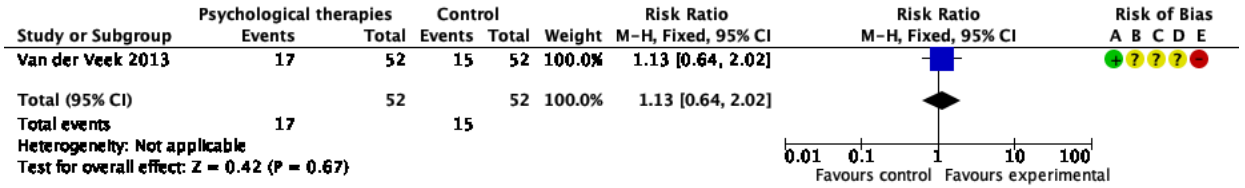
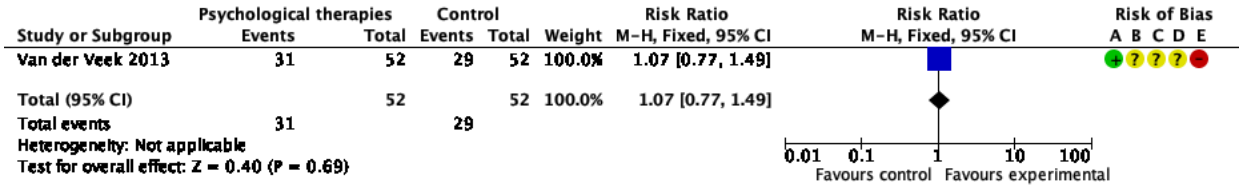
Pain intensity, follow-up *Higher scores indicate higher pain intensity*	Pain intensity, follow-up	⊕⊕◯◯ LOW

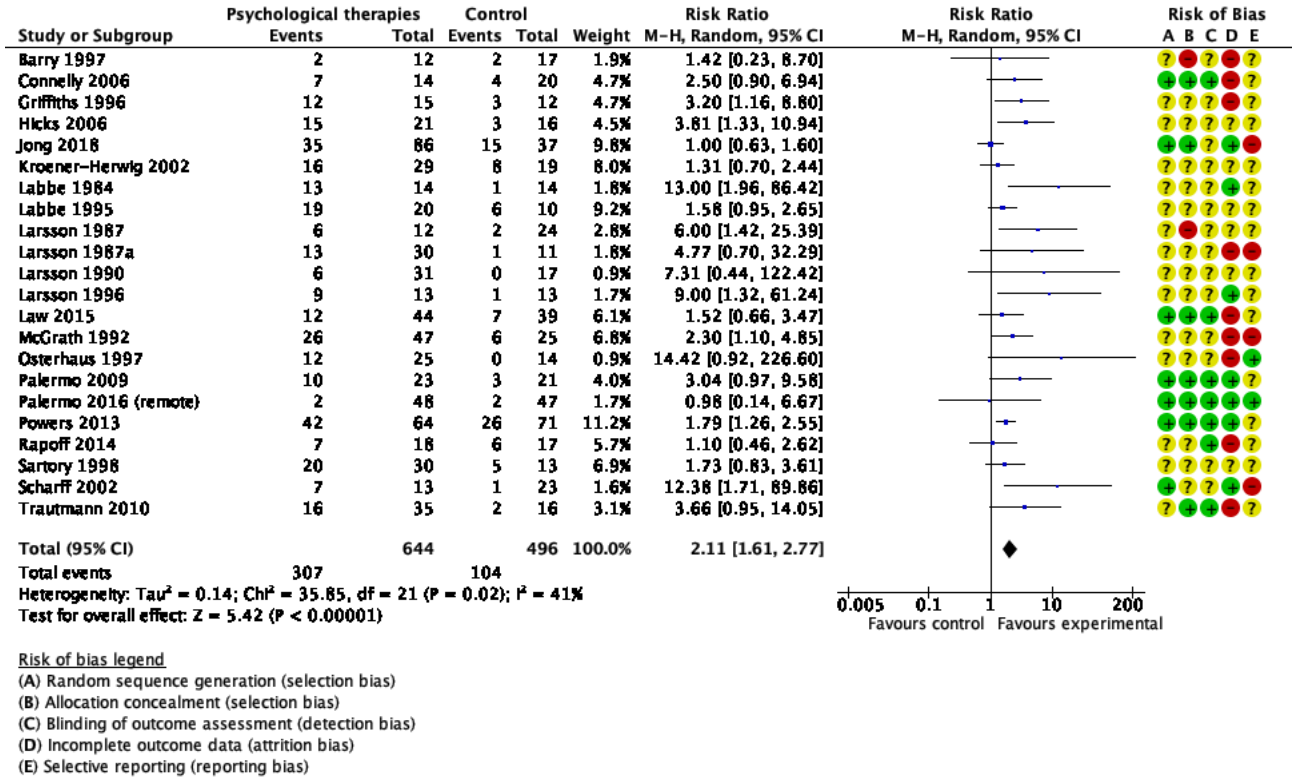
Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
Barakat 2010	16.71	23.03	17	7.84	12.31	20	3.6%	0.48 [−0.18, 1.14]	? ? ? − ?
Bussone 1988	20	18.1	20	88.8	110.3	10	2.8%	−1.04 [−1.85, −0.23]	? ? ? + −
Connelly 2019	3.1	2.5	144	2.7	2.4	145	7.6%	0.16 [−0.07, 0.39]	+ + + ? −
Grob 2013	0.08	0.31	15	1.55	1.49	14	2.7%	−1.35 [−2.17, −0.53]	? ? ? + +
Hicks 2006	2.9	2.1	25	4.9	1.3	22	3.9%	−1.11 [−1.73, −0.49]	? ? ? ? ?
Kashikar-Zuck 2012	4.9	2.2	57	5.3	2.1	55	6.1%	−0.18 [−0.56, 0.19]	+ + + + ?
Kroener-Herwig 2002	0.48	0.65	29	0.46	0.61	27	4.6%	0.03 [−0.49, 0.56]	? ? ? ? ?
Law 2015	4.19	2.45	28	3.7	2.54	22	4.4%	0.19 [−0.37, 0.75]	+ + + − ?
Lester 2020	2.67	1.9	21	3.07	2.64	18	3.8%	−0.17 [−0.80, 0.46]	+ ? + − −
Levy 2010	0.93	1.42	78	0.7	1.53	76	6.7%	0.16 [−0.16, 0.47]	+ + + ? −
Levy 2017	3.48	2.33	151	3.79	2.48	78	7.2%	−0.13 [−0.40, 0.14]	+ ? + ? −
Palermo 2016 (f2f)	5.42	2.05	31	5.3	2.12	30	4.8%	0.06 [−0.45, 0.56]	+ + + + −
Palermo 2016 (remote)	5.85	1.97	134	5.55	2.02	135	7.5%	0.15 [−0.09, 0.39]	+ + + + +
Palermo 2020	5.3	1.9	73	6.2	1.8	70	6.5%	−0.48 [−0.82, −0.15]	? + + + −
Rapoff 2014	4.46	1.88	11	3.68	2.04	11	2.6%	0.38 [−0.46, 1.23]	? ? + − ?
Richter 1986	2.02	1.48	30	2.02	1.39	12	3.6%	0.00 [−0.67, 0.67]	? ? ? ? −
Sanders 1994	0.64	1.38	22	2.11	3.56	22	4.0%	−0.53 [−1.14, 0.07]	? ? ? − ?
Trautmann 2010	4.9	1.4	12	5.5	1.9	16	3.1%	−0.34 [−1.10, 0.41]	? + + − ?
Van der Veek 2013	19.03	17.0393	52	17.72	15.19	52	6.0%	0.08 [−0.30, 0.47]	+ ? ? ? −
Wahlund 2015	2.8	1.9	31	2.8	1.6	33	5.0%	0.00 [−0.49, 0.49]	+ + ? ? ?
Wicksell 2009	3.1	2.7	16	4.5	2.4	16	3.3%	−0.53 [−1.24, 0.17]	+ + + ? ?
Total (95% CI)			997			884	100.0%	−0.14 [−0.30, 0.02]	

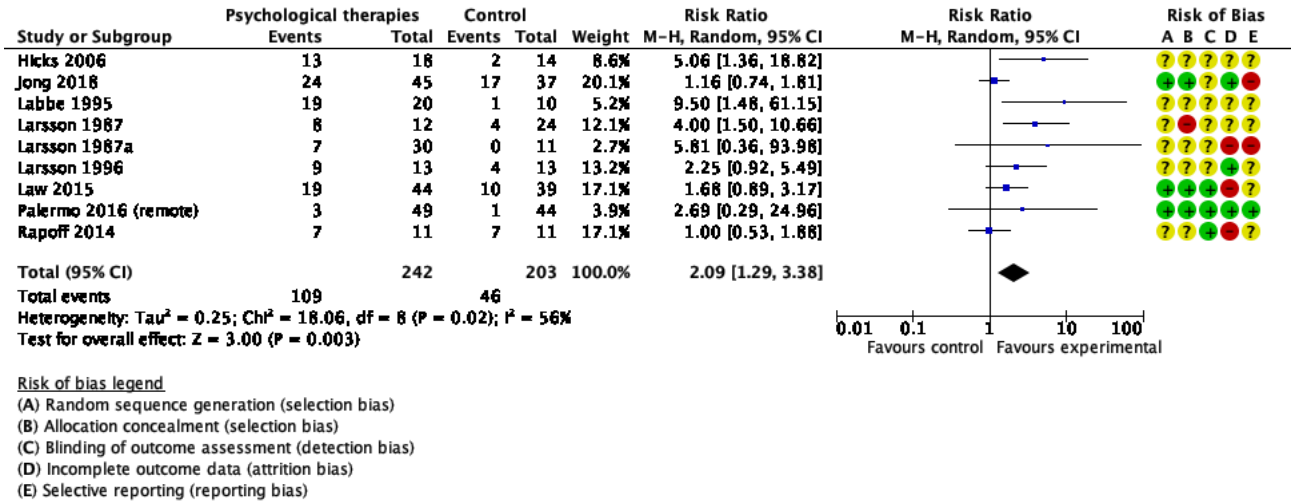
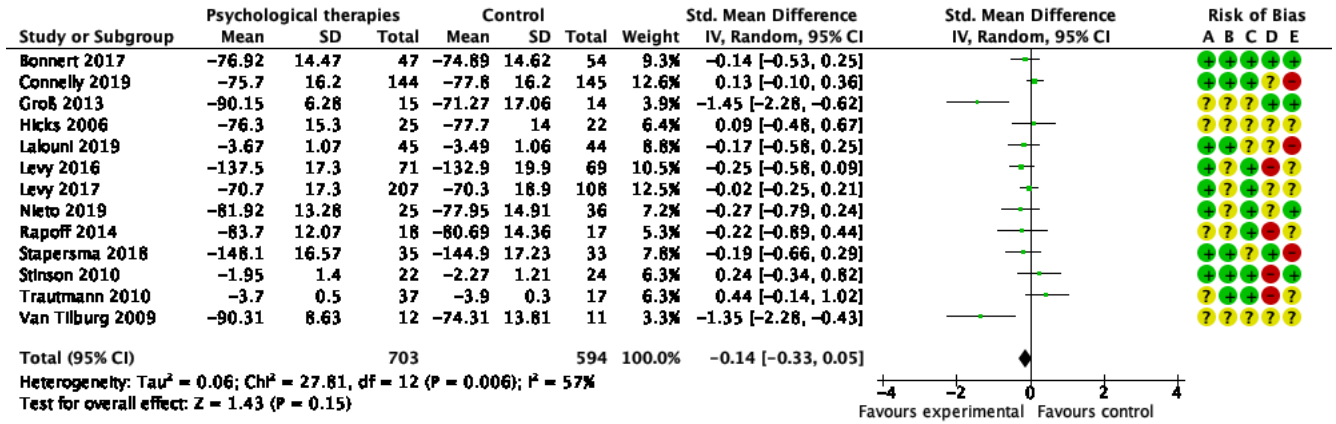
Heterogeneity: $Tau^2 = 0.08$; $Chi^2 = 51.59$, df = 20 (P = 0.0001); $I^2 = 61\%$
Test for overall effect: Z = 1.67 (P = 0.09)

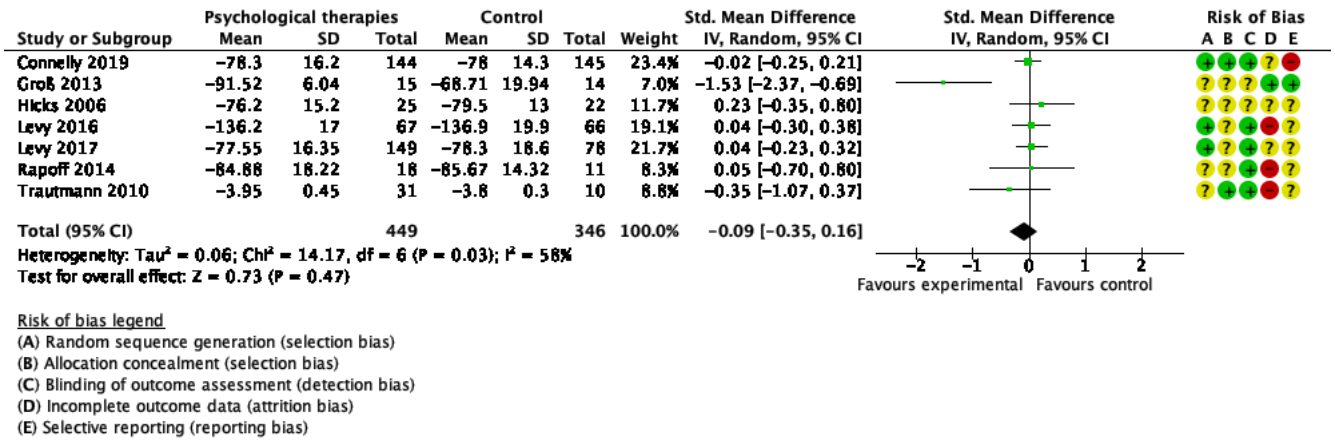


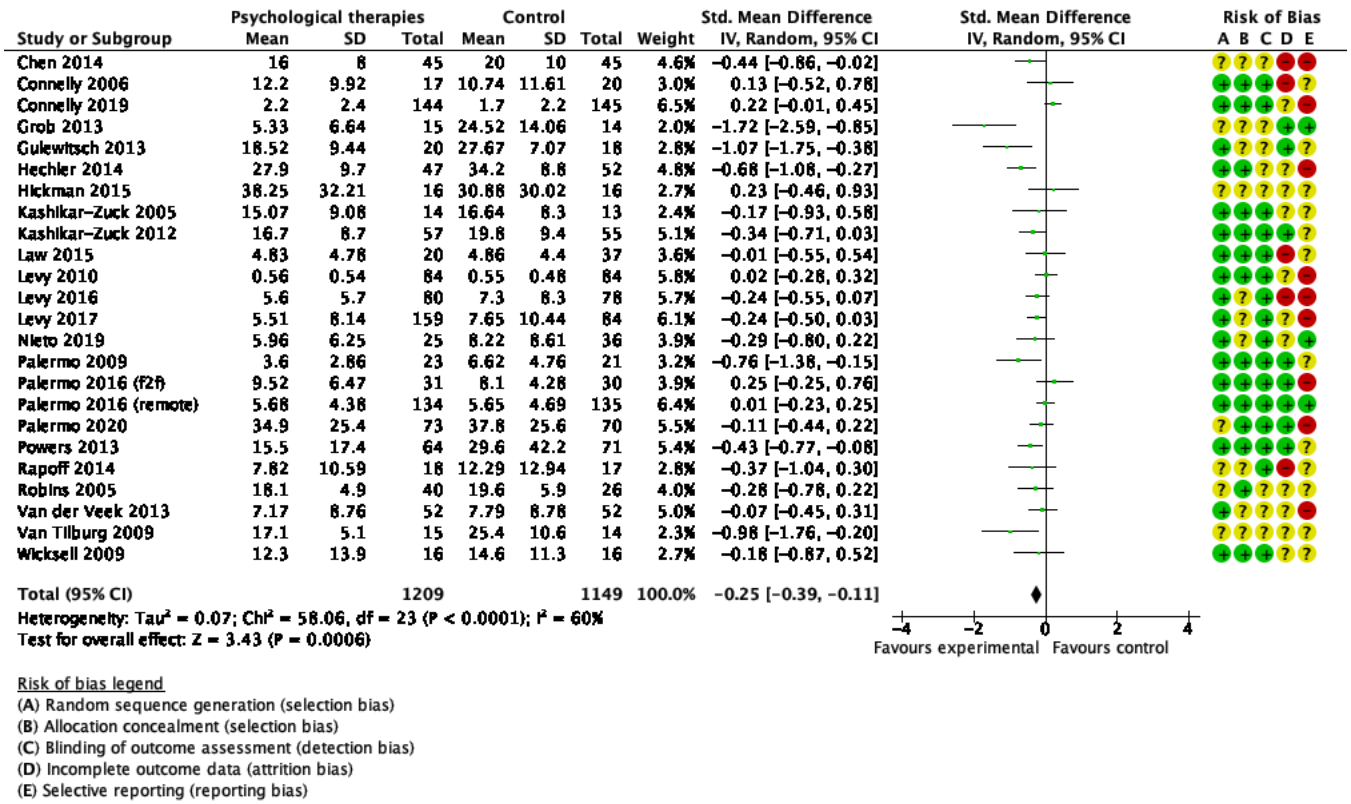
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

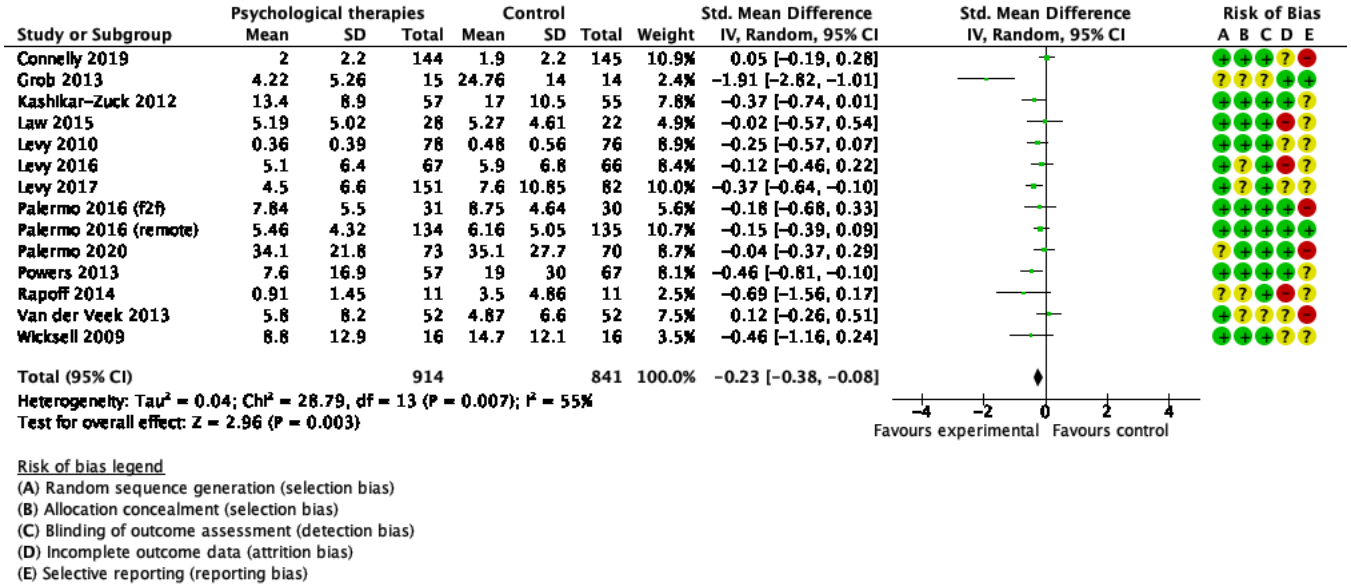
30% reduction, post-treatment	30% pain reduction, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW
30% reduction, follow-up	30% pain reduction, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW

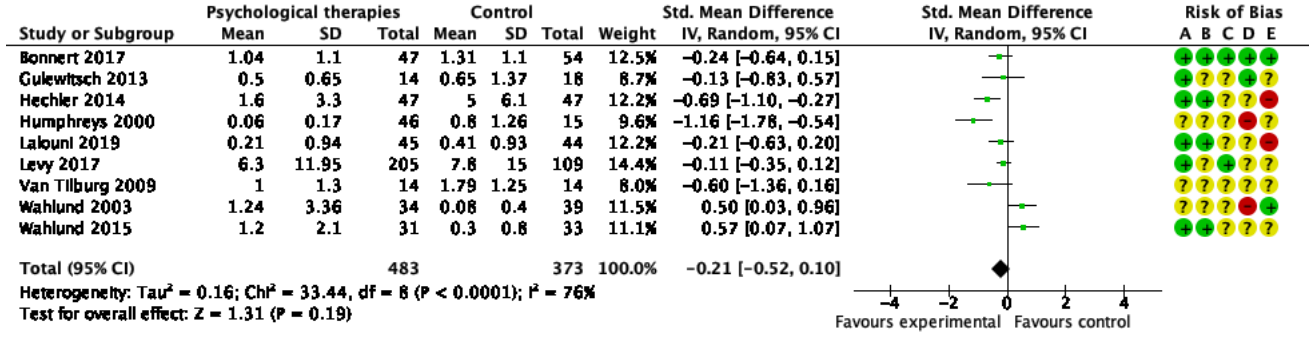
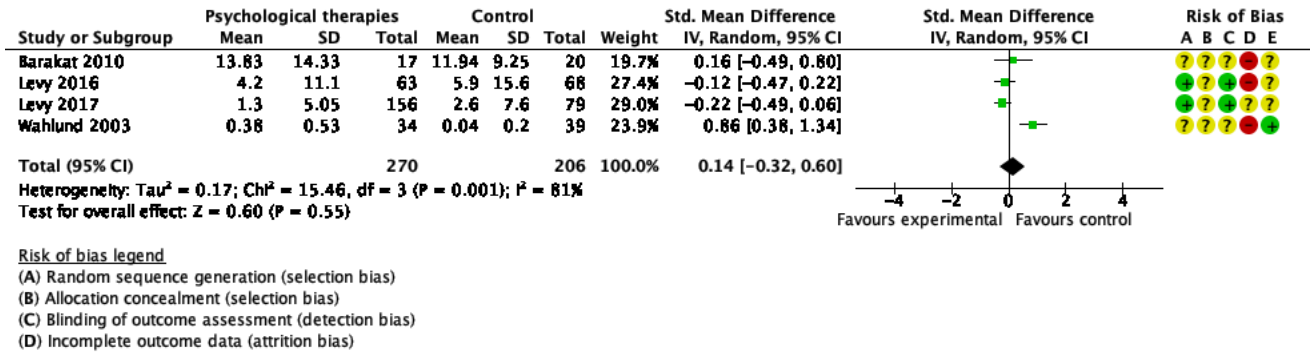
50% reduction, post-treatment	50% pain reduction, post-treatment 	⊕⊕◯◯ LOW

50% reduction, follow-up	50% pain reduction, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕◯◯◯ VERY LOW
Health-related quality of life, post-treatment *Lower scores indicate lower quality of life*	Health-related quality of life, post-treatment 	⊕⊕◯◯ LOW

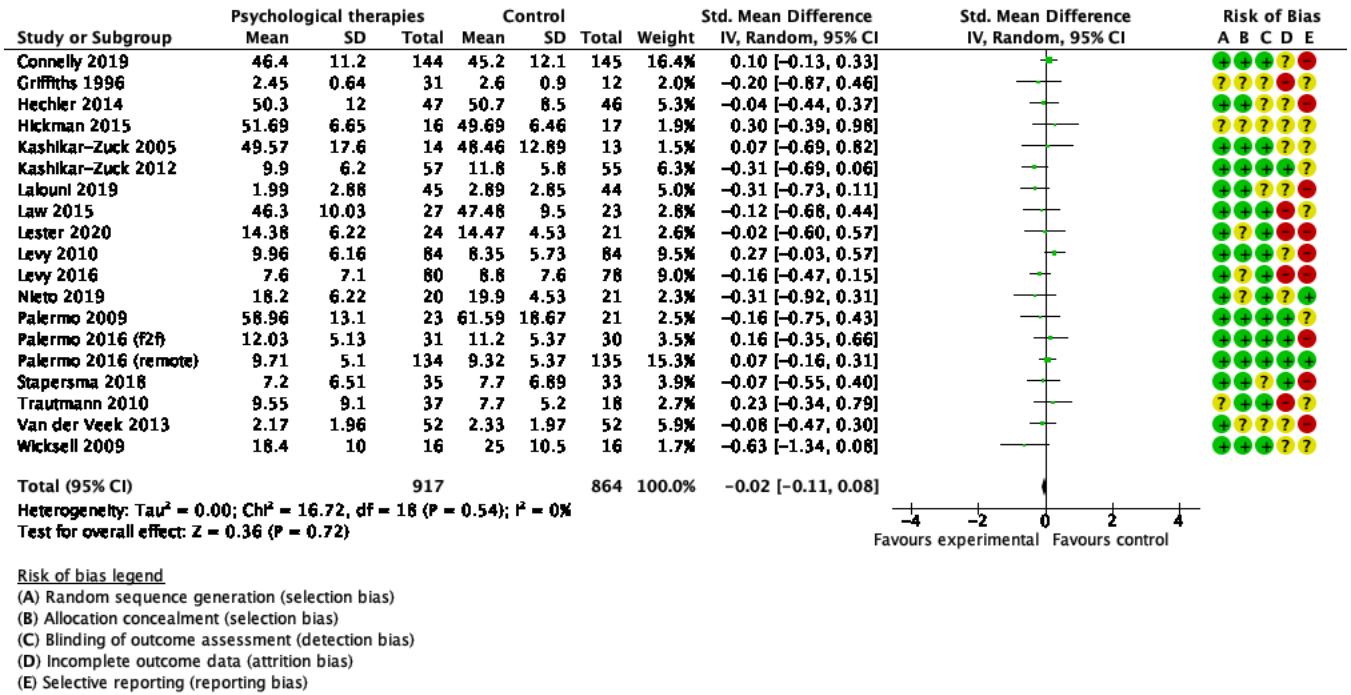
Health-related quality of life, follow-up *Lower scores indicate better quality of life*	Health-related quality of life, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)	⊕⊕◯◯ LOW

Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment 	⊕⊕◯◯ LOW

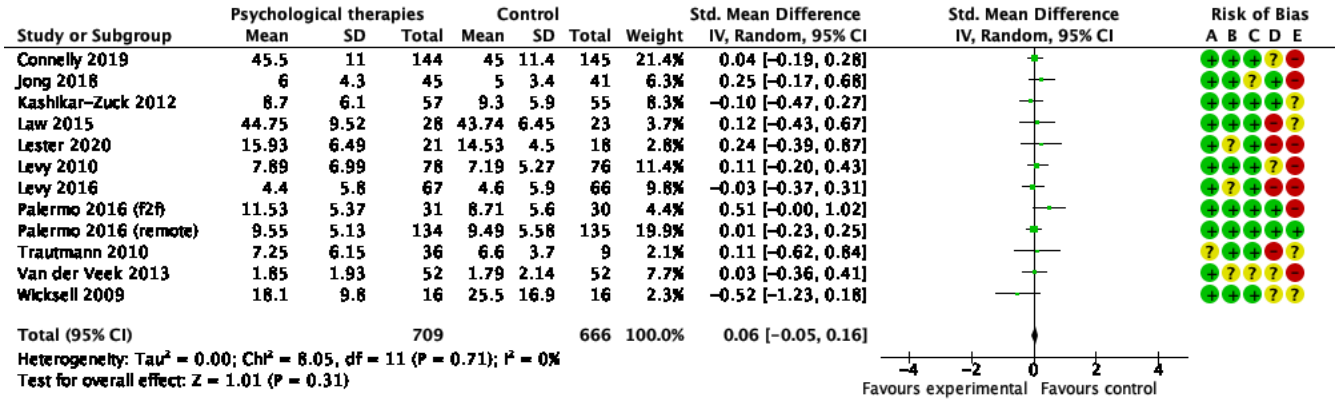
Functional disability, follow-up *Higher scores indicate lower disability*	Functional disability, follow-up  	⊕⊕⊕◯ MODERATE
Role functioning	Role functioning (school absence), post-treatment	⊕◯◯◯ VERY LOW

(school absence), post-treatment *Higher scores indicate more absence from school*		
Role functioning (school absence), post-treatment *Higher scores indicate more absence from school*	Role functioning (school absence), follow-up 	⊕◯◯◯ VERY LOW

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, post-treatment	⊕⊕⊕⊕ HIGH



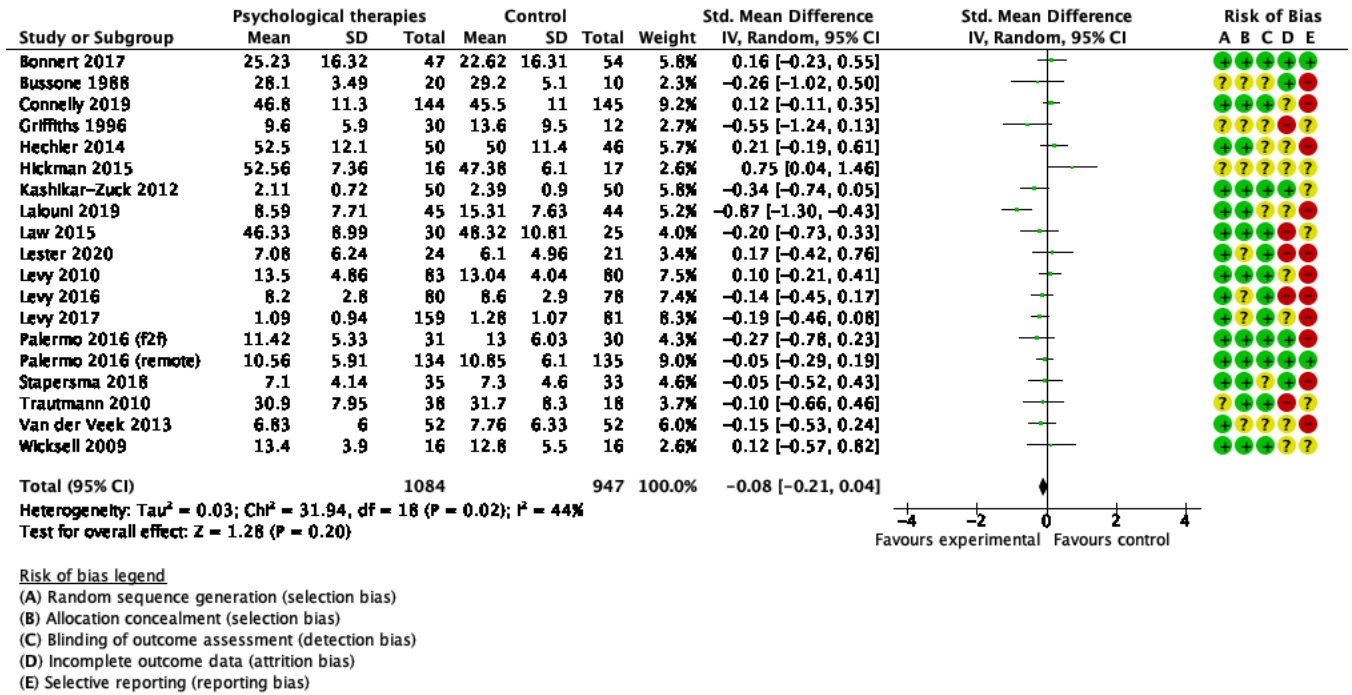
Emotional functioning: Depression, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow-up	⊕⊕⊕⊕ HIGH

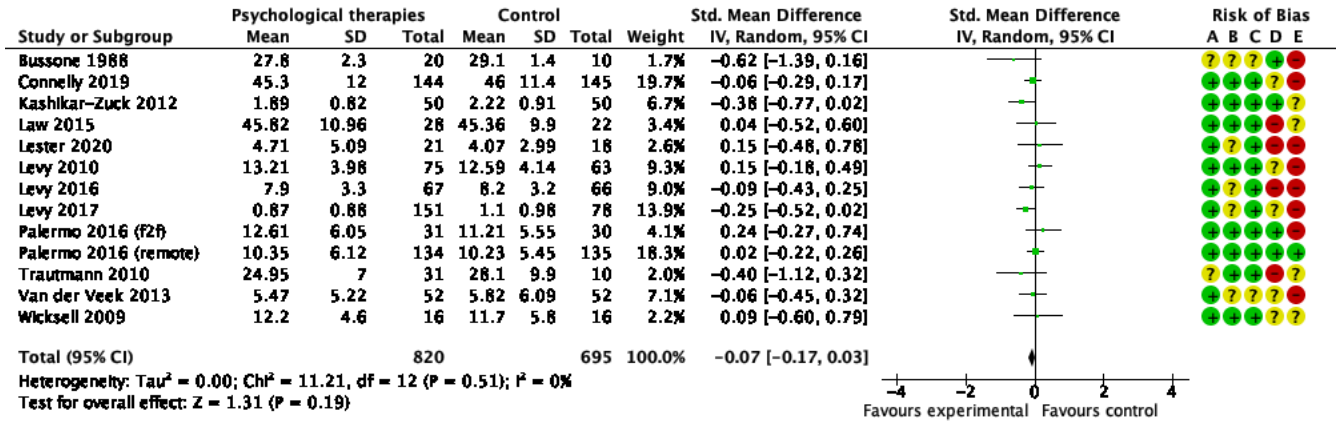


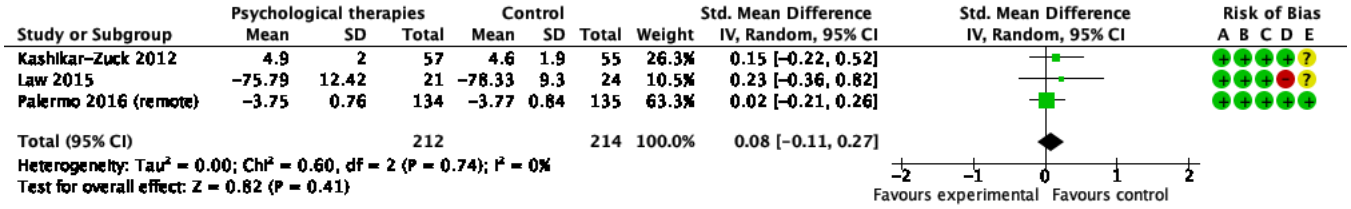
Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Connelly 2019	45.5	11	144	45	11.4	145	21.4%	0.04 [−0.19, 0.28]
Jong 2018	6	4.3	45	5	3.4	41	6.3%	0.25 [−0.17, 0.68]
Kashikar−Zuck 2012	8.7	6.1	57	9.3	5.9	55	8.3%	−0.10 [−0.47, 0.27]
Law 2015	44.75	9.52	28	43.74	6.45	23	3.7%	0.12 [−0.43, 0.67]
Lester 2020	15.93	6.49	21	14.53	4.5	18	2.8%	0.24 [−0.39, 0.87]
Levy 2010	7.89	6.99	78	7.19	5.27	76	11.4%	0.11 [−0.20, 0.43]
Levy 2016	4.4	5.8	67	4.6	5.9	66	9.8%	−0.03 [−0.37, 0.31]
Palermo 2016 (f2f)	11.53	5.37	31	8.71	5.6	30	4.4%	0.51 [−0.00, 1.02]
Palermo 2016 (remote)	9.55	5.13	134	9.49	5.58	135	19.9%	0.01 [−0.23, 0.25]
Trautmann 2010	7.25	6.15	36	6.6	3.7	9	2.1%	0.11 [−0.62, 0.84]
Van der Veek 2013	1.85	1.93	52	1.79	2.14	52	7.7%	0.03 [−0.36, 0.41]
Wicksell 2009	18.1	9.8	16	25.5	16.9	16	2.3%	−0.52 [−1.23, 0.18]
Total (95% CI)			709			666	100.0%	0.06 [−0.05, 0.16]

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 8.05$, df = 11 (P = 0.71); $I^2 = 0\%$
Test for overall effect: Z = 1.01 (P = 0.31)

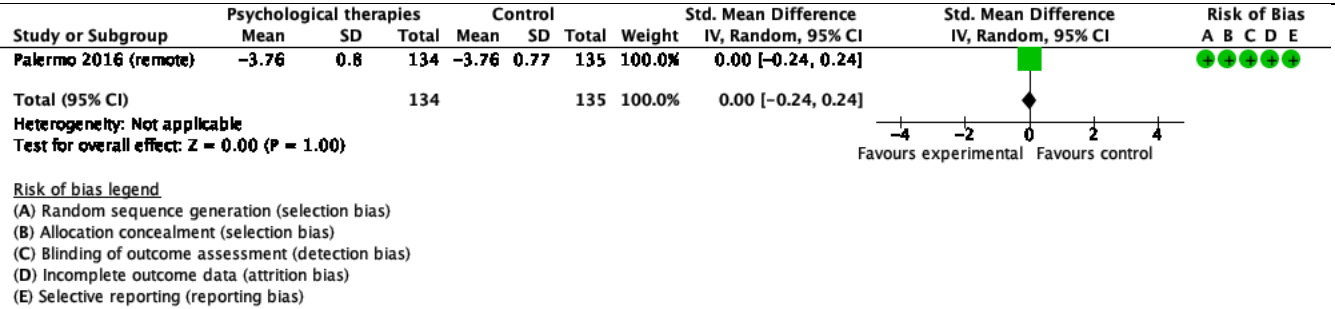
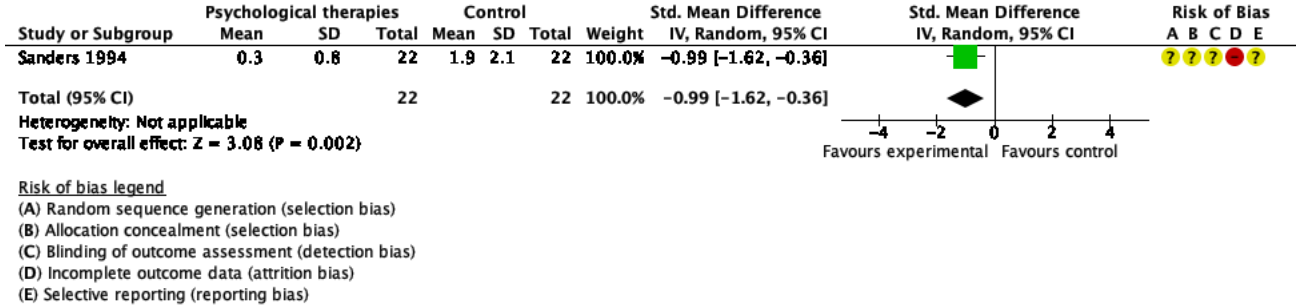
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

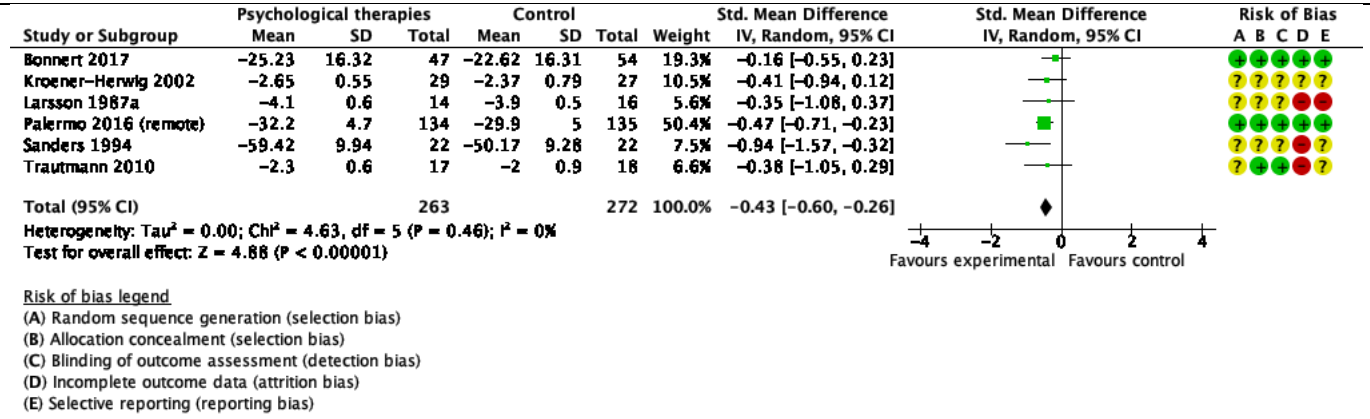
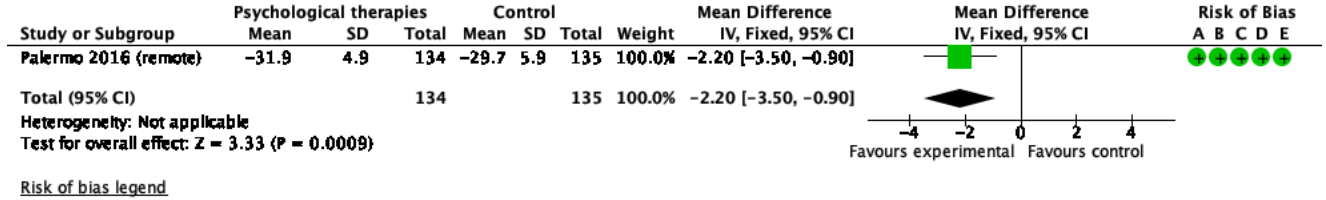
Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher anxious symptomology*	Emotional functioning: Anxiety, post-treatment  	⊕⊕⊕◯ MODERATE
Emotional functioning:	Emotional functioning: Anxiety, follow-up	⊕⊕⊕⊕ HIGH

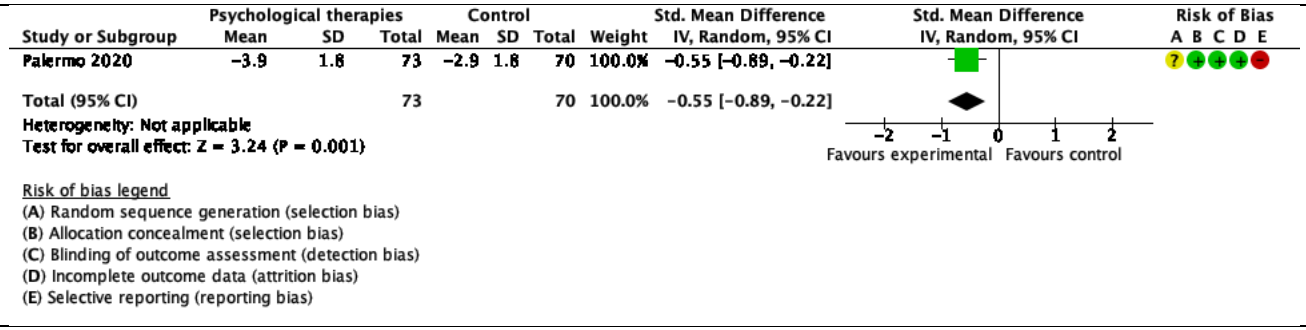
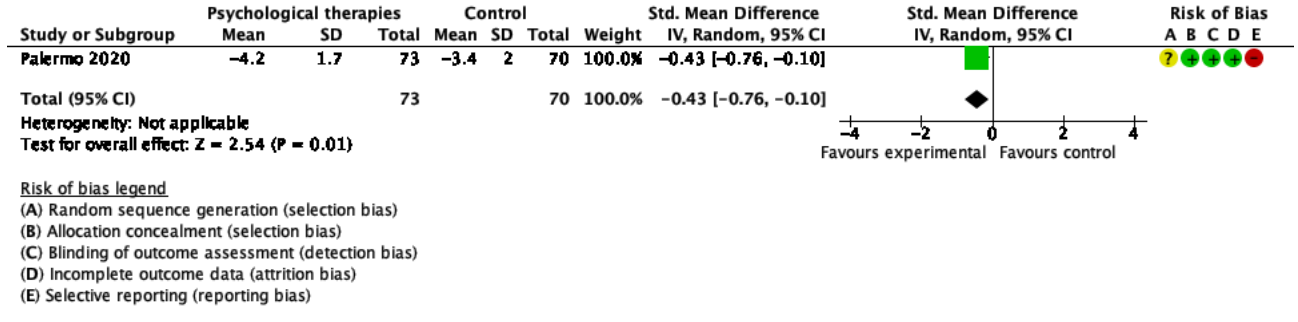
<table>
<tr>
<td>Anxiety, follow-up
Higher scores indicate higher anxious symptomology</td>
<td>


</td>
<td></td>
</tr>
<tr>
<td>Sleep quality, post-treatment
Lower scores indicate worse sleep quality</td>
<td>
Sleep quality, post-treatment


</td>
<td>⊕⊕◯◯
LOW</td>
</tr>
<tr>
<td>Sleep quality, follow-up</td>
<td>Sleep quality, follow-up</td>
<td>⊕◯◯◯
VERY LOW</td>
</tr>
</table>

Lower scores indicate worse sleep quality		
Activity participation, follow-up (no post-treatment data) *Higher scores indicate higher interference with child activities*	Activity participation, follow-up 	⊕◯◯◯ VERY LOW
Global satisfaction with treatment, post-treatment	Global satisfaction with treatment, post-treatment	⊕⊕⊕◯ MODERATE

Lower scores indicate higher satisfaction with treatment		
Global satisfaction with treatment, follow-up *Lower scores indicate higher satisfaction with treatment*	Global satisfaction with treatment, follow-up 	⊕◯◯◯ VERY LOW
Patient global impression of	Patient global impression of change, post-treatment	⊕◯◯◯ VERY LOW

change, post-treatment *Lower scores indicate higher impression of change*		
Patient global impression of change, follow-up *Lower scores indicate higher impression of change*	Patient global impression of change, follow-up 	⊕◯◯◯ VERY LOW

Appendix F.

Appendix F.1. WHO GRADE Profile: Pharmacological therapies vs. any control for children and adolescents with chronic pain

Question: Should pharmacological treatments compared to any control be used for children and adolescents with chronic pain (post-treatment)?

Setting: Any healthcare setting

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pharmacological treatment	any control	Relative (95% CI)	Absolute (95% CI)		
Pain intensity, post-treatment												
5	randomised controlled trials	serious [a]	not serious	not serious	not serious	none	277	346	-	SMD **0.19 lower** (0.35 lower to 0.03 lower)	⊕⊕⊕◯ MODERATE	CRITICAL
30% pain reduction, post-treatment												
2	randomised controlled trials	serious [a]	not serious	serious [b]	serious [c]	none	65/144 (45.1%)	49/142 (34.5%)	**RR 1.33** (1.00 to 1.77)	**114 more per 1,000** (from 0 fewer to 266 more)	⊕◯◯◯ VERY LOW	CRITICAL
50% pain reduction, post-treatment												
2	randomised controlled trials	serious [a]	not serious	serious [b]	serious [c]	none	45/144 (31.3%)	26/142 (18.3%)	**RR 1.71** (1.13 to 2.58)	**130 more per 1,000** (from 24 more to 289 more)	⊕◯◯◯ VERY LOW	CRITICAL
Health-related quality of life, post-treatment												
1	randomised controlled trials	serious [a]	not serious	serious [b]	very serious [d]	none	No studies reported data that could be analysed on health-related quality of life, post-treatment. One study (33 participants) reported the treatment group were more likely to improve quality of life from baseline, compared to placebo.				⊕◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pharmacological treatment	any control	Relative (95% CI)	Absolute (95% CI)		
Functional disability, post-treatment												
1	randomised controlled trials	serious [a]	not serious	serious [b]	very serious [d]	none	91	93	-	SMD **0.1 higher** (0.19 lower to 0.39 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Role functioning, post-treatment - not reported												
0	No studies reported data that could be analysed for role functioning, post-treatment. However, one cross-over trial reported fewer school absences compared to baseline in the treatment group compared to the control group.											CRITICAL
Emotional functioning (depression), post-treatment												
3	randomised controlled trials	not serious	not serious	serious [b]	serious [c]	none	196	193	-	SMD **0.06 lower** (0.25 lower to 0.14 higher)	⊕⊕◯◯ LOW	CRITICAL
Emotional functioning (anxiety), post-treatment												
2	randomised controlled trials	not serious	not serious	serious [b]	serious [c]	none	150	149	-	SMD **0.07 lower** (0.3 lower to 0.16 higher)	⊕⊕◯◯ LOW	CRITICAL
Sleep, post-treatment												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pharmacological treatment	any control	Relative (95% CI)	Absolute (95% CI)		
1	randomised controlled trials	serious [a]	not serious	serious [b]	very serious [d]	none	54	50	-	SMD **0.09 lower** (0.47 lower to 0.3 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Treatment-related serious adverse events												
4	randomised controlled trials	not serious	not serious	serious [b]	very serious [e]	none	3/767 (0.4%)	0/361 (0.0%)	**RD 0.00** (-0.01 to 0.01)	**0 fewer per 1,000** (from 10 fewer to 10 more)	⊕◯◯◯ VERY LOW	CRITICAL
Treatment-related adverse events												
3	randomised controlled trials	serious [a]	serious [f]	serious [b]	not serious	none	153/490 (31.2%)	104/238 (43.7%)	not estimable	**110 fewer per 1,000** (from 280 fewer to 70 more)	⊕◯◯◯ VERY LOW	CRITICAL
Other adverse events - not reported												
	No studies reported other types of adverse events.											CRITICAL
Activity participation, post-treatment - not reported												
1	Non-randomised study	serious [a]	not serious	serious [b]	very serious [d]	none	One non-randomised study (110 participants) reported no differences between groups on activity participation, post-treatment.				⊕◯◯◯ VERY LOW	IMPORTANT
Global judgement of satisfaction with treatment, post-treatment												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pharmacological treatment	any control	Relative (95% CI)	Absolute (95% CI)		
3	randomised controlled trials	not serious	serious [g]	serious [b]	not serious	none	One study (490 participants) reported a higher percentage of subjects treated with sumatriptan and naproxen versus placebo reported being satisfied/very satisfied for "how effective the medication is overall" and "overall satisfaction with medication" at 2 and 24 hours post dose (unadjusted P ≤ .014). Two further studies (205 participants) did not note any differences between groups in the ITT analyses.				⊕⊕◯◯ LOW	IMPORTANT
Patient global impression of change, post-treatment												
1	randomised controlled trials	serious [a]	not serious	serious [b]	very serious [d]	none	One study (104 participants) reported PGIC response was significantly improved with pregabalin versus placebo (P = 0.013), with 53.1% of subjects much improved or very much improved at endpoint with pregabalin, compared with 29.5% with placebo.				⊕◯◯◯ VERY LOW	IMPORTANT
Fatigue, post-treatment - not reported												
No studies reported fatigue, post-treatment												IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

Explanations

a. Downgraded one level for limitations in study design or execution: >50% of risk of bias judgements were rated unclear or high risk of bias.
b. Downgraded by one level for indirectness: few conditions presented in the meta-analysis.
c. Downgraded by one level for imprecision: small number of participants (<400 participants) or studies (<2 studies) contributing to the analyses.
d. Downgraded by two levels for imprecision: very small number of participants (<200 participants) or studies (<2 studies) contributing to the analyses.
e. Downgraded by two levels for serious imprecision: very few events
f. Downgraded one level for inconsistency: unexplained statistical heterogeneity >50%.
g. Downgraded by one level for inconsistency: unable to combine results in meta-analysis and estimates from the different studies were contradictory leading to inconsistency.

WHO GRADE Profile: Pharmacological therapies vs. any control for children and adolescents with chronic pain (follow-up)

Question: Should pharmacological treatments compared to any control be used for children and adolescents with chronic pain (follow-up, within 12 months)?

Setting: Global

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pharmacological treatment	any control	Relative (95% CI)	Absolute (95% CI)		
Pain intensity, follow-up												
2	randomised controlled trials	serious [a]	not serious	serious [b]	very serious [c]	none	75	73	-	SMD **0.22 lower** (0.54 lower to 0.1 higher)	⨁◯◯◯ VERY LOW	CRITICAL
30% pain reduction, follow-up - not reported												
	No studies reported 30% pain reduction at follow-up.										-	CRITICAL
50% pain reduction, follow-up - not reported												
-	-	-	-	-	-	-	No randomised controlled studies reported 50% pain reduction at follow-up. One crossover trial reported 56/58 participants reached 50% pain reduction in the two treatment groups, and 25/29 in the control group.					CRITICAL
Health-related quality of life, follow-up												
1	randomised controlled trials	serious [a]	not serious	serious [b]	very serious [c]	none	No studies reported data that could be analysed on health-related quality of life, post-treatment. One study (33 participants) reported the treatment group were more likely to improve quality of life from baseline, compared to placebo.				⨁◯◯◯ VERY LOW	CRITICAL
Functional disability, follow-up - not reported												
-	No studies reported functional disability at follow-up.										-	CRITICAL
Role functioning, follow-up - not reported												
-	No studies reported role functioning at follow-up.										-	CRITICAL
Emotional functioning (depression), follow-up												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pharmacological treatment	any control	Relative (95% CI)	Absolute (95% CI)		
1	randomised controlled trials	not serious	not serious	serious [b]	very serious [c]	none	59	56	-	SMD **0.26 lower** (0.63 lower to 0.11 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Emotional functioning (anxiety), follow-up												
1	randomised controlled trial	not serious	not serious	serious [b]	very serious [c]	none	59	56	-	SMD **0.03 higher** (0.34 lower to 0.39 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Sleep, follow-up - not reported												
-	No studies reported sleep at follow-up.										-	CRITICAL
Activity participation, follow-up - not reported												
-	No studies reported activity participation at follow-up.										-	IMPORTANT
Global judgement of satisfaction with treatment, follow-up												
1	randomised controlled trial	not serious	not serious	serious [b]	very serious [c]	none	One study (115 participants) reported no differences between groups in the ITT analyses at follow-up (p =0.491).				⊕◯◯◯ VERY LOW	IMPORTANT
Patient global impression of change, follow-up - not reported												
-	No studies reported patient global impression of change at follow-up.										-	IMPORTANT
Fatigue, follow-up - not reported												
-	No studies reported fatigue at follow-up.										-	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

a. Downgraded one level for limitations in study design or execution: >50% of risk of bias judgements were rated unclear or high risk of bias.
b. Downgraded by one level for indirectness: few conditions presented in the meta-analysis.
c. Downgraded by two levels for imprecision: very small number of participants (<200 participants) or studies (<2 studies) contributing to the analyses.

Appendix F.2. WHO GRADE Profiles: Physical therapies vs. any control for children and adolescents with chronic pain

Question: Should physical therapies compared to any control be used for children with chronic pain (post-treatment)?

Setting: Global

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Physical therapies	control	Relative (95% CI)	Absolute (95% CI)		
Pain intensity, post-treatment												
6	randomised trials	serious [a]	very serious [b]	not serious	serious [c]	none	179	195	-	SMD **0.6 lower** (1.15 lower to 0.04 lower)	⊕◯◯◯ VERY LOW	CRITICAL
30% pain reduction, post-treatment - not reported												
-	-	-	-	-	-	-	No studies reported 30% pain reduction.				-	CRITICAL
50% pain reduction, post-treatment - not reported												
-	-	-	-	-	-	-	No studies reported 50% pain reduction.				-	CRITICAL
Health-related quality of life, post-treatment												
2	randomised trials	serious [a]	very serious [b]	serious [d]	very serious [e]	none	81	52	-	SMD **0.64 lower** (1.91 lower to 0.63 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Functional disability, post-treatment												
4	randomised trials	serious [a]	not serious	not serious	very serious [e]	none	92	82	-	SMD **0.64 lower** (0.95 lower to 0.34 lower)	⊕◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Physical therapies	control	Relative (95% CI)	Absolute (95% CI)		
Role functioning, post-treatment												
2	randomised trials	very serious [a]	not serious	serious [d]	very serious [e]	none	No differences were found between groups in either study (93 participants) post-treatment.				⨁◯◯◯ VERY LOW	CRITICAL
Emotional functioning (depression), post-treatment												
3	randomised trials	serious [a]	not serious	not serious	very serious [e]	none	49	44	-	SMD **0.25 lower** (0.66 lower to 0.16 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Emotional functioning (anxiety), post-treatment												
2	randomised trials	serious [a]	very serious [b]	not serious	very serious [e]	none	32	25	-	SMD **0.06 higher** (1.39 lower to 1.51 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Sleep, post-treatment - not reported												
-	-	-	-	-	-	-	No studies reported sleep outcomes, post-treatment.				-	CRITICAL
Treatment-related serious adverse events, post-treatment - not reported												
-	-	-	-	-	-	-	No studies reported treatment-related serious adverse events, post-treatment.				-	
Treatment-related adverse events												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Physical therapies	control	Relative (95% CI)	Absolute (95% CI)		
4	randomised trials	serious [a]	not serious	not serious	very serious [e]	none	1/81 (1.2%)	0/80 (0.0%)	**RD 0.01** (-0.04 to 0.05)	**10 fewer per 1,000** (from 50 fewer to 40 more)	⊕◯◯◯ VERY LOW	CRITICAL
Other adverse events, post-treatment - not reported												
-	-	-	-	-	-	-	No studies reported other adverse events, post-treatment.				-	CRITICAL
Activity participation, post-treatment												
1	randomised trials	very serious [a]	not serious	serious [d]	very serious [e]	none	Fewer activity participation absences were reported in the treatment group (n = 27) compared to control group (n = 27).				⊕◯◯◯ VERY LOW	IMPORTANT
Global judgement of satisfaction with treatment - not reported												
-	-	-	-	-	-	-	No studies reported global judgement of satisfaction with treatment, post-treatment.				-	IMPORTANT
Patient global impression of change												
1	randomised trials	serious [a]	not serious	serious [d]	very serious [e]	none	18/21 reported 'slight but noticeable change' and 10/21 reported 'definite improvement' in the treatment group. 1/22 reported 'slight but noticeable' or 'definite improvement' in the control group.				⊕◯◯◯ VERY LOW	IMPORTANT
Fatigue, post-treatment - not reported												
-	-	-	-	-	-	-	No studies reported fatigue outcomes, post-treatment.				-	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

a. Downgraded by one level for limitations in study design or execution: >50% of risk of bias judgements were rated unclear or high risk of bias.
b. Downgraded by two levels for serious inconsistency: unexplained statistical heterogeneity >75%.
c. Downgraded by one level for imprecision: small number of participants (<400 participants) or studies (<2 studies) contributing to the analyses.
d. Downgraded by one level for indirectness: few conditions presented in the meta-analysis.
e. Downgraded by two levels for serious imprecision: very small number of participants (<200 participants) or studies (<2 studies) contributing to the analyses.

WHO GRADE Profiles: Physical therapies vs. any control for children and adolescents with chronic pain at follow-up (within 12 months)

Question: Should physical therapies compared to any control be used for children with chronic pain (follow-up, within 12 months)?

Setting: Global

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Physical therapies	any control	Relative (95% CI)	Absolute (95% CI)		
Pain intensity, follow-up												
3	randomised trials	serious [a]	serious [b]	not serious	very serious [c]	none	85	102	-	SMD **0.13 lower** (0.74 lower to 0.48 higher)	⊕◯◯◯ VERY LOW	CRITICAL
30% pain reduction, follow-up - not reported												
0	No studies reported 30% pain reduction at follow-up.											CRITICAL
50% pain reduction, follow-up - not reported												
0	No studies reported 50% pain reduction at follow-up.											CRITICAL
Health-related quality of life, follow-up - not reported												
0	No studies reported overall quality of life at follow-up.											CRITICAL
Role functioning, follow-up												
1	randomised trials	serious [a]	not serious	serious [d]	very serious [c]	none	One study reported role/social physical functioning at follow-up, data were not presented in a way that allowed it to be entered in a meta-analysis. No differences were reported between groups.				⊕◯◯◯ VERY LOW	CRITICAL
Functional disability, follow-up												
1	randomised trials	not serious	not serious	serious [d]	very serious [c]	none	17	19	-	SMD **0.38 lower** (1.04 lower to 0.28 higher)	⊕◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Physical therapies	any control	Relative (95% CI)	Absolute (95% CI)		
Emotional functioning (depression), follow-up												
1	randomised trials	not serious	not serious	serious [d]	very serious [c]	none	17	19	-	SMD **0.22 lower** (0.88 lower to 0.44 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Emotional functioning (anxiety), follow-up - not reported												
	No studies reported emotional functioning (anxiety) at follow-up.											CRITICAL
Sleep, follow-up - not reported												
	No studies reported sleep at follow-up.											CRITICAL
Activity participation, follow-up - not reported												
	No studies reported activity participation at follow-up.											IMPORTANT
Global judgement of satisfaction with treatment - not reported												
	No studies reported global judgement of satisfaction with treatment at follow-up.											IMPORTANT
Patient global impression of change, follow-up - not reported												
	No studies reported patient global impression of change at follow-up.											IMPORTANT
Fatigue - not reported												
	No studies reported fatigue at follow-up.											IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference

Explanations

a. Downgraded one level for limitations in study design or execution: >50% of risk of bias judgements were rated unclear or high risk of bias.
b. Downgraded one level for inconsistency: unexplained statistical heterogeneity >50%.

c. Downgraded by two levels for serious imprecision: very small number of participants (<200 participants) or studies (<2 studies) contributing to the analyses.
d. Downgraded by one level for indirectness: few conditions presented in the meta-analysis.

Appendix F.3. WHO GRADE Profile: Psychological therapies vs. any control for children and adolescents with chronic pain, post-treatment

Question: Psychological therapies compared to any control in children and adolescents with chronic pain (post-treatment)

Setting: Global

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
Pain intensity, post-treatment												
38	randomised trials	serious [a]	serious [b]	not serious	not serious	none	1584	1441	-	SMD **0.29 lower** (0.43 lower to 0.16 lower)	⊕⊕◯◯ LOW	CRITICAL
30% pain reduction, post-treatment												
1	randomised trials	very serious [c]	not serious	not serious	very serious [d]	none	17/52 (32.7%)	15/52 (28.8%)	**RR 1.13** (0.64 to 2.02)	**37 more per 1,000** (from 104 fewer to 294 more)	⊕◯◯◯ VERY LOW	CRITICAL
50% reduction in pain, post-treatment												
22	randomised trials	serious [a]	not serious	serious [e]	not serious	none	307/644 (47.7%)	104/496 (21.0%)	**RR 2.11** (1.61 to 2.77)	**233 more per 1,000** (from 128 more to 371 more)	⊕⊕◯◯ LOW	CRITICAL
Health-related quality of life, post-treatment												
13	randomised trials	serious [a]	serious [b]	not serious	not serious	none	703	594	-	SMD **0.14 SD lower** (0.33 lower to 0.05 higher)	⊕⊕◯◯ LOW	CRITICAL
Functional disability, post-treatment												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
24	randomised trials	serious [a]	serious [b]	not serious	not serious	none	1209	1149	-	SMD **0.25 lower** (0.39 lower to 0.11 lower)	⊕⊕◯◯ LOW	CRITICAL
Role functioning (school absence), post-treatment												
9	randomised trials	serious [a]	very serious [f]	not serious	not serious	none	483	373	-	SMD **0.21 SD lower** (0.52 lower to 0.1 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Emotional functioning (depression), post-treatment												
19	randomised trials	not serious	not serious	not serious	not serious	none	917	864	-	SMD **0.02 lower** (0.11 lower to 0.08 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Emotional functioning (anxiety), post-treatment												
19	randomised trials	serious [a]	not serious	not serious	not serious	none	1084	947	-	SMD **0.08 lower** (0.21 lower to 0.04 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Sleep quality, post-treatment												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials	not serious	not serious	very serious [e,g]	not serious	none	212	214	-	SMD **0.08 SD lower** (0.11 lower to 0.27 higher)	⊕⊕◯◯ LOW	CRITICAL
Adverse events												
7	randomised trials	not serious	not serious	serious [e]	very serious [h]	none	5 studies (524 participants) reported no adverse events (SAEs, TAEs, and other AEs) in any trial arm. One study (135 participants) reported more AEs in the control arm (education + amitriptyline) compared to treatment arm, and most were attributed to amitriptyline. A final study (43 participants) reported mild headache in the treatment arm when listening to CDs.				⊕◯◯◯ VERY LOW	CRITICAL
Activity participation, post-treatment												
0	randomised trials	No studies assessed activity participation post-treatment.										IMPORTANT
Global satisfaction with treatment, post-treatment												
6	randomised trials	serious [a]	not serious	not serious	not serious	none	263	272	-	SMD **0.43 lower** (0.6 lower to 0.26 lower)	⊕⊕⊕◯ MODERATE	IMPORTANT
Patient Global Impression of Change, post-treatment												
1	randomised trials	not serious	not serious	serious [e]	very serious [i]	none	73	70	-	SMD **0.55 lower** (0.89 lower to 0.22 lower)	⊕◯◯◯ VERY LOW	IMPORTANT
Fatigue, post-treatment												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
0	randomised trials	No studies assessed fatigue post-treatment.										IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

Explanations

a. Downgraded one level for limitations in study design or execution: >50% of risk of bias judgements were rated unclear or high risk of bias.
b. Downgraded one level for inconsistency: unexplained statistical heterogeneity >50%.
c. Downgraded two levels for serious limitations in study design or execution: >75% of risk of bias judgements were rated unclear or high risk of bias.
d. Downgraded by two levels for serious imprecision: very small number of participants (<200 participants) or studies (<2 studies) contributing to the analyses.
e. Downgraded by one level for indirectness: few conditions presented in the meta-analysis so estimate may not be applicable to other chronic pain conditions.
f. Downgraded by two levels for serious inconsistency: unexplained statistical heterogeneity >75%.
g. Downgraded by one level for indirectness: 2/3 studies came from same the same setting.
h. Downgraded by two levels for serious imprecision: small number of events.
i. Downgraded by one level for imprecision: small number of participants (<400 participants) or studies (<2 studies) contributing to the analyses.

WHO GRADE Profile: Psychological therapies vs. any control for children and adolescents with chronic pain, follow-up (up to 12 months)
Question: Psychological therapies compared to any control in children and adolescents with chronic pain (follow-up; up to 12 months)
Setting: Global

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
Pain intensity												
21	randomised controlled trials	serious [a]	serious [b]	not serious	not serious	none	997	884	-	SMD **0.14 lower** (0.3 lower to 0.02 higher)	⨁⨁◯◯ LOW	CRITICAL
30% pain reduction												
1	randomised controlled trials	very serious [c]	not serious	serious [d]	very serious [e]	none	31/52 (59.6%)	29/52 (55.8%)	**RR 1.07** (0.77 to 1.49)	**39 more per 1,000** (from 128 fewer to 273 more)	⨁◯◯◯ VERY LOW	CRITICAL
50% reduction in pain,												
9	randomised controlled trials	serious [a]	serious [b]	serious [d]	not serious	none	109/242 (45.0%)	46/203 (22.7%)	**RR 2.09** (1.29 to 3.38)	**247 more per 1,000** (from 66 more to 539 more)	⨁◯◯◯ VERY LOW	CRITICAL
Health-related quality of life												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
67	randomised controlled trials	serious [a]	not serious	not serious	not serious	none	449	346	-	SMD **0.09 SD higher** (0.35 lower to 0.16 higher)	⊕⊕◯◯ LOW	CRITICAL
Functional disability												
14	randomised controlled trials	not serious	serious [b]	not serious	not serious	none	914	841	-	SMD **0.23 SD lower** (0.38 lower to 0.08 lower)	⊕⊕⊕◯ MODERATE	CRITICAL
Role functioning (school absence)												
4	randomised controlled trials	serious [a]	very serious [f]	not serious	not serious	none	270	206	-	SMD **0.14 SD higher** (0.32 lower to 0.6 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Emotional functioning (depression)												
12	randomised controlled trials	not serious	not serious	not serious	not serious	none	709	666	-	SMD **0.06 higher** (0.05 lower to 0.16 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Emotional functioning (anxiety)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
13	randomised controlled trials	not serious	not serious	not serious	not serious	none	820	695	-	SMD **0.07 lower** (0.17 lower to 0.03 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Sleep quality												
1	randomised controlled trials	not serious	not serious	serious [d]	very serious [g]	none	134	135	-	SMD **0 SD** (0.24 lower to 0.24 higher)	⊕⊕⊕◯ VERY LOW	CRITICAL
Activity participation												
1	randomised controlled trials	very serious [c]	not serious	serious [d]	very serious [g]	none	22	22	-	SMD **0.99 lower** (1.62 lower to 0.36 lower)	⊕◯◯◯ VERY LOW	IMPORTANT
Global satisfaction with treatment												
1	randomised controlled trials	not serious	not serious	serious [d]	very serious [g]	none	134	135	-	MD **2.2 lower** (3.5 lower to 0.9 lower)	⊕◯◯◯ VERY LOW	IMPORTANT
Patient Global Impression of Change												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychological therapies	any control	Relative (95% CI)	Absolute (95% CI)		
1	randomised controlled trials	not serious	not serious	serious [d]	very serious [g]	none	73	70	-	SMD **0.43 lower** (0.76 lower to 0.1 lower)	⨁◯◯◯ VERY LOW	IMPORTANT

Fatigue, follow-up - not reported

0	No studies assessed fatigue at follow-up.	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio; **MD:** Mean difference

Explanations

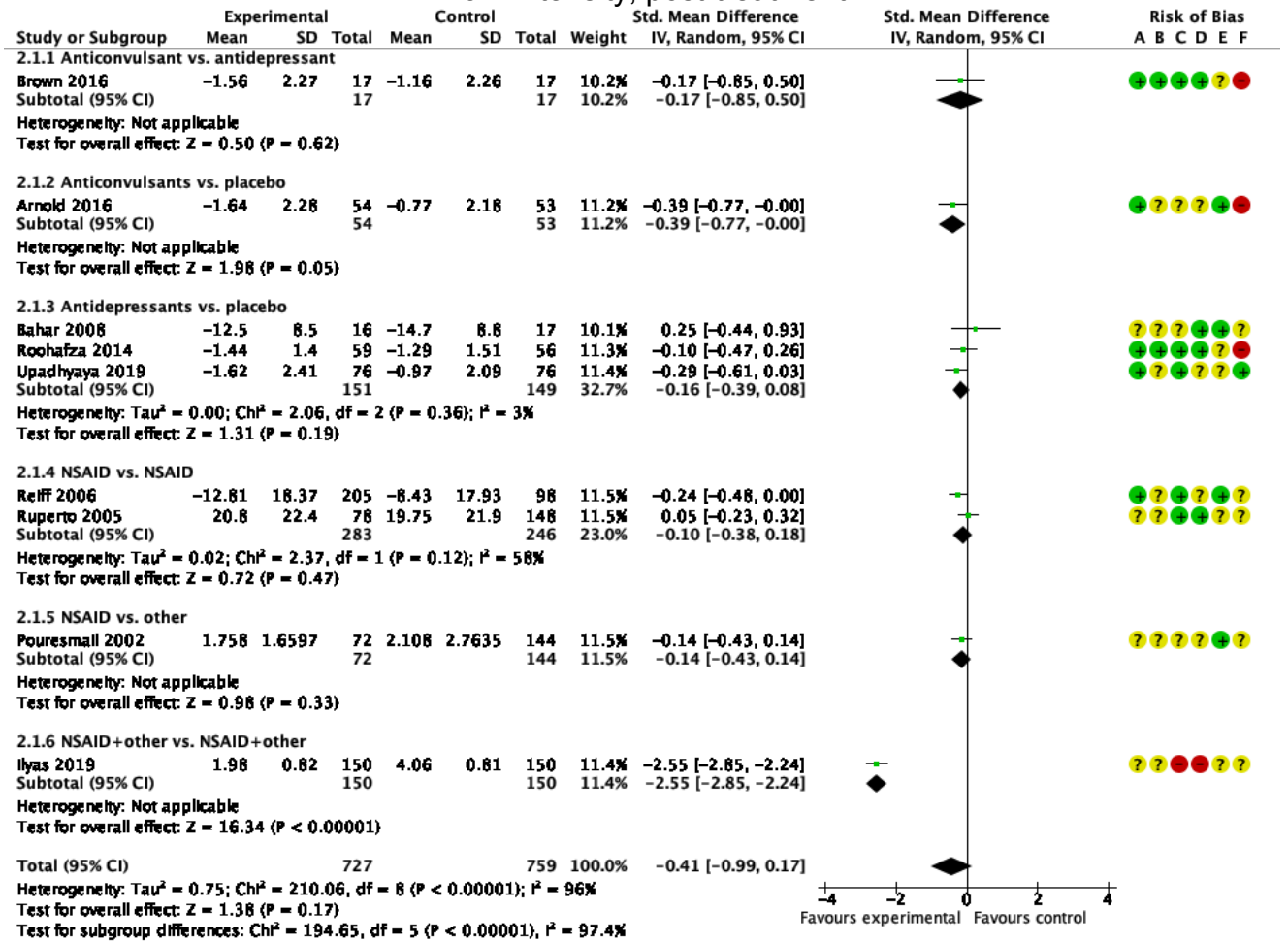
a. Downgraded one level for limitations in study design or execution: >50% of risk of bias judgements were rated unclear or high risk of bias.
b. Downgraded one level for inconsistency: unexplained statistical heterogeneity >50%.
c. Downgraded two levels for serious limitations in study design or execution: >75% of risk of bias judgements were rated unclear or high risk of bias.
d. Downgraded by one level for indirectness: few conditions presented in the meta-analysis so estimate may not be applicable to other chronic pain conditions.
e. Downgraded by one level for imprecision: small number of participants (<400 participants) or studies (<2 studies) contributing to the analyses.
f. Downgraded two levels for serious inconsistency: unexplained statistical heterogeneity >75%.
g. Downgraded by two levels for serious imprecision: very small number of participants (<200 participants) or studies (<2 studies) contributing to the analyses.

Appendix G

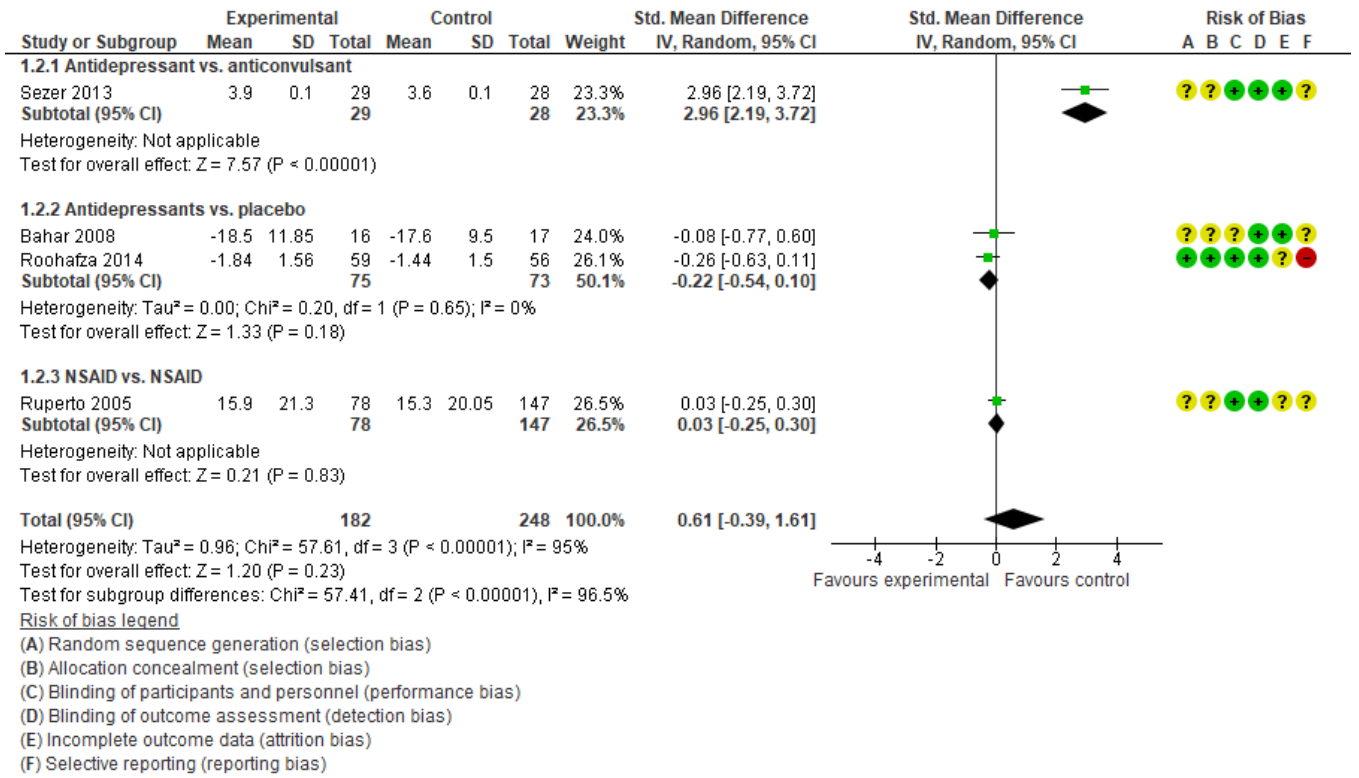
Appendix G.1. WHO review: Pharmacological interventions for children with chronic pain

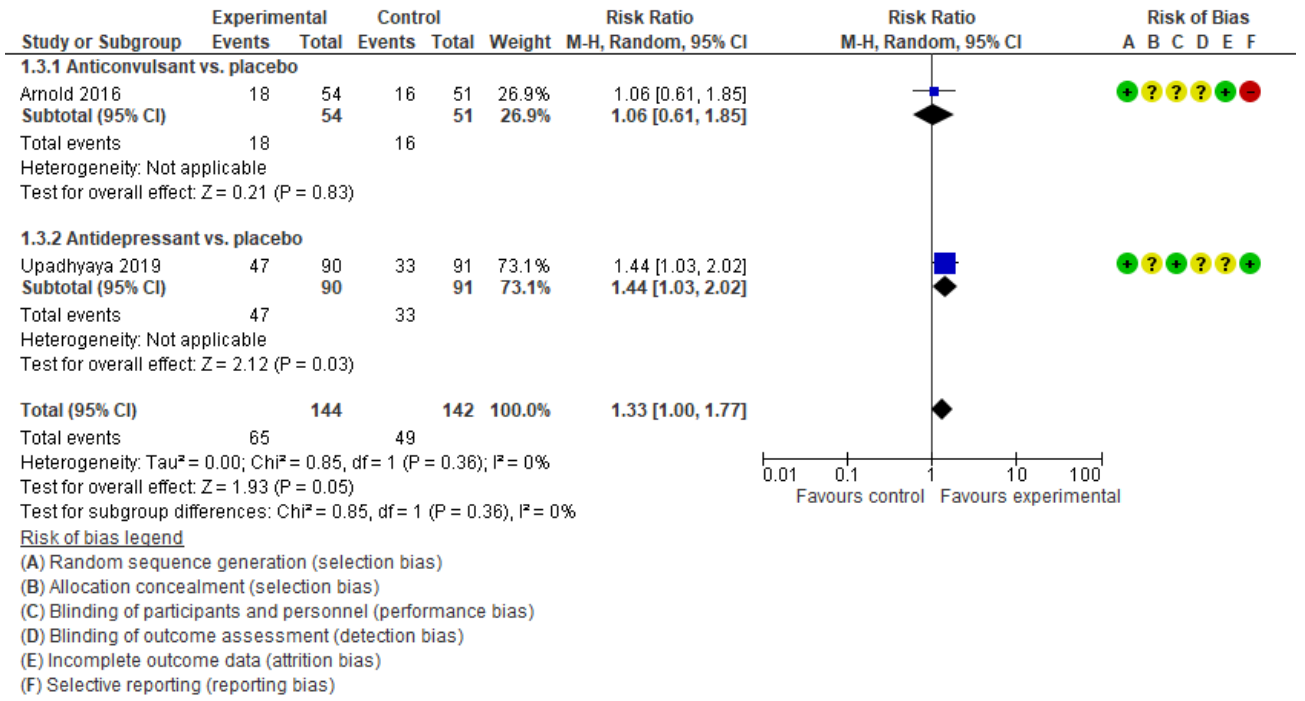
Comparison: Pharmacological therapies versus placebo, waitlist control, or other pharmacological control
Population: Children and adolescents with chronic pain
Setting: Any setting
Studies: Randomised controlled trials

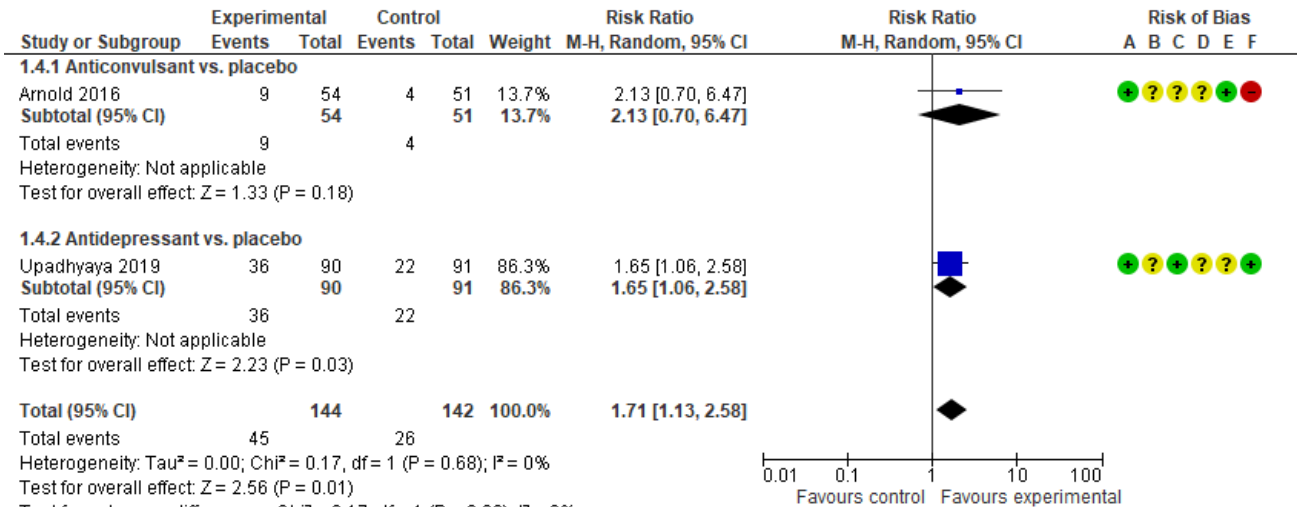
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

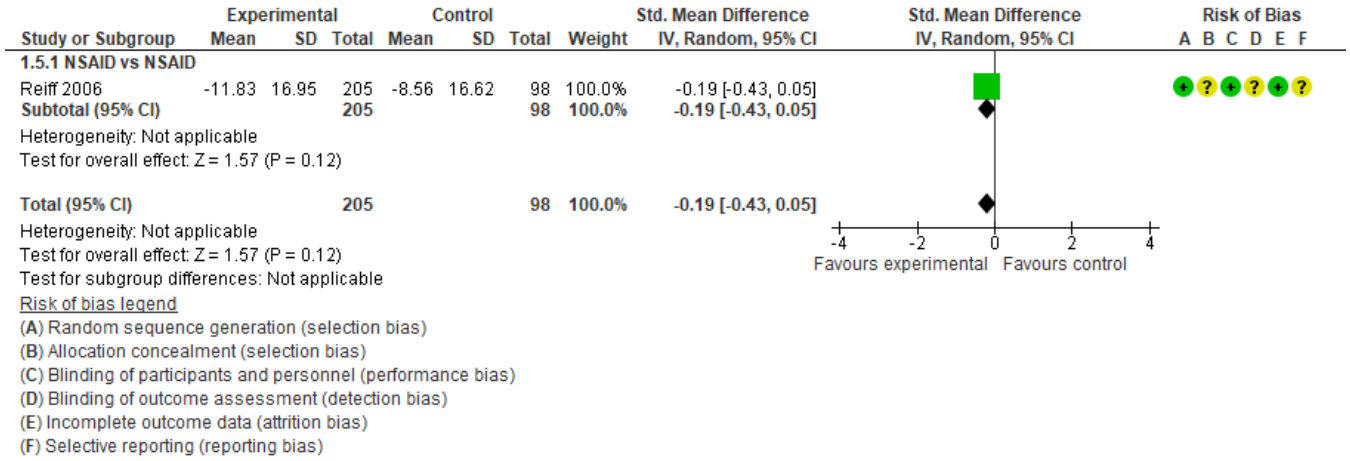
Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Bahar 2008: Amitriptyline vs. placebo; Chronic secondary visceral pain (IBS), 14 years.	Anticonvulsants vs. antidepressants: ⊕◯◯◯ VERY LOW Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕⊕◯◯ LOW NSAID vs NSAID: ⊕◯◯◯ VERY LOW NSAID vs. other: ⊕◯◯◯ VERY LOW

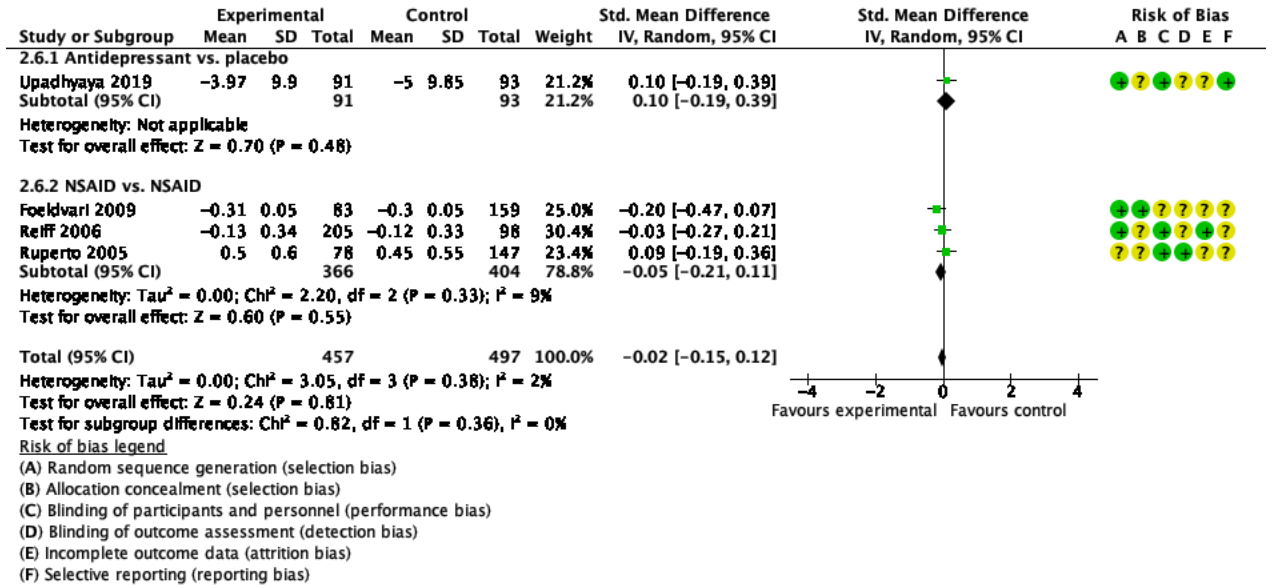
	• Brown 2016: Gabapentin vs. amitriptyline; CRPS/neuropathic pain, 13 years. • Ilyas 201: Mefenamic acid plus vitamin E vs. mefenamic acid; dysmenorrhea, 15 years. • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Reiff 2006: Rofecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years. • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years. • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years • Pouresmall 2002: Ibuprofen vs. acupressure or sham acupressure; Chronic primary visceral pain (Dysmenorrhea), 14-18 years	

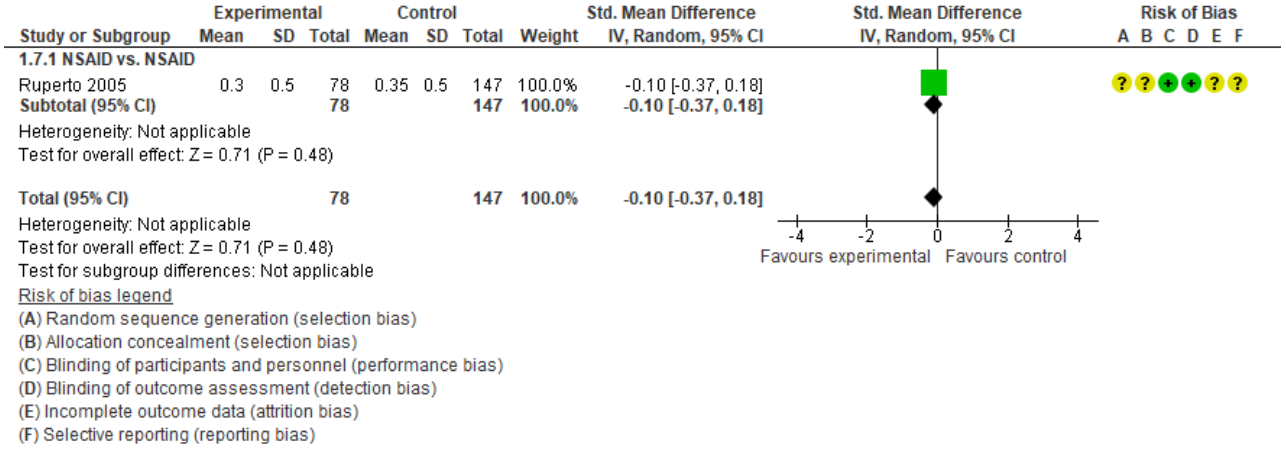
Pain intensity, follow-up *Higher scores indicate higher pain intensity*	Pain intensity, follow-up  • Sezer 2014: Amitriptyline vs. topiramate; mixed pain conditions, 15 years. • Bahar 2008: Amitriptyline vs. placebo; Chronic secondary visceral pain (IBS), 14 years • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Ruperto 2005: Meloxicam vs. naproxen	Antidepressants vs. anticonvulsants: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕⊕◯◯ LOW NSAID vs. NSAID: ⊕◯◯◯ VERY LOW

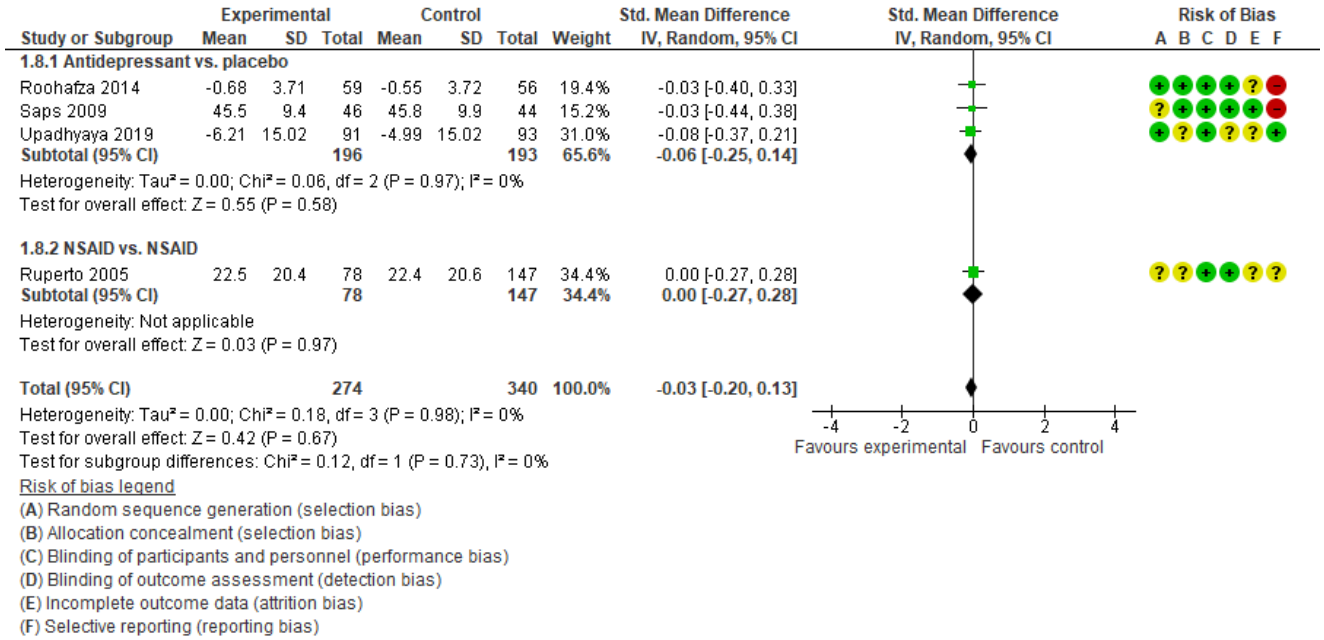
30% reduction, post-treatment	30% reduction, post-treatment  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

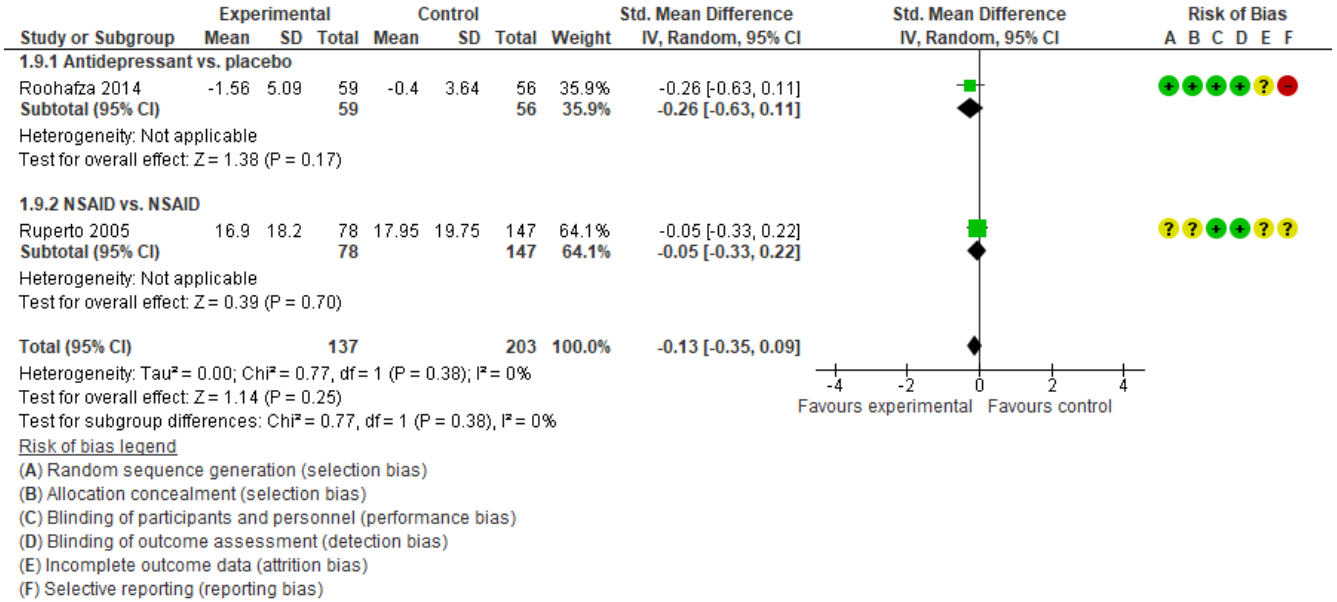
50% pain reduction, post-treatment	50% pain reduction  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Anticonvulsants vs. placebo: ⨁◯◯◯ VERY LOW Antidepressants vs. placebo: ⨁◯◯◯ VERY LOW

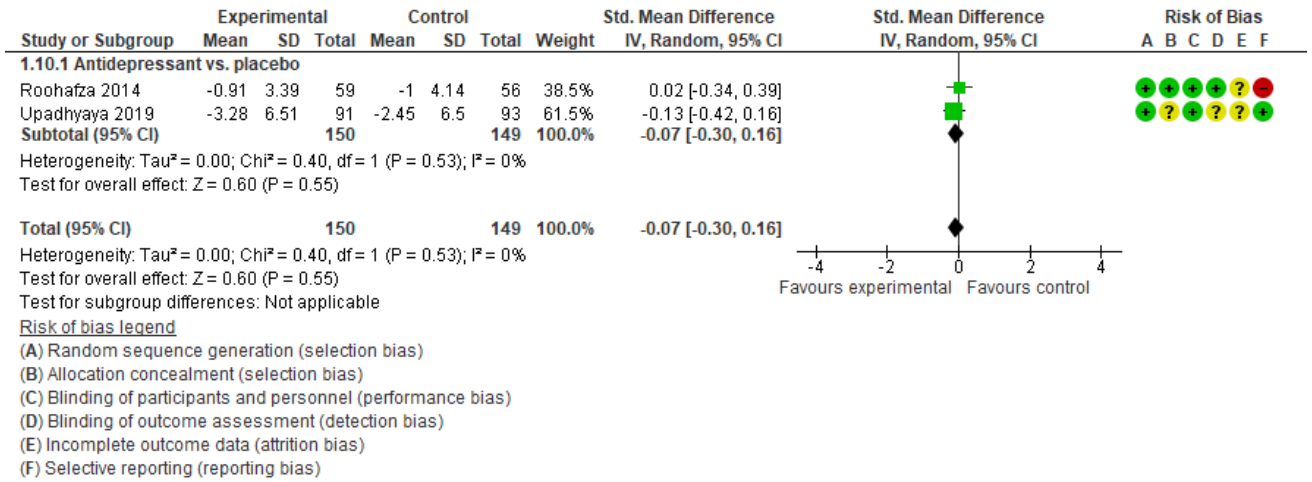
Health-related quality of life, post-treatment *Higher scores indicate better quality of life*	Health-related quality of life, post-treatment  • Reiff 2006: Rofecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years.	NSAID vs. NSAID: ⊕◯◯◯ VERY LOW

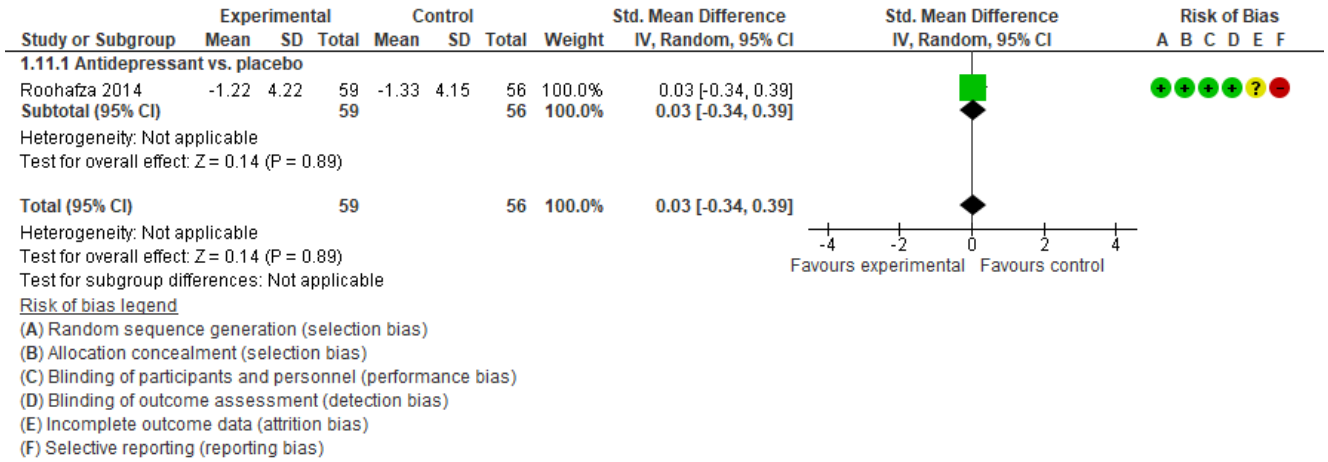
Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment  • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years. • Foedvari 2009: Celecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years. • Reiff 2006: Rofecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years. • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years.	Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW NSAID vs. NSAID: ⊕◯◯◯ VERY LOW

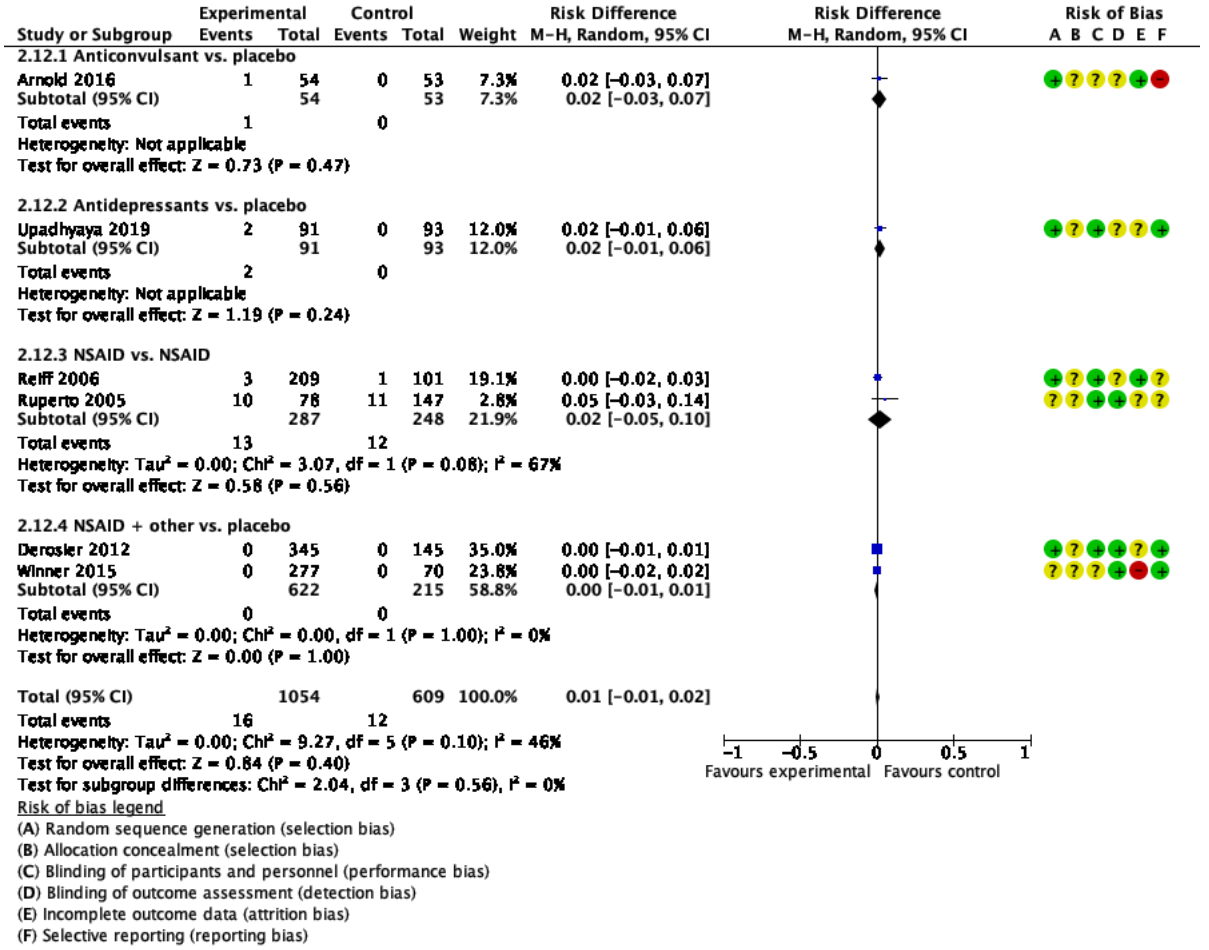
Functional disability, follow-up *Higher scores indicate lower Disability*	Functional disability, follow-up  • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years.	NSAID vs. NSAID: ⊕◯◯◯ VERY LOW

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, post-treatment  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years. • Saps 2009: Amitriptyline vs. placebo; Mixed pain (FAP, Functional dyspepsia, IBS), 12 years. • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years. • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years.	Antidepressants vs. placebo: ⊕⊕◯◯ LOW NSAID vs. NSAID: ⊕◯◯◯ VERY LOW

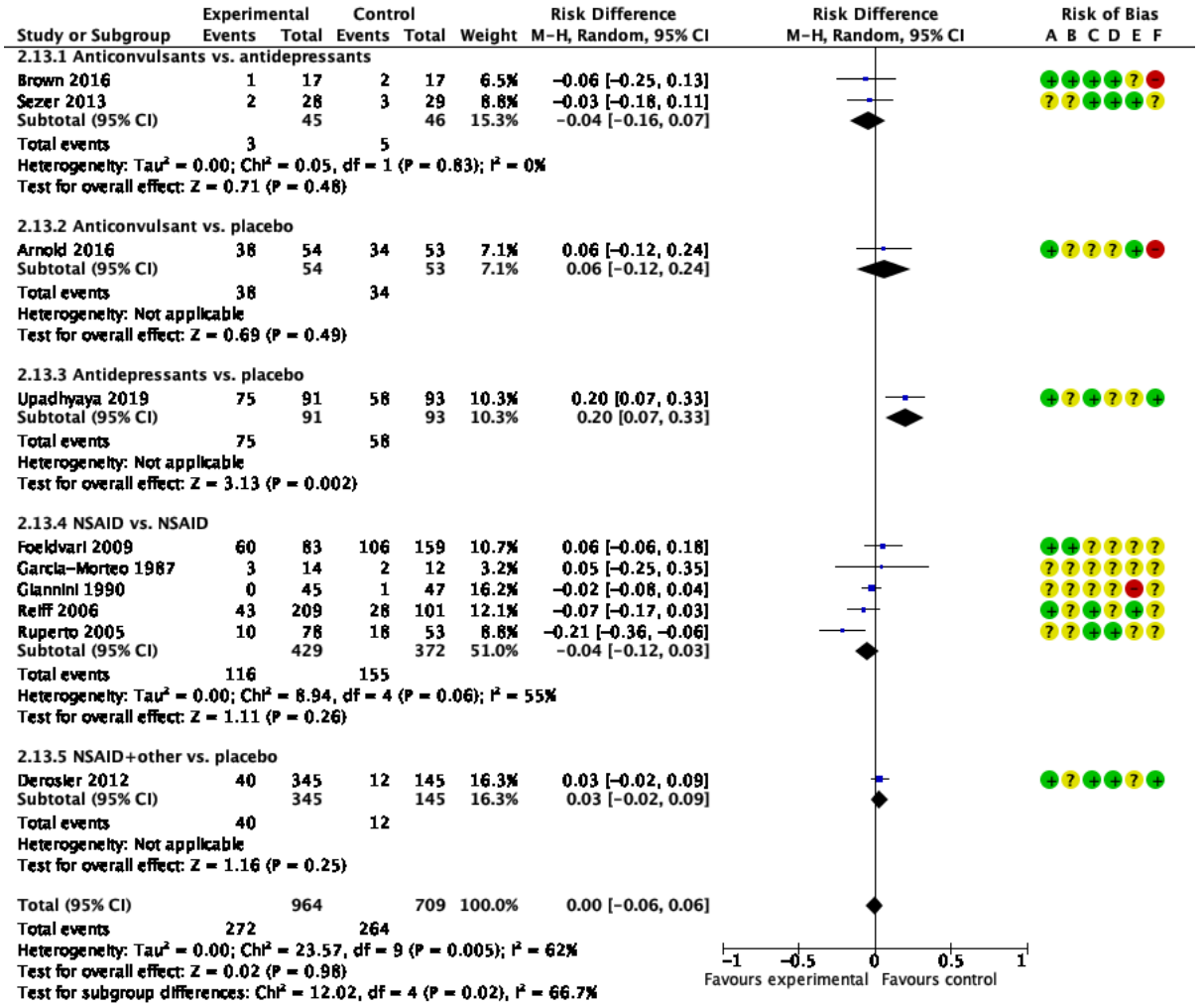
Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow up  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years.	Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW NSAID vs. NSAID: ⊕◯◯◯ VERY LOW

Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Anxiety, post-treatment  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years	Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

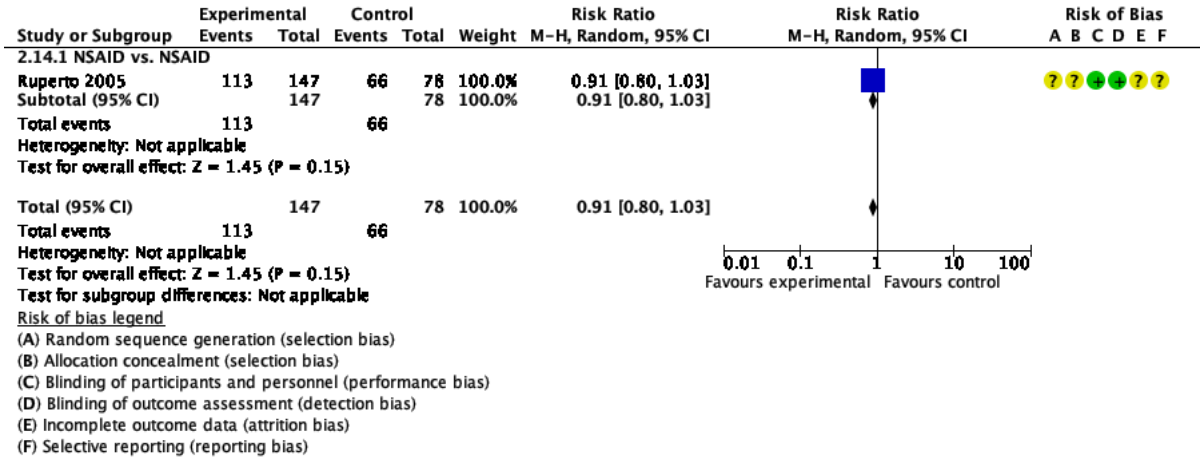
Emotional functioning: Anxiety, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Anxiety, follow-up  • Roohafza 2014: Citalopram vs. placebo; Chronic primary visceral pain (Functional abdominal pain), 9 years	Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW

Treatment-related serious adverse events	Treatment-related serious adverse events  • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years • Reiff 2006: Rofecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years.	Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW NSAID vs NSAID: ⊕◯◯◯ VERY LOW NSAID+ other vs. NSAID: ⊕◯◯◯ VERY LOW

	• Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years. • Derosier 2012: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine). • Winner 2015: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine), 15 years, 15 years.	

Treatment-related adverse events	Treatment-related adverse events  • Brown 2016: Amitriptyline vs gabapentin, Chronic regional pain syndrome, 13 years.	Anticonvulsants vs. placebo: ⊕◯◯◯ VERY LOW Antidepressants vs. placebo: ⊕◯◯◯ VERY LOW NSAID vs NSAID: ⊕◯◯◯ VERY LOW NSAID+ other vs. placebo: ⊕◯◯◯ VERY LOW

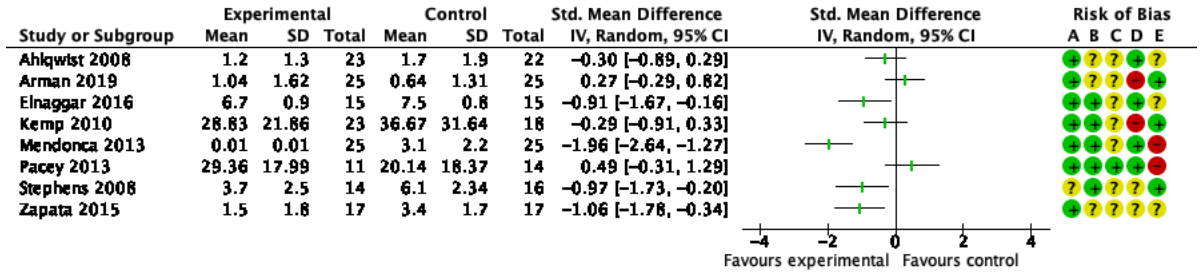
	• Sezer 2013: Amitriptyline vs topiramate, mixed chronic pain conditions, 15 years. • Arnold 2016: Pregabalin vs. placebo; Chronic widespread pain (Fibromyalgia), 14 years • Upadhyaya 2019: Duloxetine vs. placebo; Chronic widespread pain (Fibromyalgia), 15 years • Foeldvari 2009: Celecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years. • Garcia-Morteo 1978: Piroxicam vs. naproxen; juvenile idiopathic arthritis, 8 years. • Giannini 1990: Aspirin vs. ibuprofen; juvenile idiopathic arthritis, 7 years. • Reiff 2006: Rofecoxib vs. naproxen; juvenile idiopathic arthritis, 10 years. • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years. • Derosier 2012: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine) • Winner 2015: Sumatriptan and Naproxen (varying doses) vs. placebo; Non-chronic headache (Migraine), 15 years, 15 years.	

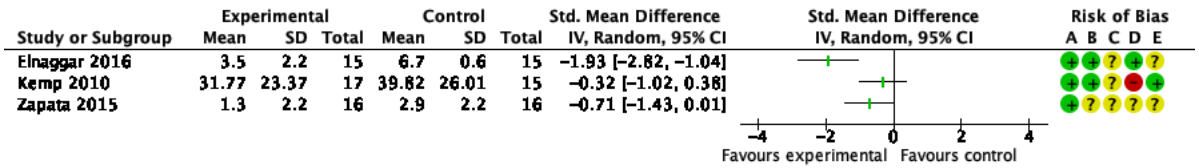
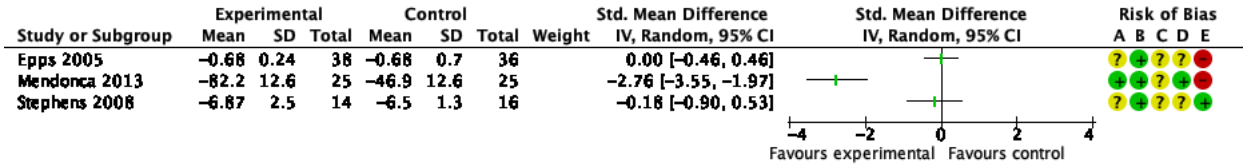
Other adverse events	Other adverse events  • Ruperto 2005: Meloxicam vs. naproxen, functional abdominal pain, 8 years.	NSAID vs NSAID: ⊕◯◯◯ VERY LOW

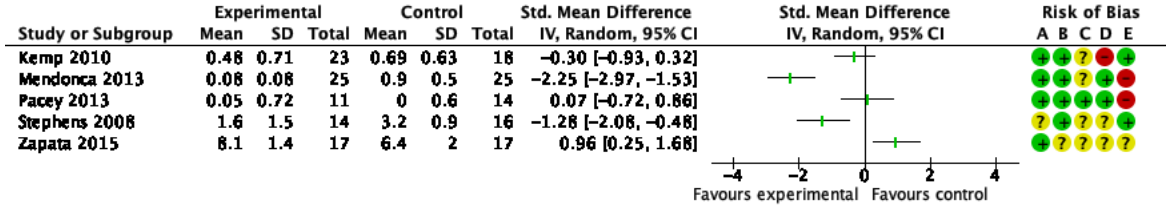
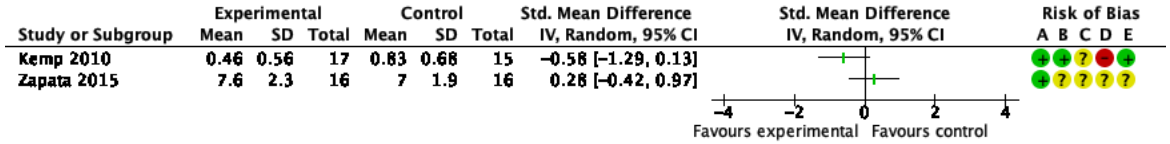
Appendix G.2. WHO review: Physical interventions for children with chronic pain

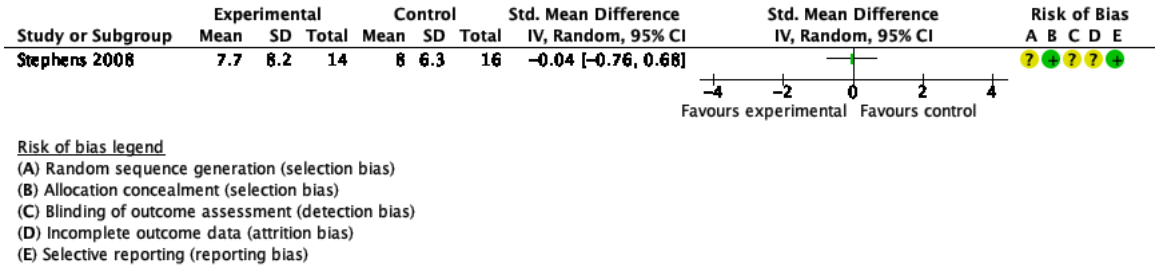
Comparison: Physical therapies versus other physical therapies
Population: Children with any chronic pain
Setting: Any setting
Studies: Randomised controlled trials

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Outcome	Forest plot
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)

Pain intensity, follow-up *Higher scores indicate higher pain intensity*	Pain intensity, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)
Health-related quality of life, post-treatment *Higher scores indicate better quality of life*	Health-related quality of life, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)

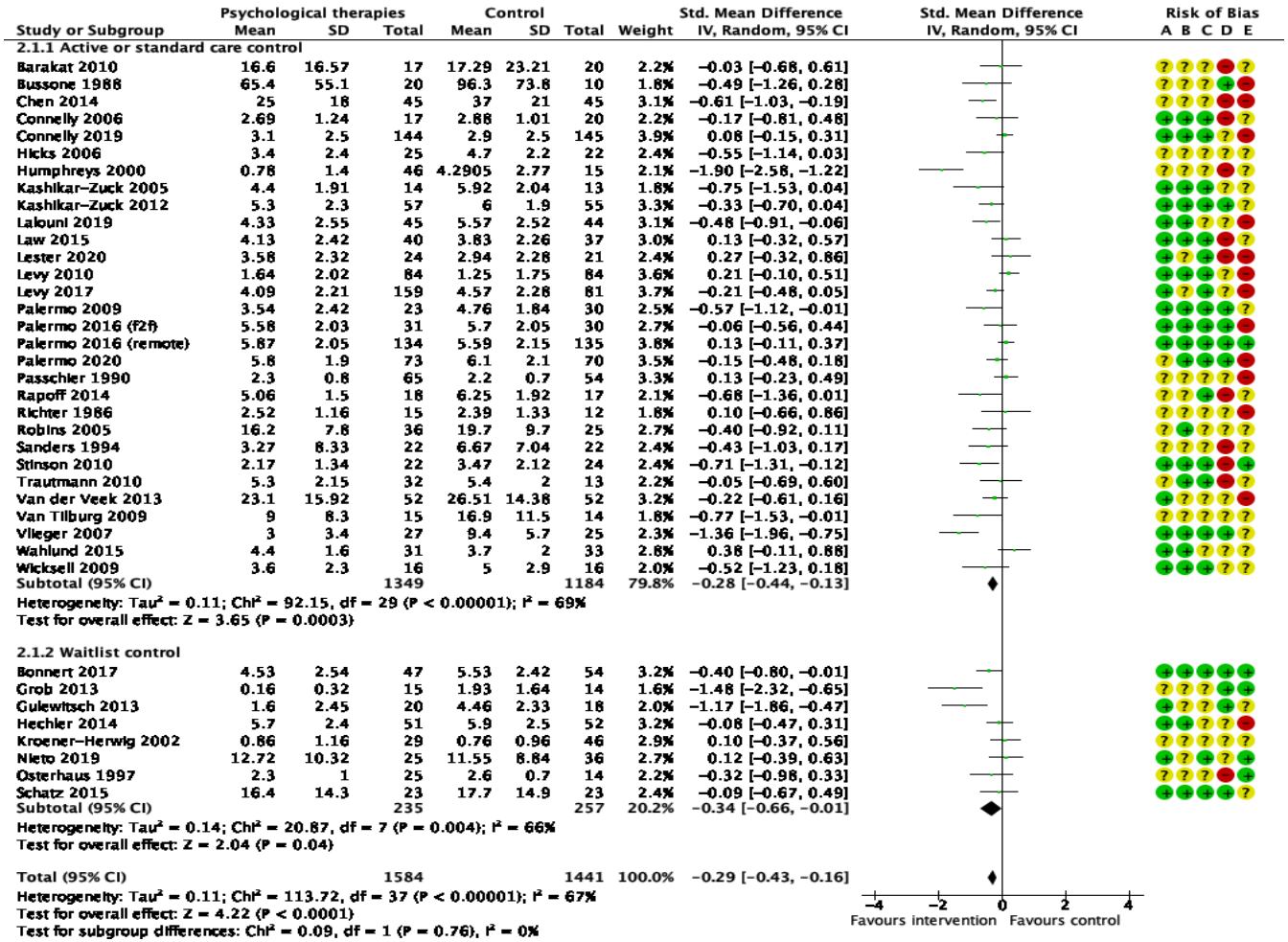
Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)
Functional disability, follow-up *Higher scores indicate lower disability*	Functional disability, follow-up  Risk of bias legend (A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias) (C) Blinding of outcome assessment (detection bias) (D) Incomplete outcome data (attrition bias) (E) Selective reporting (reporting bias)

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, post-treatment 

Appendix G.3. WHO review: Psychological interventions for children with chronic pain
Subgroup analysis: by control type

Comparison: Psychological therapies versus active/standard care control or waitlist control
Population: children with any chronic pain
Setting: Any setting
Studies: Randomised controlled trials

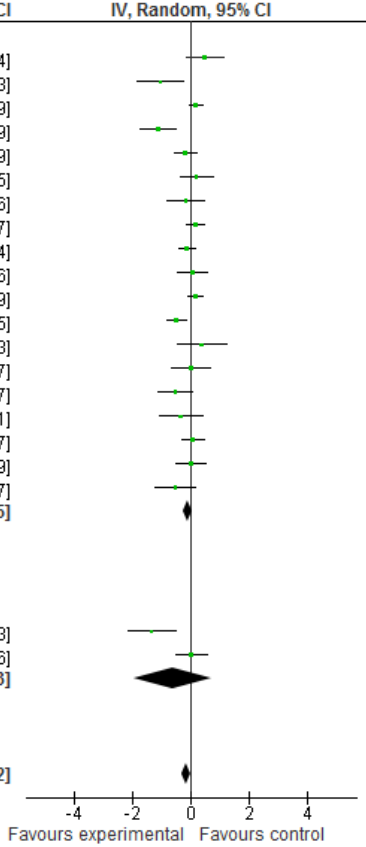
Risk of bias legend
(**A**) Random sequence generation (selection bias)
(**B**) Allocation concealment (selection bias)
(**C**) Blinding of outcome assessment (detection bias)
(**D**) Incomplete outcome data (attrition bias)
(**E**) Selective reporting (reporting bias)

Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*		Active or standard care control: ⊕⊕◯◯ LOW Waitlist control: ⊕⊕◯◯ LOW

Pain intensity, follow-up *Higher scores indicate higher pain intensity*

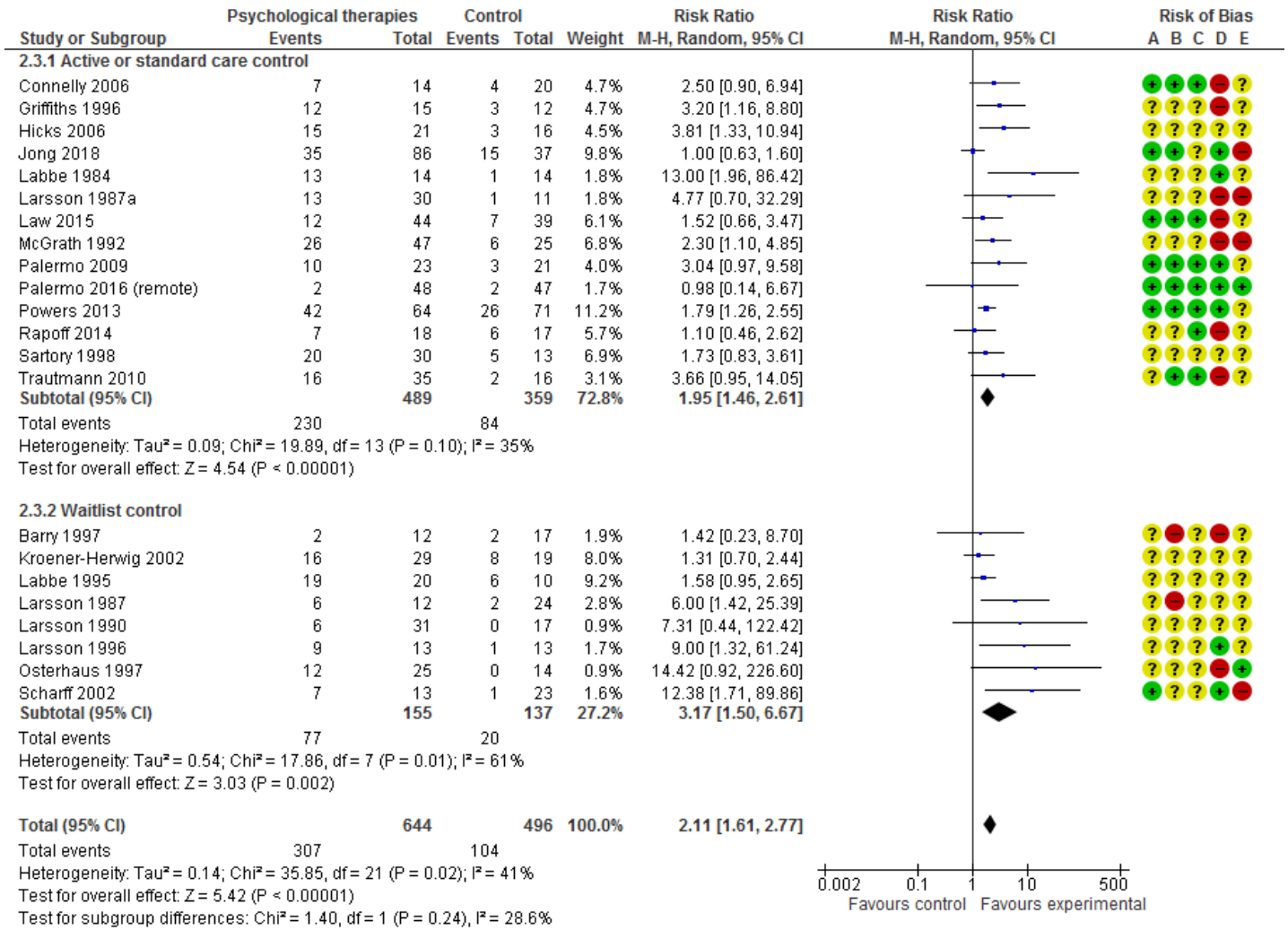
Pain intensity, follow-up

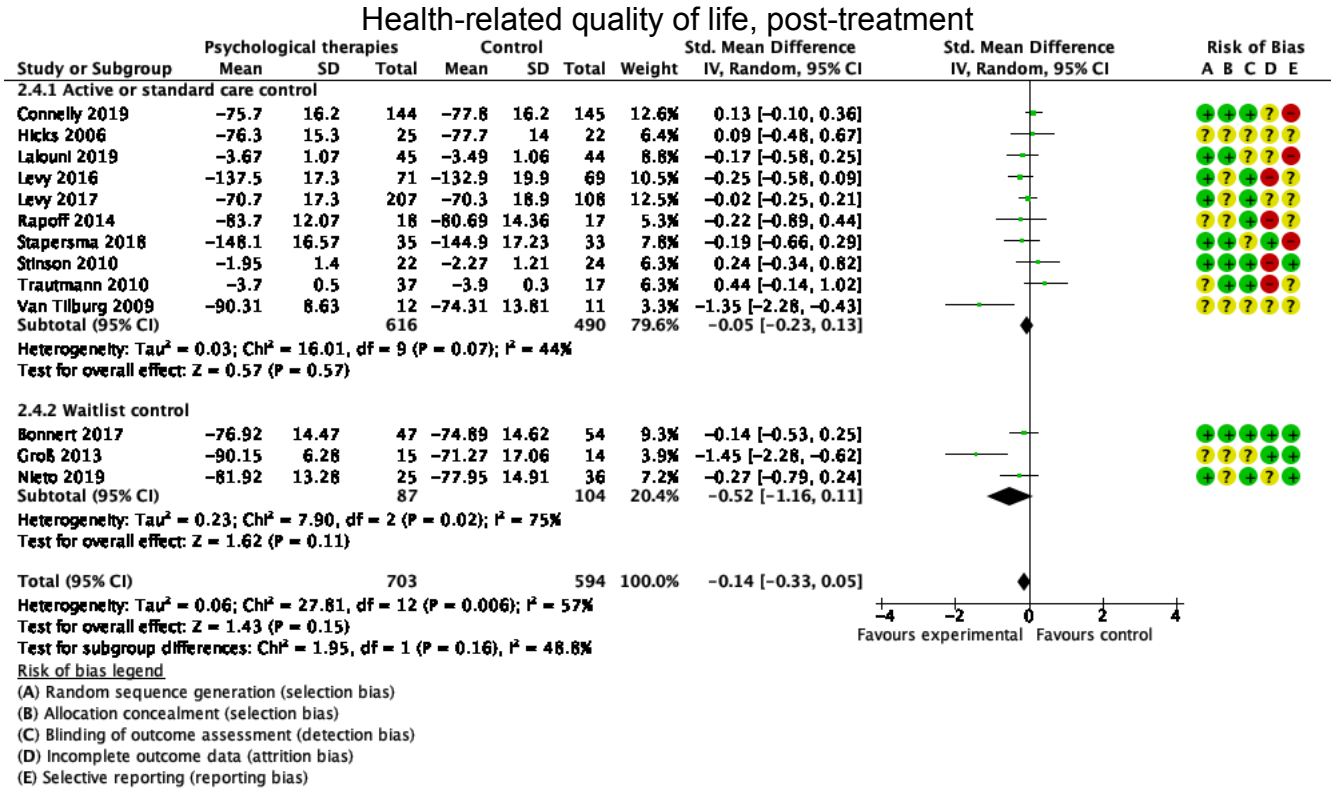
Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
2.2.1 Active or standard care control									
Barakat 2010	16.71	23.03	17	7.84	12.31	20	3.6%	0.48 [-0.18, 1.14]	? ? ? − ?
Bussone 1988	20	18.1	20	88.8	110.3	10	2.8%	-1.04 [-1.85, -0.23]	? ? ? + −
Connelly 2019	3.1	2.5	144	2.7	2.4	145	7.6%	0.16 [-0.07, 0.39]	+ + + ? −
Hicks 2006	2.9	2.1	25	4.9	1.3	22	3.9%	-1.11 [-1.73, -0.49]	? ? ? ? ?
Kashikar-Zuck 2012	4.9	2.2	57	5.3	2.1	55	6.1%	-0.18 [-0.56, 0.19]	+ + + + ?
Law 2015	4.19	2.45	28	3.7	2.54	22	4.4%	0.19 [-0.37, 0.75]	+ + + − ?
Lester 2020	2.67	1.9	21	3.07	2.64	18	3.8%	-0.17 [-0.80, 0.46]	+ ? + − −
Levy 2010	0.93	1.42	78	0.7	1.53	76	6.7%	0.16 [-0.16, 0.47]	+ + + ? −
Levy 2017	3.48	2.33	151	3.79	2.48	78	7.2%	-0.13 [-0.40, 0.14]	+ ? + ? −
Palermo 2016 (f2f)	5.42	2.05	31	5.3	2.12	30	4.8%	0.06 [-0.45, 0.56]	+ + + + −
Palermo 2016 (remote)	5.85	1.97	134	5.55	2.02	135	7.5%	0.15 [-0.09, 0.39]	+ + + + +
Palermo 2020	5.3	1.9	73	6.2	1.8	70	6.5%	-0.48 [-0.82, -0.15]	? + + + −
Rapoff 2014	4.46	1.88	11	3.68	2.04	11	2.6%	0.38 [-0.46, 1.23]	? ? + − ?
Richter 1986	2.02	1.48	30	2.02	1.39	12	3.6%	0.00 [-0.67, 0.67]	? ? ? ? −
Sanders 1994	0.64	1.38	22	2.11	3.56	22	4.0%	-0.53 [-1.14, 0.07]	? ? ? − ?
Trautmann 2010	4.9	1.4	12	5.5	1.9	16	3.1%	-0.34 [-1.10, 0.41]	? + + − ?
Van der Veek 2013	19.03	17.0393	52	17.72	15.19	52	6.0%	0.08 [-0.30, 0.47]	+ ? ? ? −
Wahlund 2015	2.8	1.9	31	2.8	1.6	33	5.0%	0.00 [-0.49, 0.49]	+ + ? ? ?
Wicksell 2009	3.1	2.7	16	4.5	2.4	16	3.3%	-0.53 [-1.24, 0.17]	+ + + ? ?
Subtotal (95% CI)			**953**			**843**	**92.6%**	**-0.11 [-0.26, 0.05]**	
Heterogeneity: Tau² = 0.06; Chi² = 41.84, df = 18 (P = 0.001); I² = 57%									
Test for overall effect: Z = 1.32 (P = 0.19)									
2.2.2 Wait-list control									
Grob 2013	0.08	0.31	15	1.55	1.49	14	2.7%	-1.35 [-2.17, -0.53]	? ? ? + +
Kroener-Herwig 2002	0.48	0.65	29	0.46	0.61	27	4.6%	0.03 [-0.49, 0.56]	? ? ? ? ?
Subtotal (95% CI)			**44**			**41**	**7.4%**	**-0.62 [-1.97, 0.73]**	
Heterogeneity: Tau² = 0.83; Chi² = 7.76, df = 1 (P = 0.005); I² = 87%									
Test for overall effect: Z = 0.90 (P = 0.37)									
Total (95% CI)			**997**			**884**	**100.0%**	**-0.14 [-0.30, 0.02]**	
Heterogeneity: Tau² = 0.08; Chi² = 51.59, df = 20 (P = 0.0001); I² = 61%									
Test for overall effect: Z = 1.67 (P = 0.09)									
Test for subgroup differences: Chi² = 0.55, df = 1 (P = 0.46), I² = 0%									



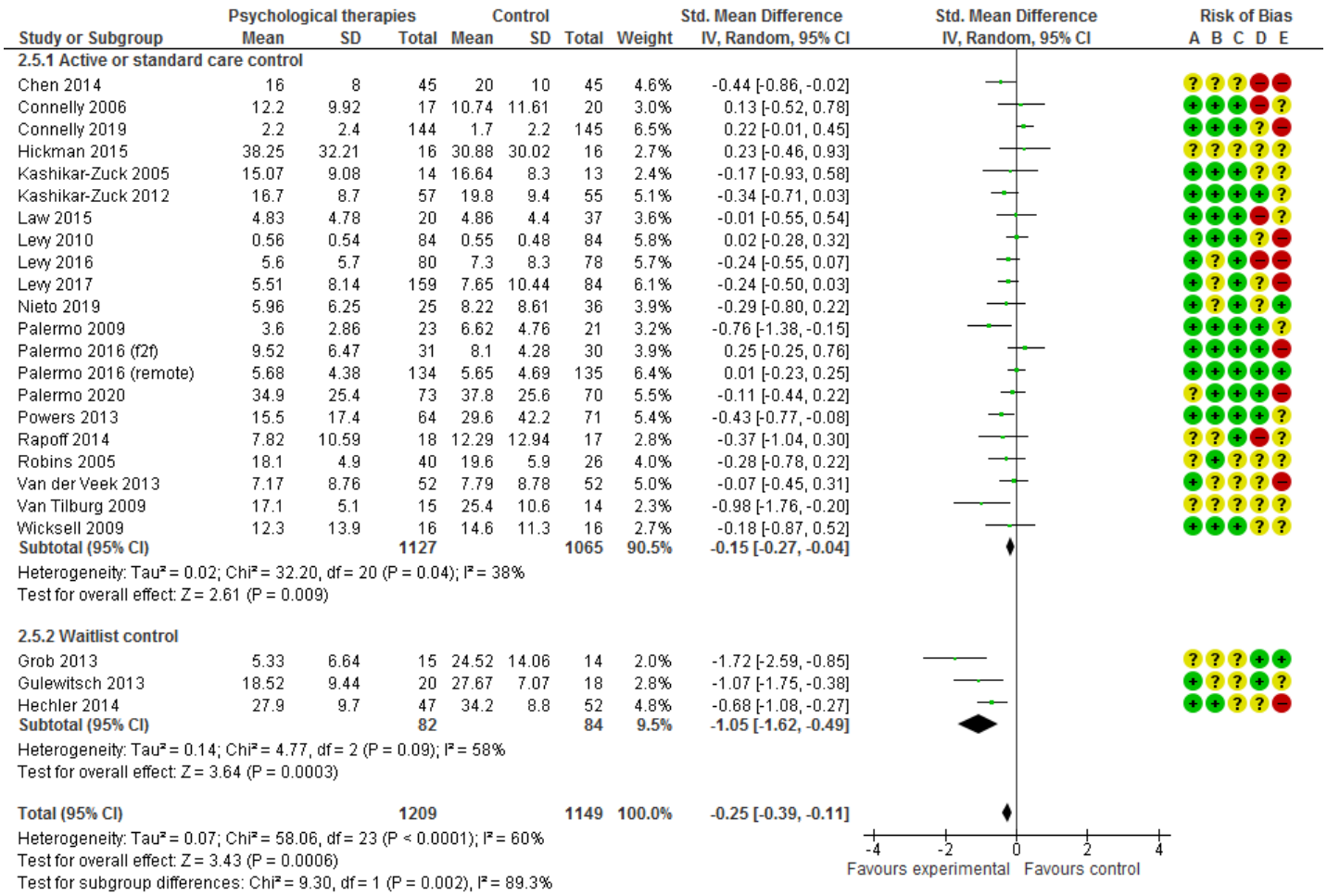
Active or standard care control: ⊕⊕◯◯ LOW

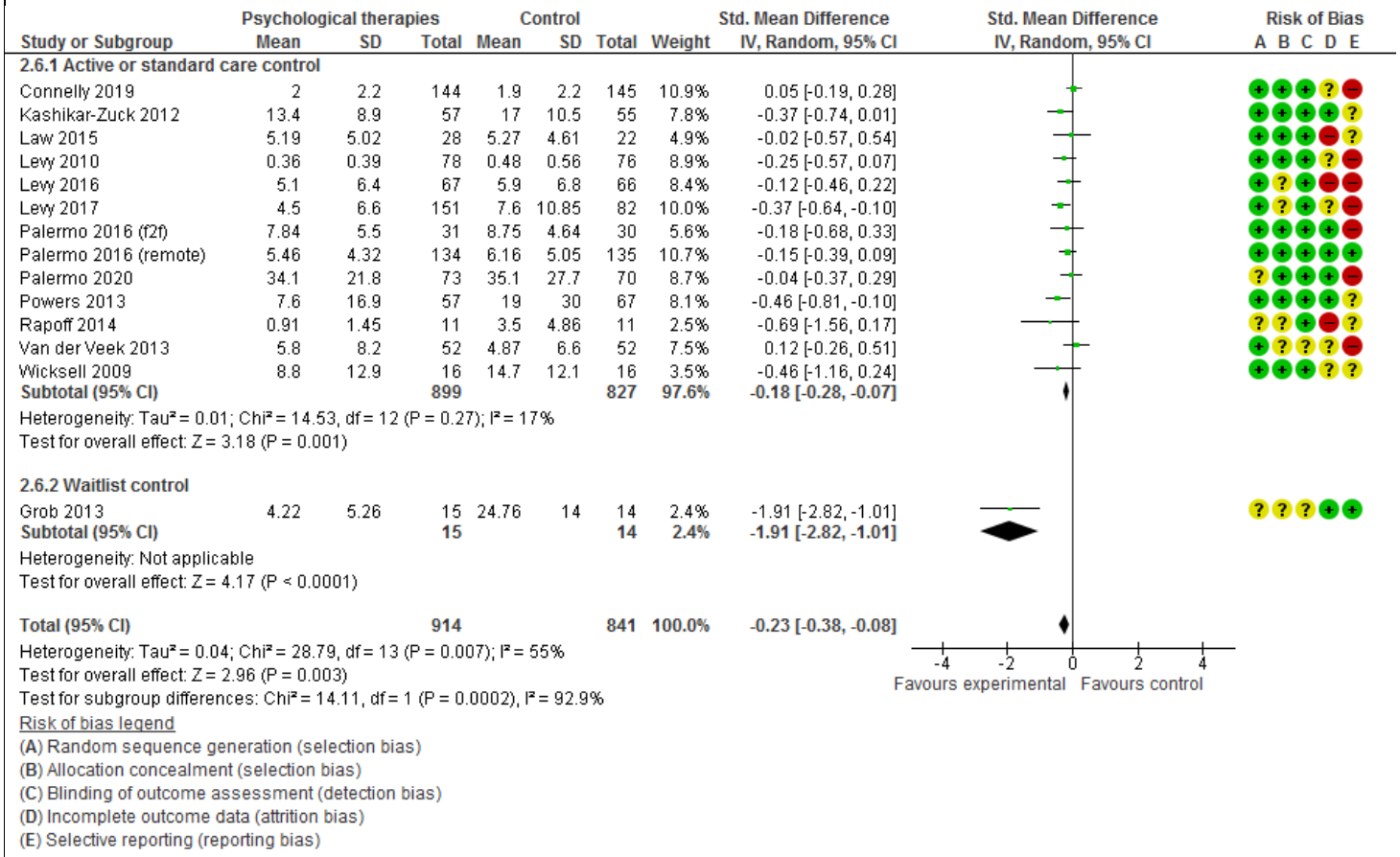
Waitlist control: ⊕◯◯◯ VERY LOW

50% reduction, post-treatment	50% reduction, post-treatment 	Active or standard care control: ⊕⊕◯◯ LOW Waitlist control: ⊕◯◯◯ VERY LOW

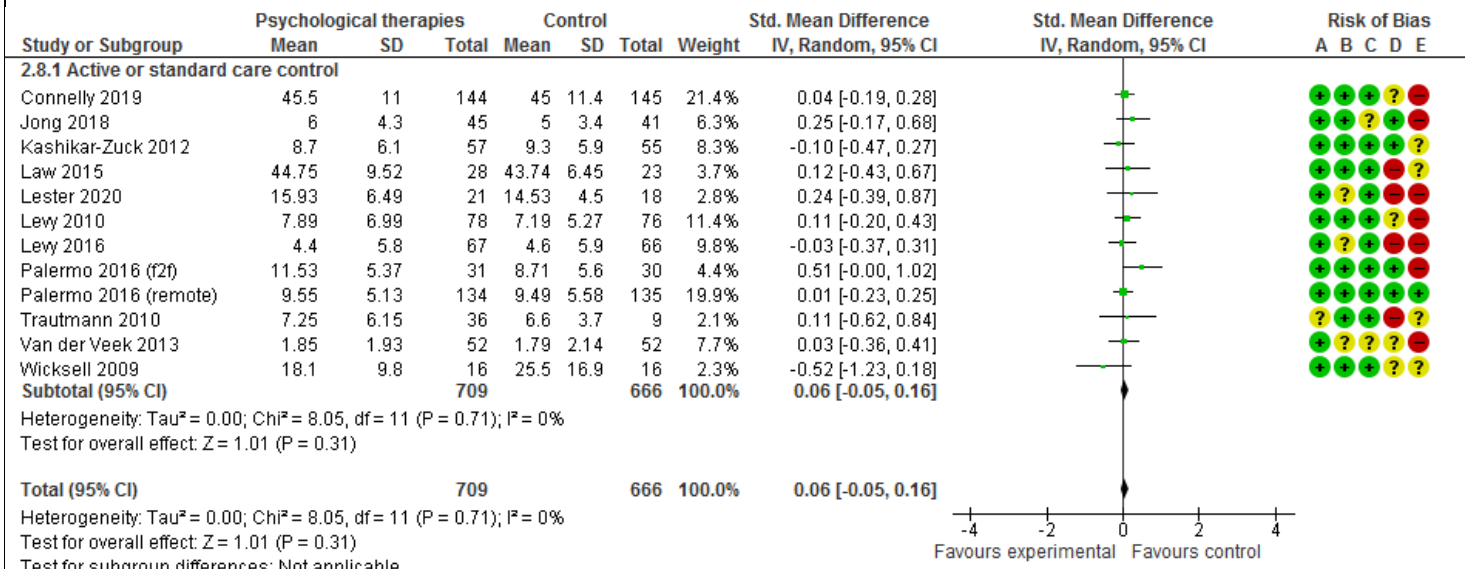
Health-related quality of life, post-treatment *Lower scores indicate better quality of life*		Active or standard care control: ⊕⊕⊕◯ MODERATE Waitlist control: ⊕◯◯◯ VERY LOW

Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment	Active or standard care control: ⊕⊕⊕◯ MODERATE Waitlist control: ⊕◯◯◯ VERY LOW

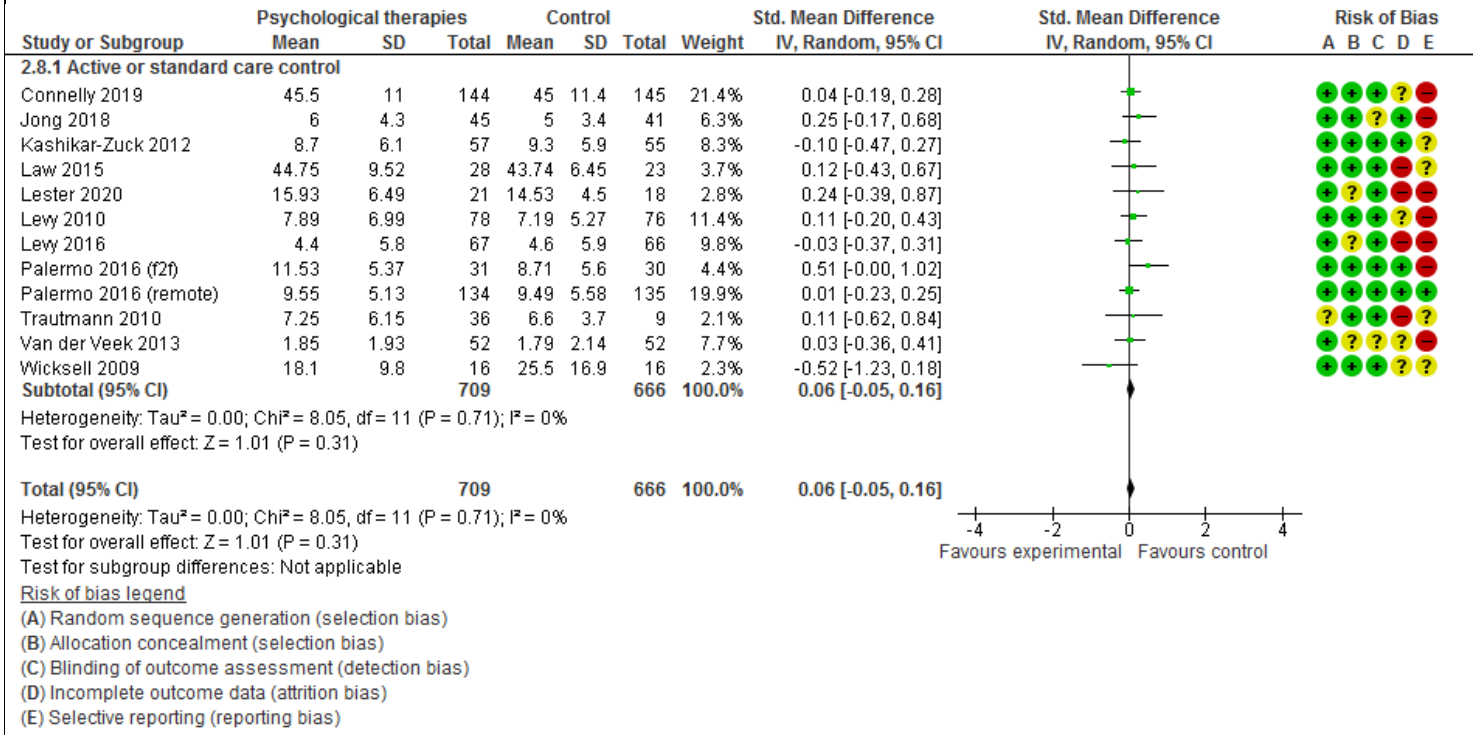


Functional disability, follow-up *Higher scores indicate lower disability*	Functional disability, follow-up 	Active or standard care control: ⊕⊕⊕⊕ HIGH

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, post-treatment	Active or standard care control: ⊕⊕⊕⊕ HIGH



Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow up	Active or standard care control: ⊕⊕⊕⊕ HIGH



Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.8.1 Active or standard care control								
Connelly 2019	45.5	11	144	45	11.4	145	21.4%	0.04 [-0.19, 0.28]
Jong 2018	6	4.3	45	5	3.4	41	6.3%	0.25 [-0.17, 0.68]
Kashikar-Zuck 2012	8.7	6.1	57	9.3	5.9	55	8.3%	-0.10 [-0.47, 0.27]
Law 2015	44.75	9.52	28	43.74	6.45	23	3.7%	0.12 [-0.43, 0.67]
Lester 2020	15.93	6.49	21	14.53	4.5	18	2.8%	0.24 [-0.39, 0.87]
Lewy 2010	7.89	6.99	78	7.19	5.27	76	11.4%	0.11 [-0.20, 0.43]
Lewy 2016	4.4	5.8	67	4.6	5.9	66	9.8%	-0.03 [-0.37, 0.31]
Palermo 2016 (f2f)	11.53	5.37	31	8.71	5.6	30	4.4%	0.51 [-0.00, 1.02]
Palermo 2016 (remote)	9.55	5.13	134	9.49	5.58	135	19.9%	0.01 [-0.23, 0.25]
Trautmann 2010	7.25	6.15	36	6.6	3.7	9	2.1%	0.11 [-0.62, 0.84]
Van der Veek 2013	1.85	1.93	52	1.79	2.14	52	7.7%	0.03 [-0.36, 0.41]
Wicksell 2009	18.1	9.8	16	25.5	16.9	16	2.3%	-0.52 [-1.23, 0.18]
Subtotal (95% CI)			**709**			**666**	**100.0%**	**0.06 [-0.05, 0.16]**
Heterogeneity: Tau² = 0.00; Chi² = 8.05, df = 11 (P = 0.71); I² = 0%								
Test for overall effect: Z = 1.01 (P = 0.31)								
Total (95% CI)			**709**			**666**	**100.0%**	**0.06 [-0.05, 0.16]**
Heterogeneity: Tau² = 0.00; Chi² = 8.05, df = 11 (P = 0.71); I² = 0%								
Test for overall effect: Z = 1.01 (P = 0.31)								
Test for subgroup differences: Not applicable								

Std. Mean Difference IV, Random, 95% CI

-4 -2 0 2 4

Favours experimental Favours control

Risk of Bias A B C D E

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of outcome assessment (detection bias)

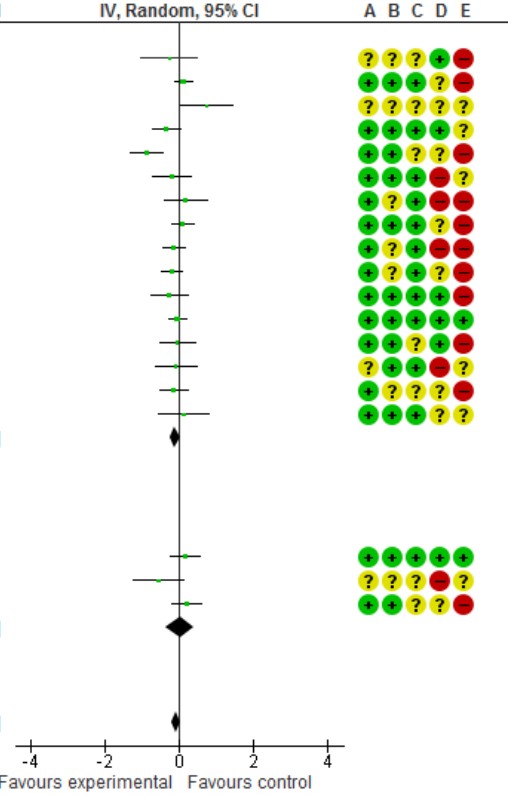
(D) Incomplete outcome data (attrition bias)

(E) Selective reporting (reporting bias)

Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher depressive symptomology*

Emotional functioning: Anxiety, post-treatment

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
2.9.1 Active or standard care control									
Bussone 1988	28.1	3.49	20	29.2	5.1	10	2.3%	-0.26 [-1.02, 0.50]	? ? ? + −
Connelly 2019	46.8	11.3	144	45.5	11	145	9.2%	0.12 [-0.11, 0.35]	+ + + ? −
Hickman 2015	52.56	7.36	16	47.38	6.1	17	2.6%	0.75 [0.04, 1.46]	? ? ? ? ?
Kashikar-Zuck 2012	2.11	0.72	50	2.39	0.9	50	5.8%	-0.34 [-0.74, 0.05]	+ + + + ?
Lalouni 2019	8.59	7.71	45	15.31	7.63	44	5.2%	-0.87 [-1.30, -0.43]	+ + ? ? −
Law 2015	46.33	8.99	30	48.32	10.81	25	4.0%	-0.20 [-0.73, 0.33]	+ + + − ?
Lester 2020	7.08	6.24	24	6.1	4.96	21	3.4%	0.17 [-0.42, 0.76]	+ ? + − −
Lewy 2010	13.5	4.86	83	13.04	4.04	80	7.5%	0.10 [-0.21, 0.41]	+ + + ? −
Lewy 2016	8.2	2.8	80	8.6	2.9	78	7.4%	-0.14 [-0.45, 0.17]	+ ? + − −
Lewy 2017	1.09	0.94	159	1.28	1.07	81	8.3%	-0.19 [-0.46, 0.08]	+ ? + ? −
Palermo 2016 (f2f)	11.42	5.33	31	13	6.03	30	4.3%	-0.27 [-0.78, 0.23]	+ + + + −
Palermo 2016 (remote)	10.56	5.91	134	10.85	6.1	135	9.0%	-0.05 [-0.29, 0.19]	+ + + + +
Stapersma 2018	7.1	4.14	35	7.3	4.6	33	4.6%	-0.05 [-0.52, 0.43]	+ + ? + −
Trautmann 2010	30.9	7.95	38	31.7	8.3	18	3.7%	-0.10 [-0.66, 0.46]	? + + − ?
Van der Veek 2013	6.83	6	52	7.76	6.33	52	6.0%	-0.15 [-0.53, 0.24]	+ ? ? ? −
Wicksell 2009	13.4	3.9	16	12.8	5.5	16	2.6%	0.12 [-0.57, 0.82]	+ + + ? ?
Subtotal (95% CI)			**957**			**835**	**85.7%**	**-0.10 [-0.24, 0.03]**	
Heterogeneity: Tau² = 0.03; Chi² = 26.66, df = 15 (P = 0.03); I² = 44%									
Test for overall effect: Z = 1.50 (P = 0.13)									
2.9.2 Waitlist control									
Bonnert 2017	25.23	16.32	47	22.62	16.31	54	5.8%	0.16 [-0.23, 0.55]	+ + + + +
Griffiths 1996	9.6	5.9	30	13.6	9.5	12	2.7%	-0.55 [-1.24, 0.13]	? ? ? − ?
Hechler 2014	52.5	12.1	50	50	11.4	46	5.7%	0.21 [-0.19, 0.61]	+ + ? ? −
Subtotal (95% CI)			**127**			**112**	**14.3%**	**0.03 [-0.35, 0.41]**	
Heterogeneity: Tau² = 0.05; Chi² = 3.90, df = 2 (P = 0.14); I² = 49%									
Test for overall effect: Z = 0.14 (P = 0.89)									
Total (95% CI)			**1084**			**947**	**100.0%**	**-0.08 [-0.21, 0.04]**	
Heterogeneity: Tau² = 0.03; Chi² = 31.94, df = 18 (P = 0.02); I² = 44%									
Test for overall effect: Z = 1.28 (P = 0.20)									
Test for subgroup differences: Chi² = 0.40, df = 1 (P = 0.53), I² = 0%									



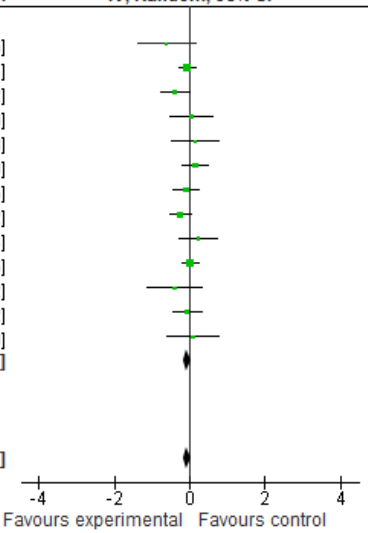
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Active or standard care control: ⊕⊕⊕◯ MODERATE

Waitlist control: ⊕◯◯◯ VERY LOW

Emotional functioning: Anxiety, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Anxiety, follow-up	Active or standard care control: ⊕⊕⊕◯ MODERATE

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.10.1 Active or standard care control								
Bussone 1988	27.8	2.3	20	29.1	1.4	10	1.7%	-0.62 [-1.39, 0.16]
Connelly 2019	45.3	12	144	46	11.4	145	19.7%	-0.06 [-0.29, 0.17]
Kashikar-Zuck 2012	1.89	0.82	50	2.22	0.91	50	6.7%	-0.38 [-0.77, 0.02]
Law 2015	45.82	10.96	28	45.36	9.9	22	3.4%	0.04 [-0.52, 0.60]
Lester 2020	4.71	5.09	21	4.07	2.99	18	2.6%	0.15 [-0.48, 0.78]
Lewy 2010	13.21	3.98	75	12.59	4.14	63	9.3%	0.15 [-0.18, 0.49]
Lewy 2016	7.9	3.3	67	8.2	3.2	66	9.0%	-0.09 [-0.43, 0.25]
Lewy 2017	0.87	0.88	151	1.1	0.98	78	13.9%	-0.25 [-0.52, 0.02]
Palermo 2016 (f2f)	12.61	6.05	31	11.21	5.55	30	4.1%	0.24 [-0.27, 0.74]
Palermo 2016 (remote)	10.35	6.12	134	10.23	5.45	135	18.3%	0.02 [-0.22, 0.26]
Trautmann 2010	24.95	7	31	28.1	9.9	10	2.0%	-0.40 [-1.12, 0.32]
Van der Veek 2013	5.47	5.22	52	5.82	6.09	52	7.1%	-0.06 [-0.45, 0.32]
Wicksell 2009	12.2	4.6	16	11.7	5.8	16	2.2%	0.09 [-0.60, 0.79]
Subtotal (95% CI)			**820**			**695**	**100.0%**	**-0.07 [-0.17, 0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 11.21, df = 12 (P = 0.51); I² = 0%								
Test for overall effect: Z = 1.31 (P = 0.19)								
Total (95% CI)			**820**			**695**	**100.0%**	**-0.07 [-0.17, 0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 11.21, df = 12 (P = 0.51); I² = 0%								
Test for overall effect: Z = 1.31 (P = 0.19)								
Test for subgroup differences: Not applicable								



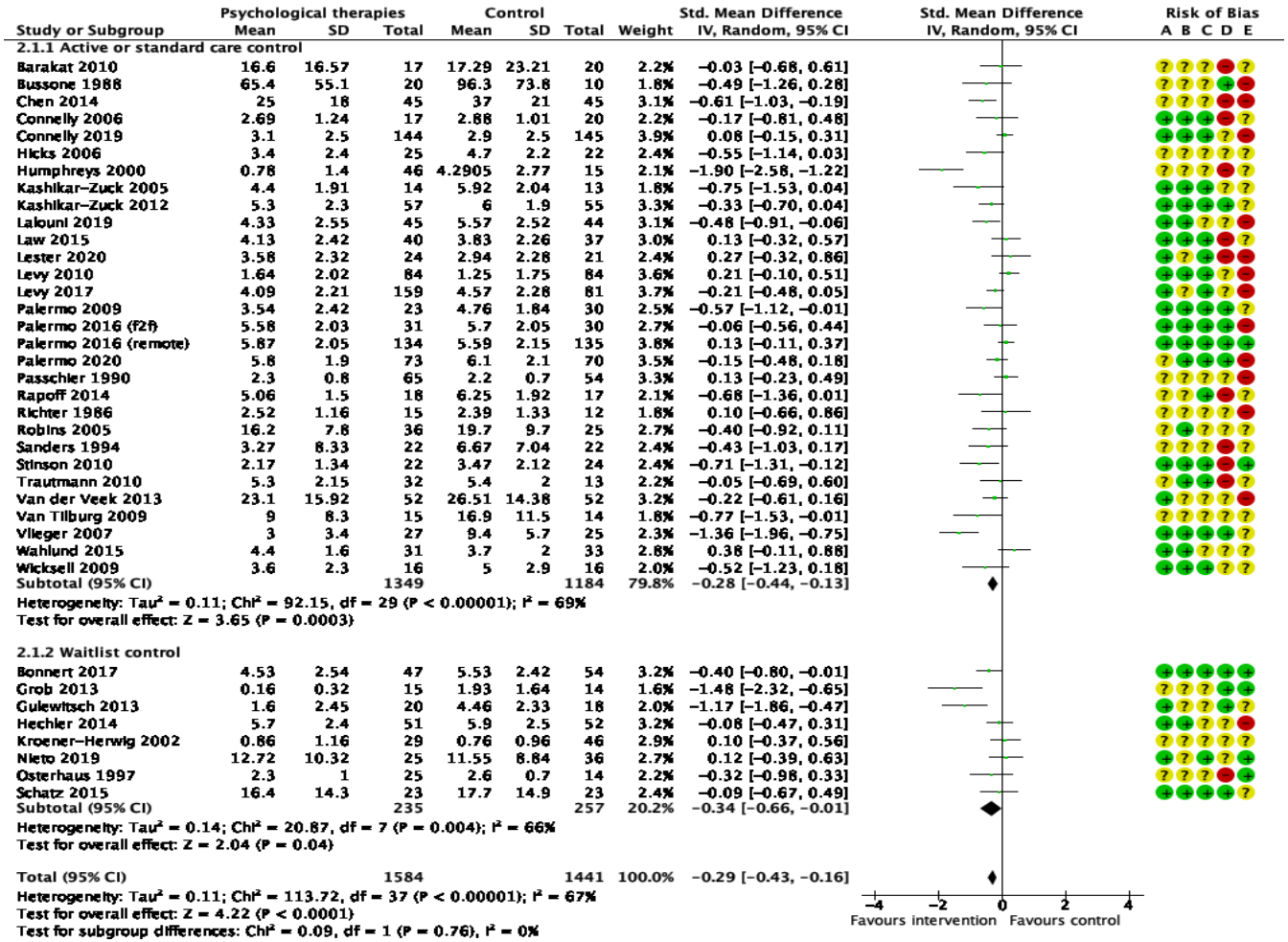
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Appendix G.4. WHO review: Psychological interventions for children with chronic pain
Subgroup analysis: by control type

Comparison: Psychological therapies versus active/standard care control or waitlist control
Population: children with any chronic pain
Setting: Any setting
Studies: Randomised controlled trials

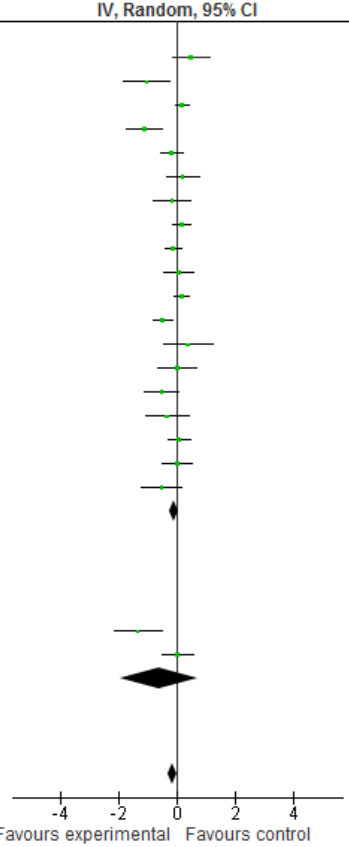
Risk of bias legend
(**A**) Random sequence generation (selection bias)
(**B**) Allocation concealment (selection bias)
(**C**) Blinding of outcome assessment (detection bias)
(**D**) Incomplete outcome data (attrition bias)
(**E**) Selective reporting (reporting bias)

Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*		Active or standard care control: ⊕⊕◯◯ LOW Waitlist control: ⊕⊕◯◯ LOW

Pain intensity, follow-up *Higher scores indicate higher pain intensity*

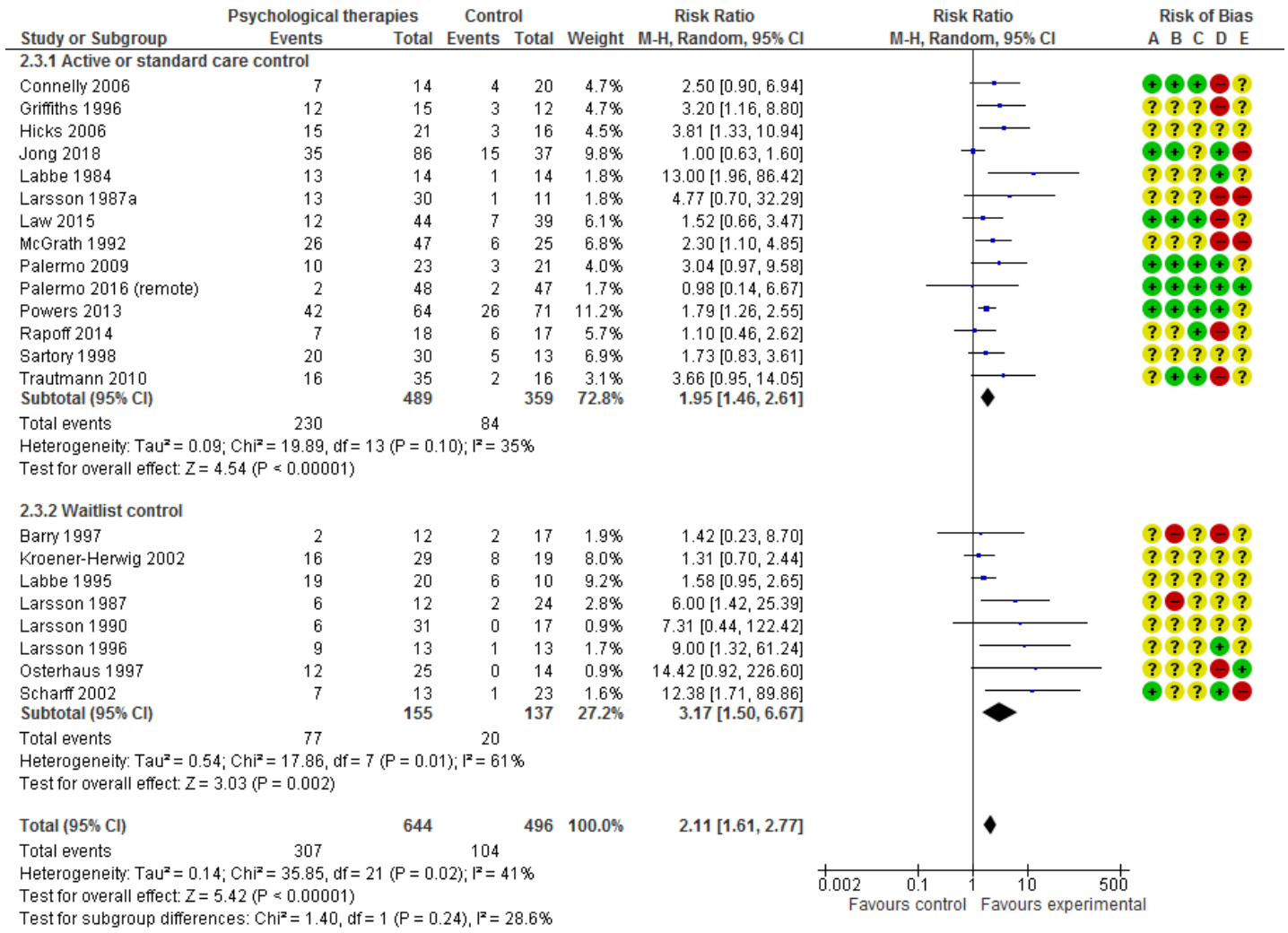
Pain intensity, follow-up

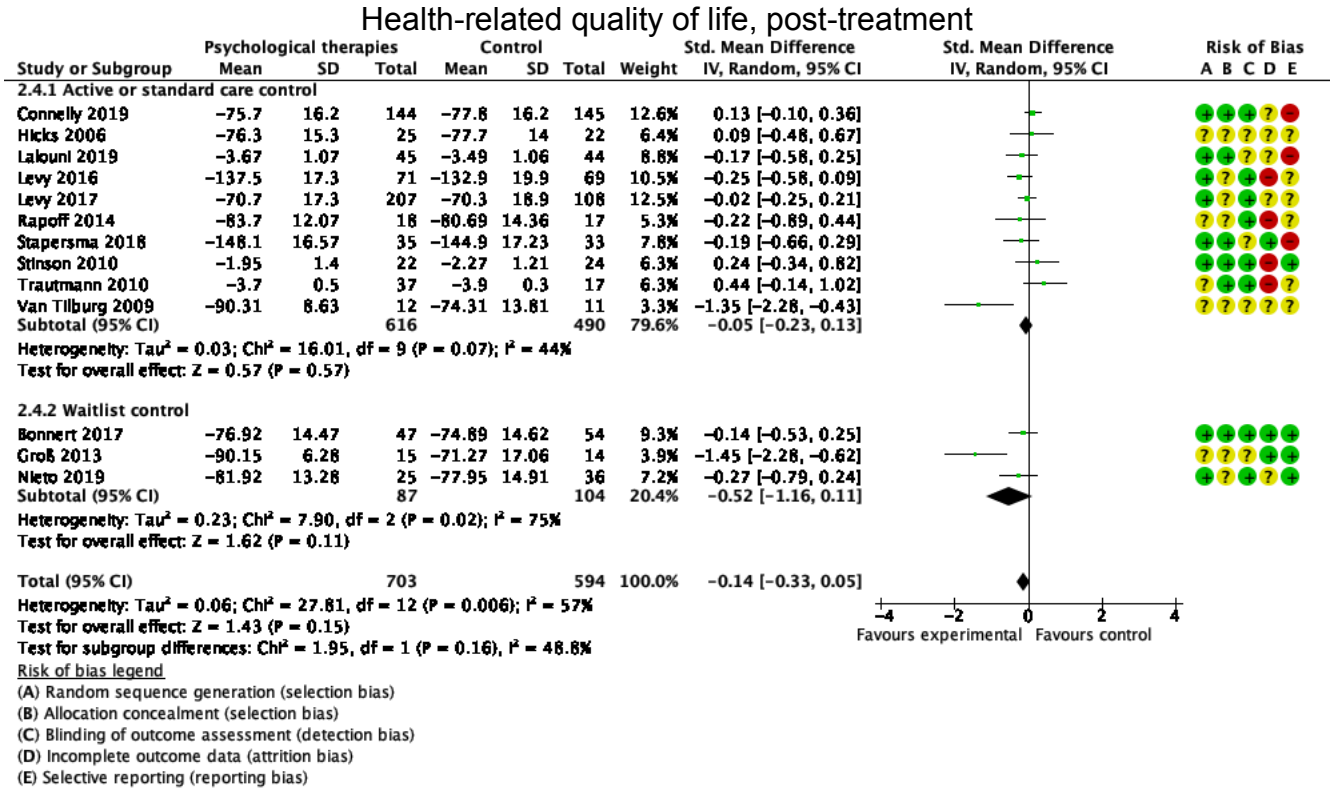
Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
2.2.1 Active or standard care control									
Barakat 2010	16.71	23.03	17	7.84	12.31	20	3.6%	0.48 [-0.18, 1.14]	? ? ? − ?
Bussone 1988	20	18.1	20	88.8	110.3	10	2.8%	-1.04 [-1.85, -0.23]	? ? ? + −
Connelly 2019	3.1	2.5	144	2.7	2.4	145	7.6%	0.16 [-0.07, 0.39]	+ + + ? −
Hicks 2006	2.9	2.1	25	4.9	1.3	22	3.9%	-1.11 [-1.73, -0.49]	? ? ? ? ?
Kashikar-Zuck 2012	4.9	2.2	57	5.3	2.1	55	6.1%	-0.18 [-0.56, 0.19]	+ + + + ?
Law 2015	4.19	2.45	28	3.7	2.54	22	4.4%	0.19 [-0.37, 0.75]	+ + + − ?
Lester 2020	2.67	1.9	21	3.07	2.64	18	3.8%	-0.17 [-0.80, 0.46]	+ ? + − −
Lewy 2010	0.93	1.42	78	0.7	1.53	76	6.7%	0.16 [-0.16, 0.47]	+ + + ? −
Lewy 2017	3.48	2.33	151	3.79	2.48	78	7.2%	-0.13 [-0.40, 0.14]	+ ? + ? −
Palermo 2016 (f2f)	5.42	2.05	31	5.3	2.12	30	4.8%	0.06 [-0.45, 0.56]	+ + + + −
Palermo 2016 (remote)	5.85	1.97	134	5.55	2.02	135	7.5%	0.15 [-0.09, 0.39]	+ + + + +
Palermo 2020	5.3	1.9	73	6.2	1.8	70	6.5%	-0.48 [-0.82, -0.15]	? + + + −
Rapoff 2014	4.46	1.88	11	3.68	2.04	11	2.6%	0.38 [-0.46, 1.23]	? ? + − ?
Richter 1986	2.02	1.48	30	2.02	1.39	12	3.6%	0.00 [-0.67, 0.67]	? ? ? ? −
Sanders 1994	0.64	1.38	22	2.11	3.56	22	4.0%	-0.53 [-1.14, 0.07]	? ? ? − ?
Trautmann 2010	4.9	1.4	12	5.5	1.9	16	3.1%	-0.34 [-1.10, 0.41]	? + + − ?
Van der Veek 2013	19.03	17.0393	52	17.72	15.19	52	6.0%	0.08 [-0.30, 0.47]	+ ? ? ? −
Wahlund 2015	2.8	1.9	31	2.8	1.6	33	5.0%	0.00 [-0.49, 0.49]	+ + ? ? ?
Wicksell 2009	3.1	2.7	16	4.5	2.4	16	3.3%	-0.53 [-1.24, 0.17]	+ + + ? ?
Subtotal (95% CI)			**953**			**843**	**92.6%**	**-0.11 [-0.26, 0.05]**	
Heterogeneity: Tau² = 0.06; Chi² = 41.84, df = 18 (P = 0.001); I² = 57%									
Test for overall effect: Z = 1.32 (P = 0.19)									
2.2.2 Wait-list control									
Grob 2013	0.08	0.31	15	1.55	1.49	14	2.7%	-1.35 [-2.17, -0.53]	? ? ? + +
Kroener-Herwig 2002	0.48	0.65	29	0.46	0.61	27	4.6%	0.03 [-0.49, 0.56]	? ? ? ? ?
Subtotal (95% CI)			**44**			**41**	**7.4%**	**-0.62 [-1.97, 0.73]**	
Heterogeneity: Tau² = 0.83; Chi² = 7.76, df = 1 (P = 0.005); I² = 87%									
Test for overall effect: Z = 0.90 (P = 0.37)									
Total (95% CI)			**997**			**884**	**100.0%**	**-0.14 [-0.30, 0.02]**	
Heterogeneity: Tau² = 0.08; Chi² = 51.59, df = 20 (P = 0.0001); I² = 61%									
Test for overall effect: Z = 1.67 (P = 0.09)									
Test for subgroup differences: Chi² = 0.55, df = 1 (P = 0.46), I² = 0%									



Active or standard care control: ⊕⊕◯◯ LOW

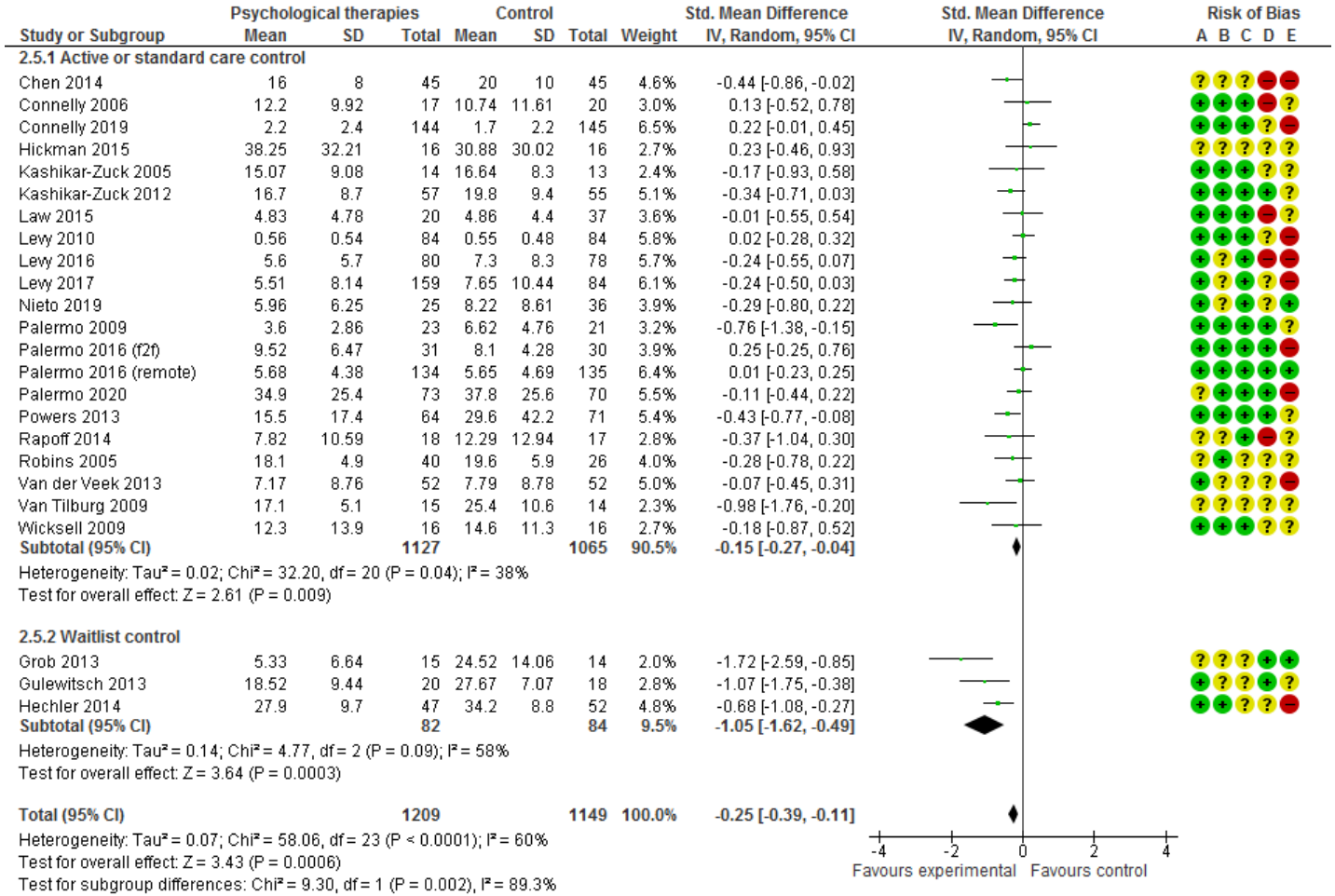
Waitlist control: ⊕◯◯◯ VERY LOW

50% reduction, post-treatment	50% reduction, post-treatment 	Active or standard care control: ⊕⊕◯◯ LOW Waitlist control: ⊕◯◯◯ VERY LOW

Health-related quality of life, post-treatment *Lower scores indicate better quality of life*		Active or standard care control: ⊕⊕⊕◯ MODERATE Waitlist control: ⊕◯◯◯ VERY LOW

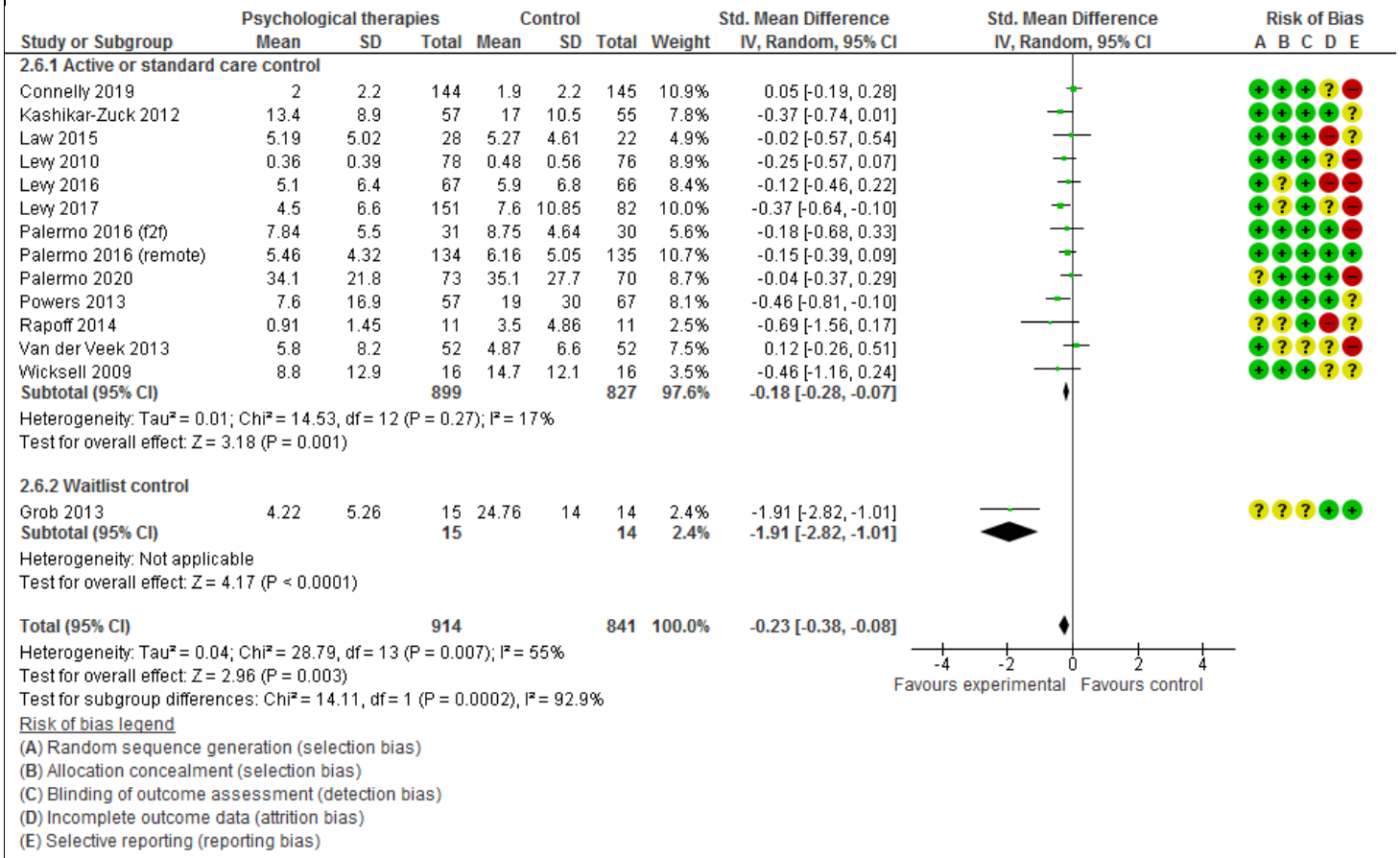
Functional disability, post-treatment *Higher scores indicate lower disability*

Functional disability, post-treatment



Active or standard care control: ⊕⊕⊕◯ MODERATE

Waitlist control: ⊕◯◯◯ VERY LOW

Functional disability, follow-up *Higher scores indicate lower disability*	Functional disability, follow-up 	Active or standard care control: ⊕⊕⊕⊕ HIGH

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*

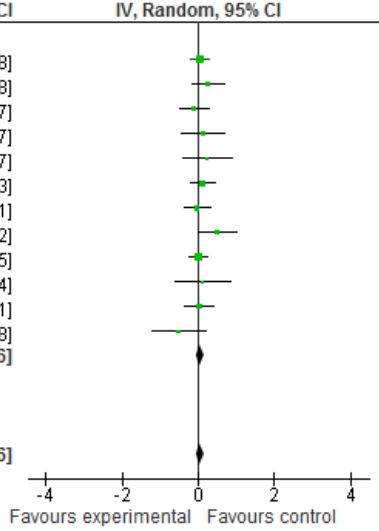
Emotional functioning: Depression, post-treatment

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.8.1 Active or standard care control								
Connelly 2019	45.5	11	144	45	11.4	145	21.4%	0.04 [-0.19, 0.28]
Jong 2018	6	4.3	45	5	3.4	41	6.3%	0.25 [-0.17, 0.68]
Kashikar-Zuck 2012	8.7	6.1	57	9.3	5.9	55	8.3%	-0.10 [-0.47, 0.27]
Law 2015	44.75	9.52	28	43.74	6.45	23	3.7%	0.12 [-0.43, 0.67]
Lester 2020	15.93	6.49	21	14.53	4.5	18	2.8%	0.24 [-0.39, 0.87]
Lewy 2010	7.89	6.99	78	7.19	5.27	76	11.4%	0.11 [-0.20, 0.43]
Lewy 2016	4.4	5.8	67	4.6	5.9	66	9.8%	-0.03 [-0.37, 0.31]
Palermo 2016 (f2f)	11.53	5.37	31	8.71	5.6	30	4.4%	0.51 [-0.00, 1.02]
Palermo 2016 (remote)	9.55	5.13	134	9.49	5.58	135	19.9%	0.01 [-0.23, 0.25]
Trautmann 2010	7.25	6.15	36	6.6	3.7	9	2.1%	0.11 [-0.62, 0.84]
Van der Veek 2013	1.85	1.93	52	1.79	2.14	52	7.7%	0.03 [-0.36, 0.41]
Wicksell 2009	18.1	9.8	16	25.5	16.9	16	2.3%	-0.52 [-1.23, 0.18]
Subtotal (95% CI)			**709**			**666**	**100.0%**	**0.06 [-0.05, 0.16]**

Heterogeneity: Tau² = 0.00; Chi² = 8.05, df = 11 (P = 0.71); I² = 0%
Test for overall effect: Z = 1.01 (P = 0.31)

Total (95% CI) **709** **666** **100.0%** **0.06 [-0.05, 0.16]**

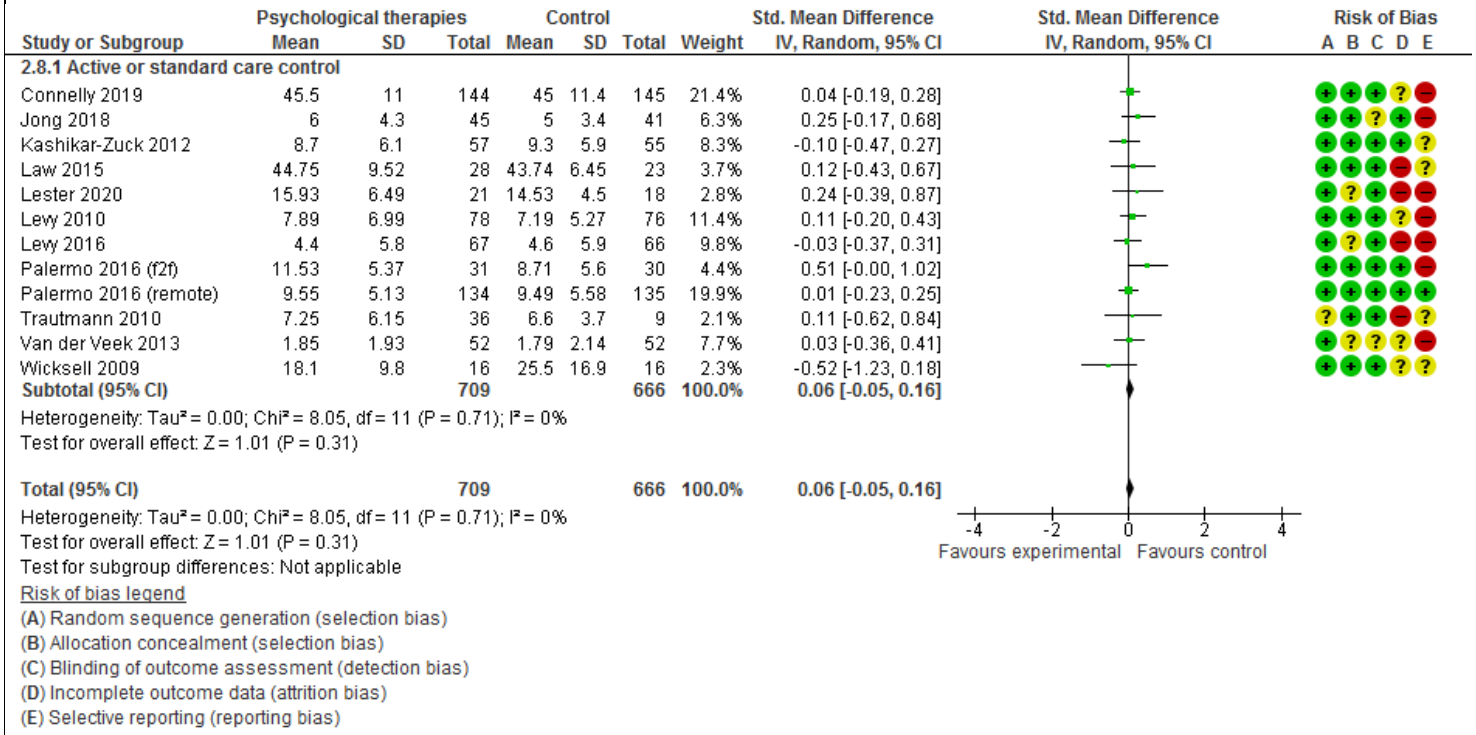
Heterogeneity: Tau² = 0.00; Chi² = 8.05, df = 11 (P = 0.71); I² = 0%
Test for overall effect: Z = 1.01 (P = 0.31)
Test for subgroup differences: Not applicable



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Active or standard care control: ⊕⊕⊕⊕ HIGH

Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow up	Active or standard care control: ⊕⊕⊕⊕ HIGH



Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.8.1 Active or standard care control								
Connelly 2019	45.5	11	144	45	11.4	145	21.4%	0.04 [-0.19, 0.28]
Jong 2018	6	4.3	45	5	3.4	41	6.3%	0.25 [-0.17, 0.68]
Kashikar-Zuck 2012	8.7	6.1	57	9.3	5.9	55	8.3%	-0.10 [-0.47, 0.27]
Law 2015	44.75	9.52	28	43.74	6.45	23	3.7%	0.12 [-0.43, 0.67]
Lester 2020	15.93	6.49	21	14.53	4.5	18	2.8%	0.24 [-0.39, 0.87]
Lewy 2010	7.89	6.99	78	7.19	5.27	76	11.4%	0.11 [-0.20, 0.43]
Lewy 2016	4.4	5.8	67	4.6	5.9	66	9.8%	-0.03 [-0.37, 0.31]
Palermo 2016 (f2f)	11.53	5.37	31	8.71	5.6	30	4.4%	0.51 [-0.00, 1.02]
Palermo 2016 (remote)	9.55	5.13	134	9.49	5.58	135	19.9%	0.01 [-0.23, 0.25]
Trautmann 2010	7.25	6.15	36	6.6	3.7	9	2.1%	0.11 [-0.62, 0.84]
Van der Veek 2013	1.85	1.93	52	1.79	2.14	52	7.7%	0.03 [-0.36, 0.41]
Wicksell 2009	18.1	9.8	16	25.5	16.9	16	2.3%	-0.52 [-1.23, 0.18]
Subtotal (95% CI)			**709**			**666**	**100.0%**	**0.06 [-0.05, 0.16]**
Heterogeneity: Tau² = 0.00; Chi² = 8.05, df = 11 (P = 0.71); I² = 0%								
Test for overall effect: Z = 1.01 (P = 0.31)								
Total (95% CI)			**709**			**666**	**100.0%**	**0.06 [-0.05, 0.16]**
Heterogeneity: Tau² = 0.00; Chi² = 8.05, df = 11 (P = 0.71); I² = 0%								
Test for overall effect: Z = 1.01 (P = 0.31)								
Test for subgroup differences: Not applicable								

Std. Mean Difference IV, Random, 95% CI

-4 -2 0 2 4

Favours experimental Favours control

Risk of Bias A B C D E

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of outcome assessment (detection bias)

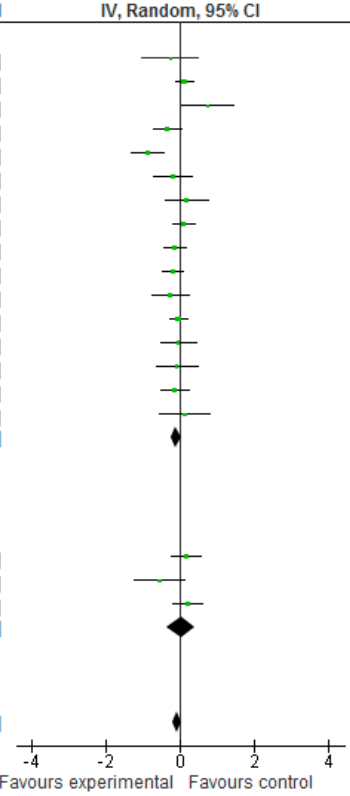
(D) Incomplete outcome data (attrition bias)

(E) Selective reporting (reporting bias)

Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher depressive symptomology*

Emotional functioning: Anxiety, post-treatment

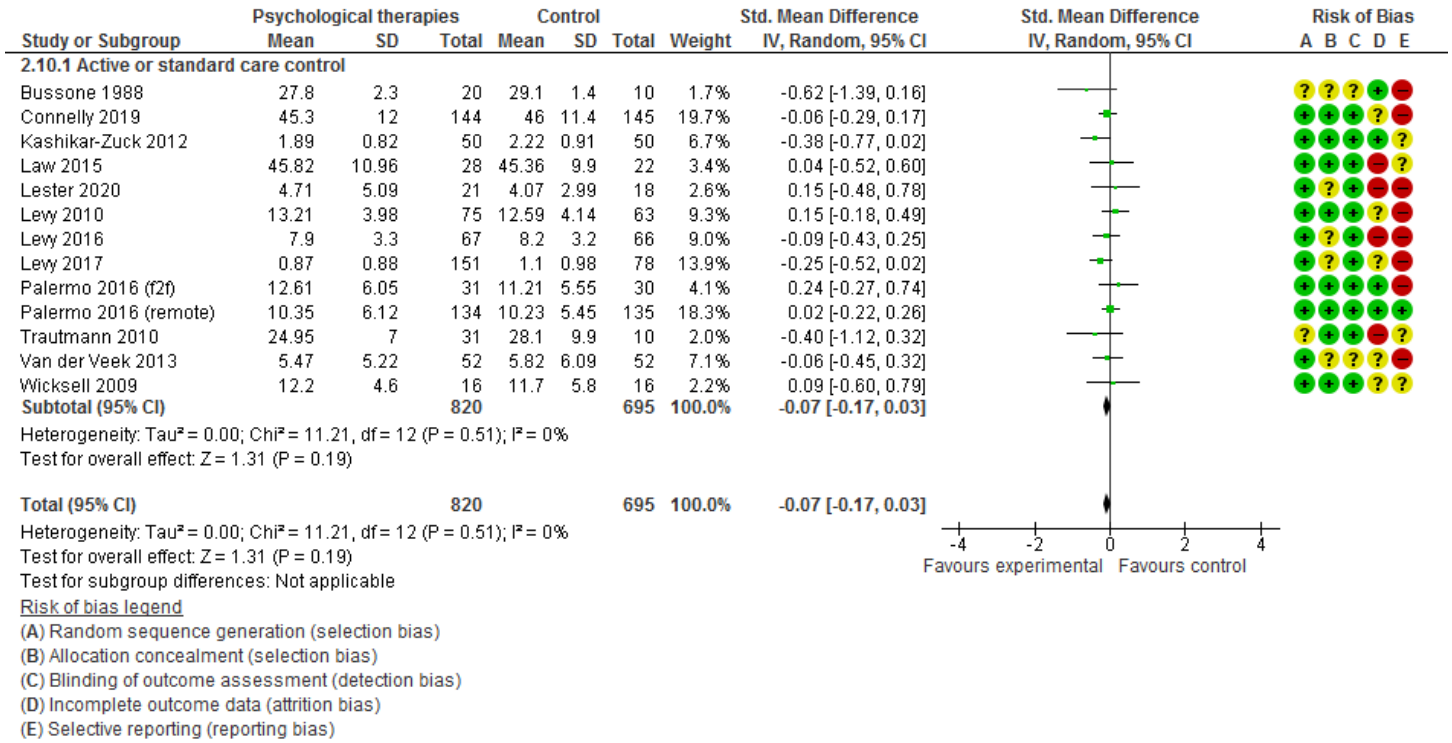
Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
2.9.1 Active or standard care control									
Bussone 1988	28.1	3.49	20	29.2	5.1	10	2.3%	-0.26 [-1.02, 0.50]	? ? ? + −
Connelly 2019	46.8	11.3	144	45.5	11	145	9.2%	0.12 [-0.11, 0.35]	+ + + ? −
Hickman 2015	52.56	7.36	16	47.38	6.1	17	2.6%	0.75 [0.04, 1.46]	? ? ? ? ?
Kashikar-Zuck 2012	2.11	0.72	50	2.39	0.9	50	5.8%	-0.34 [-0.74, 0.05]	+ + + + ?
Lalouni 2019	8.59	7.71	45	15.31	7.63	44	5.2%	-0.87 [-1.30, -0.43]	+ + ? ? −
Law 2015	46.33	8.99	30	48.32	10.81	25	4.0%	-0.20 [-0.73, 0.33]	+ + + − ?
Lester 2020	7.08	6.24	24	6.1	4.96	21	3.4%	0.17 [-0.42, 0.76]	+ ? + − −
Levy 2010	13.5	4.86	83	13.04	4.04	80	7.5%	0.10 [-0.21, 0.41]	+ + + ? −
Levy 2016	8.2	2.8	80	8.6	2.9	78	7.4%	-0.14 [-0.45, 0.17]	+ ? + − −
Levy 2017	1.09	0.94	159	1.28	1.07	81	8.3%	-0.19 [-0.46, 0.08]	+ ? + ? −
Palermo 2016 (f2f)	11.42	5.33	31	13	6.03	30	4.3%	-0.27 [-0.78, 0.23]	+ + + + −
Palermo 2016 (remote)	10.56	5.91	134	10.85	6.1	135	9.0%	-0.05 [-0.29, 0.19]	+ + + + +
Stapersma 2018	7.1	4.14	35	7.3	4.6	33	4.6%	-0.05 [-0.52, 0.43]	+ + ? + −
Trautmann 2010	30.9	7.95	38	31.7	8.3	18	3.7%	-0.10 [-0.66, 0.46]	? + + − ?
Van der Veek 2013	6.83	6	52	7.76	6.33	52	6.0%	-0.15 [-0.53, 0.24]	+ ? ? ? −
Wicksell 2009	13.4	3.9	16	12.8	5.5	16	2.6%	0.12 [-0.57, 0.82]	+ + + ? ?
Subtotal (95% CI)			**957**			**835**	**85.7%**	**-0.10 [-0.24, 0.03]**	
Heterogeneity: Tau² = 0.03; Chi² = 26.66, df = 15 (P = 0.03); I² = 44%									
Test for overall effect: Z = 1.50 (P = 0.13)									
2.9.2 Waitlist control									
Bonnert 2017	25.23	16.32	47	22.62	16.31	54	5.8%	0.16 [-0.23, 0.55]	+ + + + +
Griffiths 1996	9.6	5.9	30	13.6	9.5	12	2.7%	-0.55 [-1.24, 0.13]	? ? ? − ?
Hechler 2014	52.5	12.1	50	50	11.4	46	5.7%	0.21 [-0.19, 0.61]	+ + ? ? −
Subtotal (95% CI)			**127**			**112**	**14.3%**	**0.03 [-0.35, 0.41]**	
Heterogeneity: Tau² = 0.05; Chi² = 3.90, df = 2 (P = 0.14); I² = 49%									
Test for overall effect: Z = 0.14 (P = 0.89)									
Total (95% CI)			**1084**			**947**	**100.0%**	**-0.08 [-0.21, 0.04]**	
Heterogeneity: Tau² = 0.03; Chi² = 31.94, df = 18 (P = 0.02); I² = 44%									
Test for overall effect: Z = 1.28 (P = 0.20)									
Test for subgroup differences: Chi² = 0.40, df = 1 (P = 0.53), I² = 0%									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Active or standard care control: ⊕⊕⊕◯ MODERATE

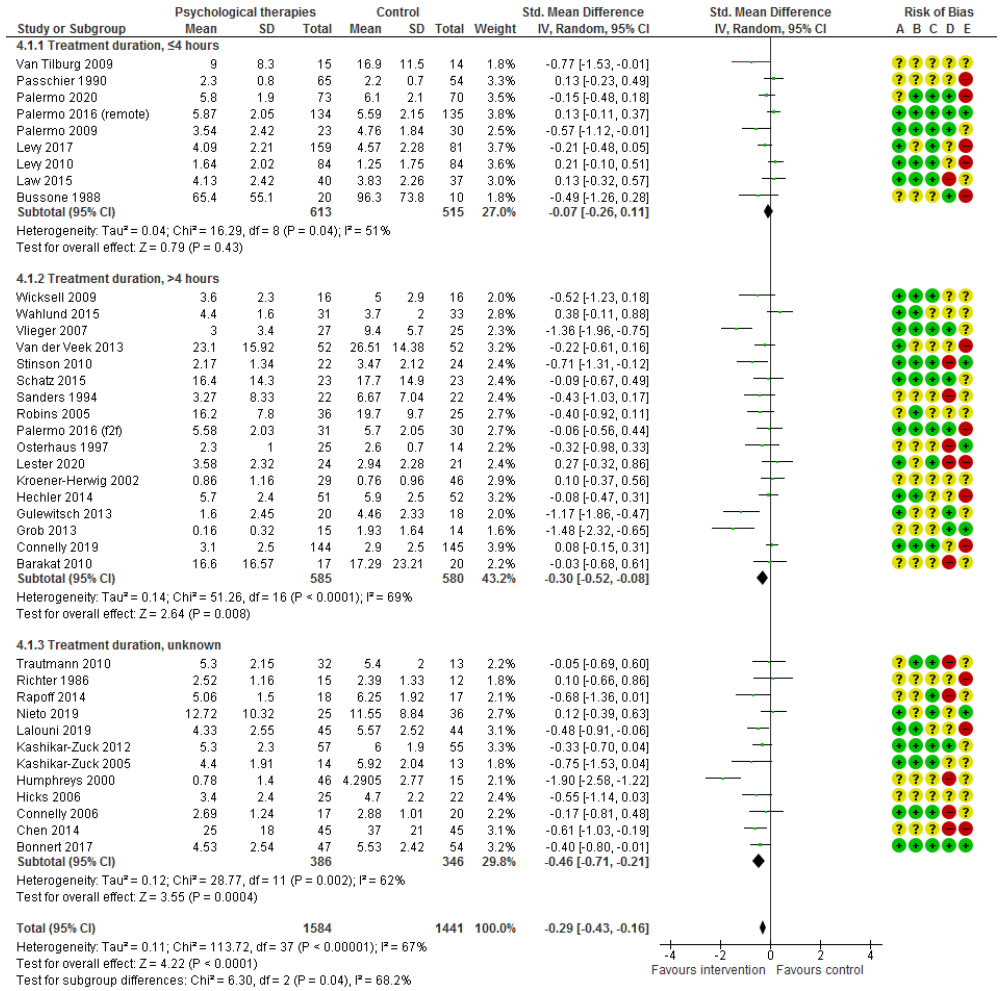
Waitlist control: ⊕◯◯◯ VERY LOW

Emotional functioning: Anxiety, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Anxiety, follow-up 	Active or standard care control: ⊕⊕⊕◯ MODERATE

Appendix G.5. WHO review: Psychological interventions for children with chronic pain
Subgroup analysis: by treatment duration

Comparison: Psychological therapies versus active (non-psychological), standard care or waitlist control; by treatment duration
Population: children and adolescents with chronic pain
Setting: Any setting
Studies: Randomised controlled trials

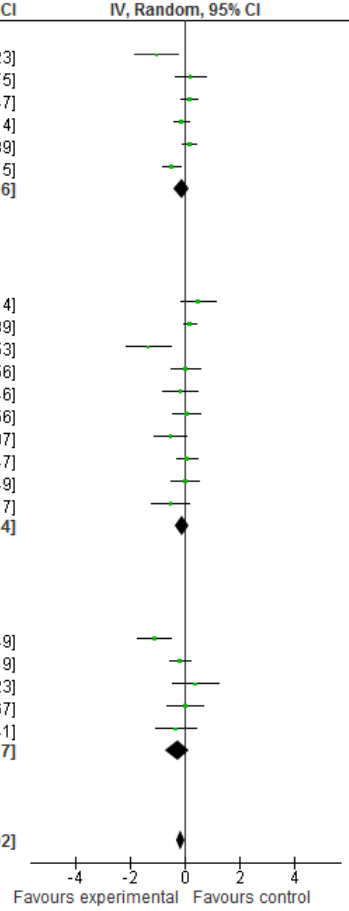
Risk of bias legend
(**A**) Random sequence generation (selection bias)
(**B**) Allocation concealment (selection bias)
(**C**) Blinding of outcome assessment (detection bias)
(**D**) Incomplete outcome data (attrition bias)
(**E**) Selective reporting (reporting bias)

Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment 	Less than 4 hours ⊕⊕◯◯ LOW More than 4 hours ⊕⊕◯◯ LOW

Pain intensity, follow-up
Higher scores indicate higher pain intensity

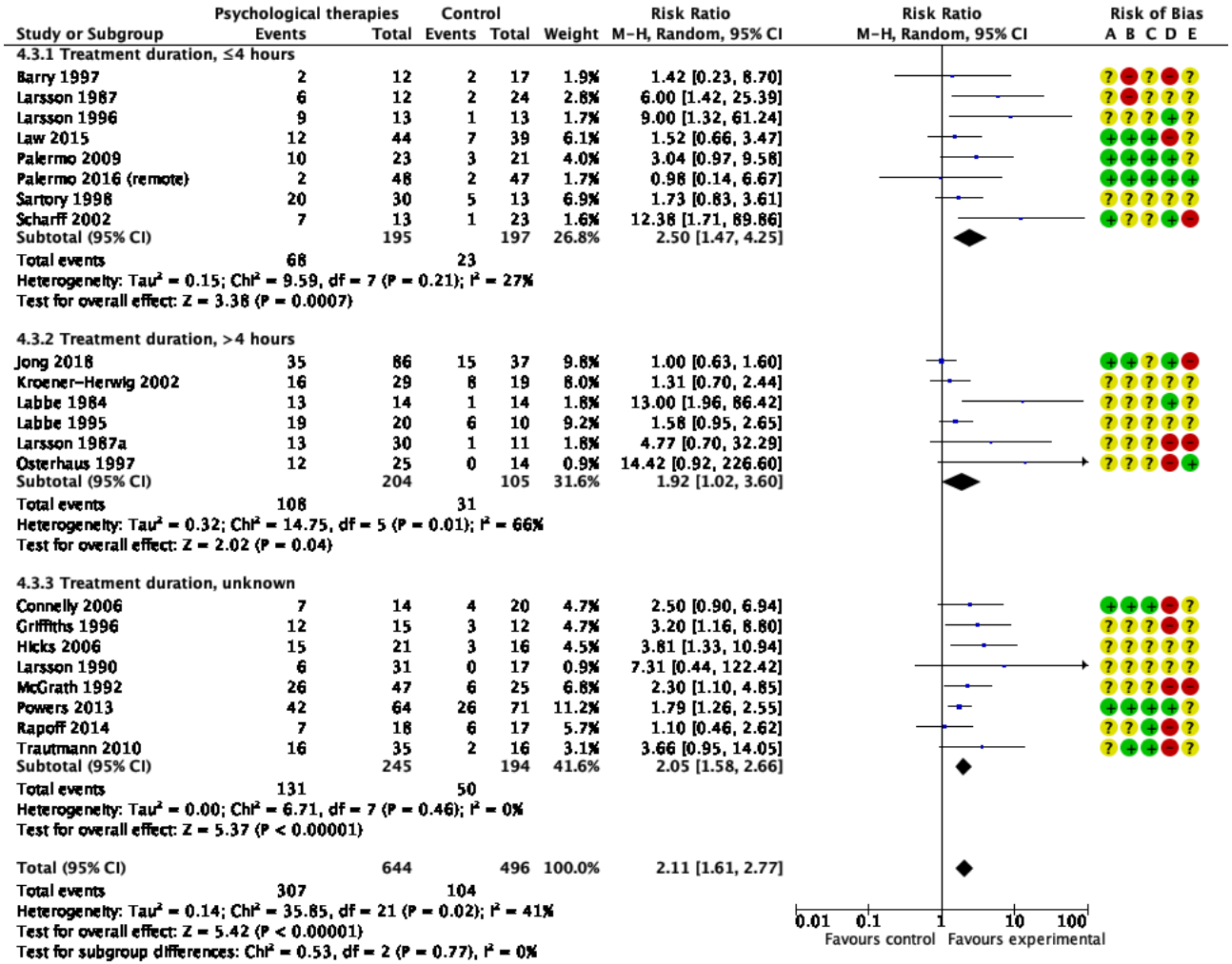
Pain intensity, follow-up

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.2.1 Treatment duration, ≤4 hours								
Bussone 1988	20	18.1	20	88.8	110.3	10	2.8%	-1.04 [-1.85, -0.23]
Law 2015	4.19	2.45	28	3.7	2.54	22	4.4%	0.19 [-0.37, 0.75]
Lewy 2010	0.93	1.42	78	0.7	1.53	76	6.7%	0.16 [-0.16, 0.47]
Lewy 2017	3.48	2.33	151	3.79	2.48	78	7.2%	-0.13 [-0.40, 0.14]
Palermo 2016 (remote)	5.85	1.97	134	5.55	2.02	135	7.5%	0.15 [-0.09, 0.39]
Palermo 2020	5.3	1.9	73	6.2	1.8	70	6.5%	-0.48 [-0.82, -0.15]
Subtotal (95% CI)			**484**			**391**	**35.1%**	**-0.11 [-0.39, 0.16]**
Heterogeneity: Tau² = 0.08; Chi² = 17.62, df = 5 (P = 0.003); I² = 72%								
Test for overall effect: Z = 0.81 (P = 0.42)								
4.2.2 Treatment duration, >4 hours								
Barakat 2010	16.71	23.03	17	7.84	12.31	20	3.6%	0.48 [-0.18, 1.14]
Connelly 2019	3.1	2.5	144	2.7	2.4	145	7.6%	0.16 [-0.07, 0.39]
Grob 2013	0.08	0.31	15	1.55	1.49	14	2.7%	-1.35 [-2.17, -0.53]
Kroener-Herwig 2002	0.48	0.65	29	0.46	0.61	27	4.6%	0.03 [-0.49, 0.56]
Lester 2020	2.67	1.9	21	3.07	2.64	18	3.8%	-0.17 [-0.80, 0.46]
Palermo 2016 (f2f)	5.42	2.05	31	5.3	2.12	30	4.8%	0.06 [-0.45, 0.56]
Sanders 1994	0.64	1.38	22	2.11	3.56	22	4.0%	-0.53 [-1.14, 0.07]
Van der Veek 2013	19.03	17.0393	52	17.72	15.19	52	6.0%	0.08 [-0.30, 0.47]
Wahlund 2015	2.8	1.9	31	2.8	1.6	33	5.0%	0.00 [-0.49, 0.49]
Wicksell 2009	3.1	2.7	16	4.5	2.4	16	3.3%	-0.53 [-1.24, 0.17]
Subtotal (95% CI)			**378**			**377**	**45.6%**	**-0.10 [-0.34, 0.14]**
Heterogeneity: Tau² = 0.08; Chi² = 20.17, df = 9 (P = 0.02); I² = 55%								
Test for overall effect: Z = 0.78 (P = 0.43)								
4.2.3 Treatment duration, unknown								
Hicks 2006	2.9	2.1	25	4.9	1.3	22	3.9%	-1.11 [-1.73, -0.49]
Kashikar-Zuck 2012	4.9	2.2	57	5.3	2.1	55	6.1%	-0.18 [-0.56, 0.19]
Rapoff 2014	4.46	1.88	11	3.68	2.04	11	2.6%	0.38 [-0.46, 1.23]
Richter 1986	2.02	1.48	30	2.02	1.39	12	3.6%	0.00 [-0.67, 0.67]
Trautmann 2010	4.9	1.4	12	5.5	1.9	16	3.1%	-0.34 [-1.10, 0.41]
Subtotal (95% CI)			**135**			**116**	**19.3%**	**-0.28 [-0.72, 0.17]**
Heterogeneity: Tau² = 0.15; Chi² = 10.21, df = 4 (P = 0.04); I² = 61%								
Test for overall effect: Z = 1.23 (P = 0.22)								
Total (95% CI)			**997**			**884**	**100.0%**	**-0.14 [-0.30, 0.02]**
Heterogeneity: Tau² = 0.08; Chi² = 51.59, df = 20 (P = 0.0001); I² = 61%								
Test for overall effect: Z = 1.67 (P = 0.09)								
Test for subgroup differences: Chi² = 0.52, df = 2 (P = 0.77), I² = 0%								



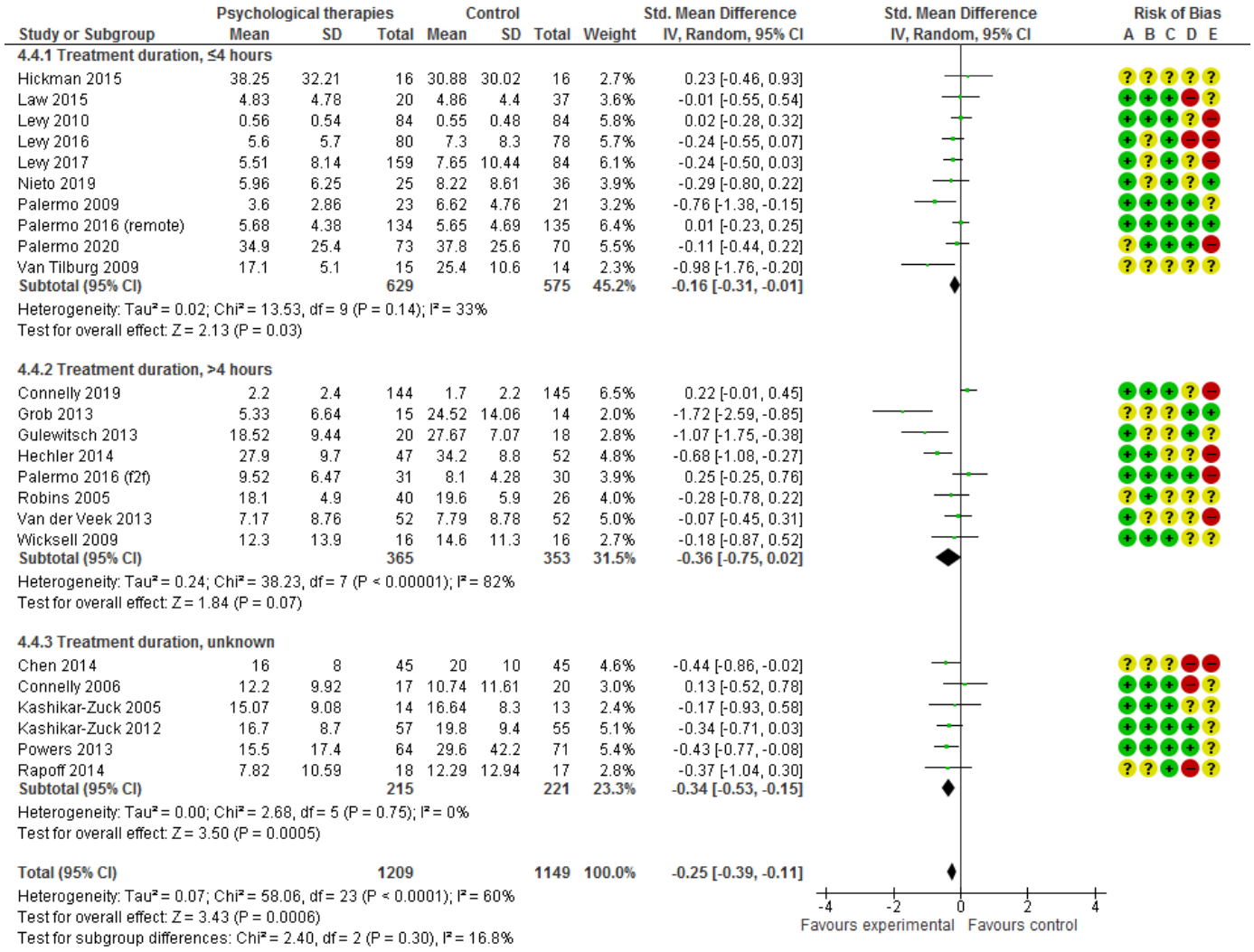
Less than 4 hours
⊕⊕⊕◯
MODERATE

More than 4 hours
⊕⊕◯◯
LOW

50% reduction, post-treatment	50% reduction, post-treatment 	Less than 4 hours ⊕◯◯◯ VERY LOW More than 4 hours ⊕◯◯◯ VERY LOW

Functional disability, post-treatment
Higher scores indicate lower disability

Functional disability, post-treatment

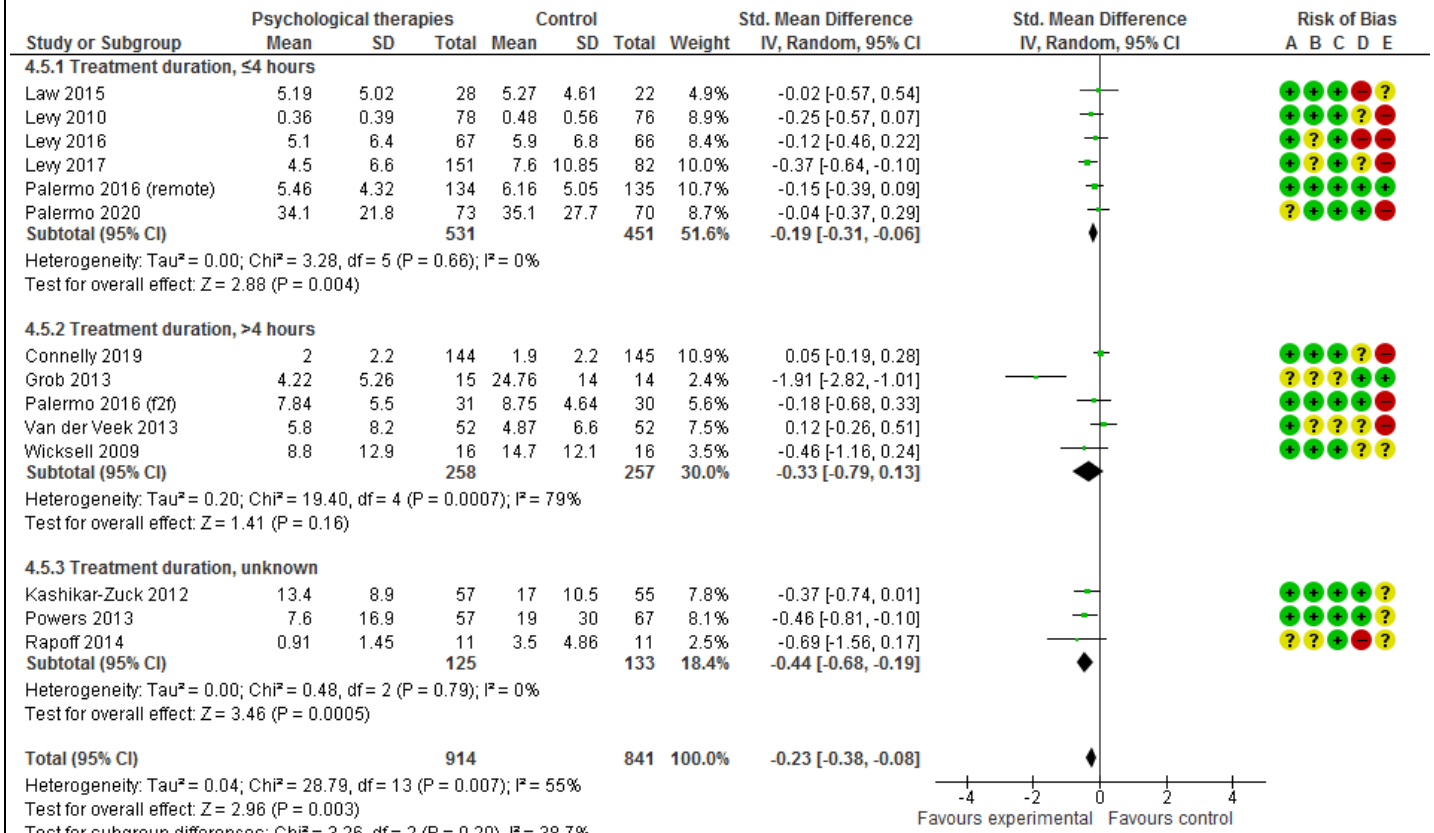


Less than 4 hours
⊕⊕⊕◯
MODERATE

More than 4 hours
⊕◯◯◯
VERY LOW

Functional disability, follow-up
Higher scores indicate lower disability

Functional disability, follow-up

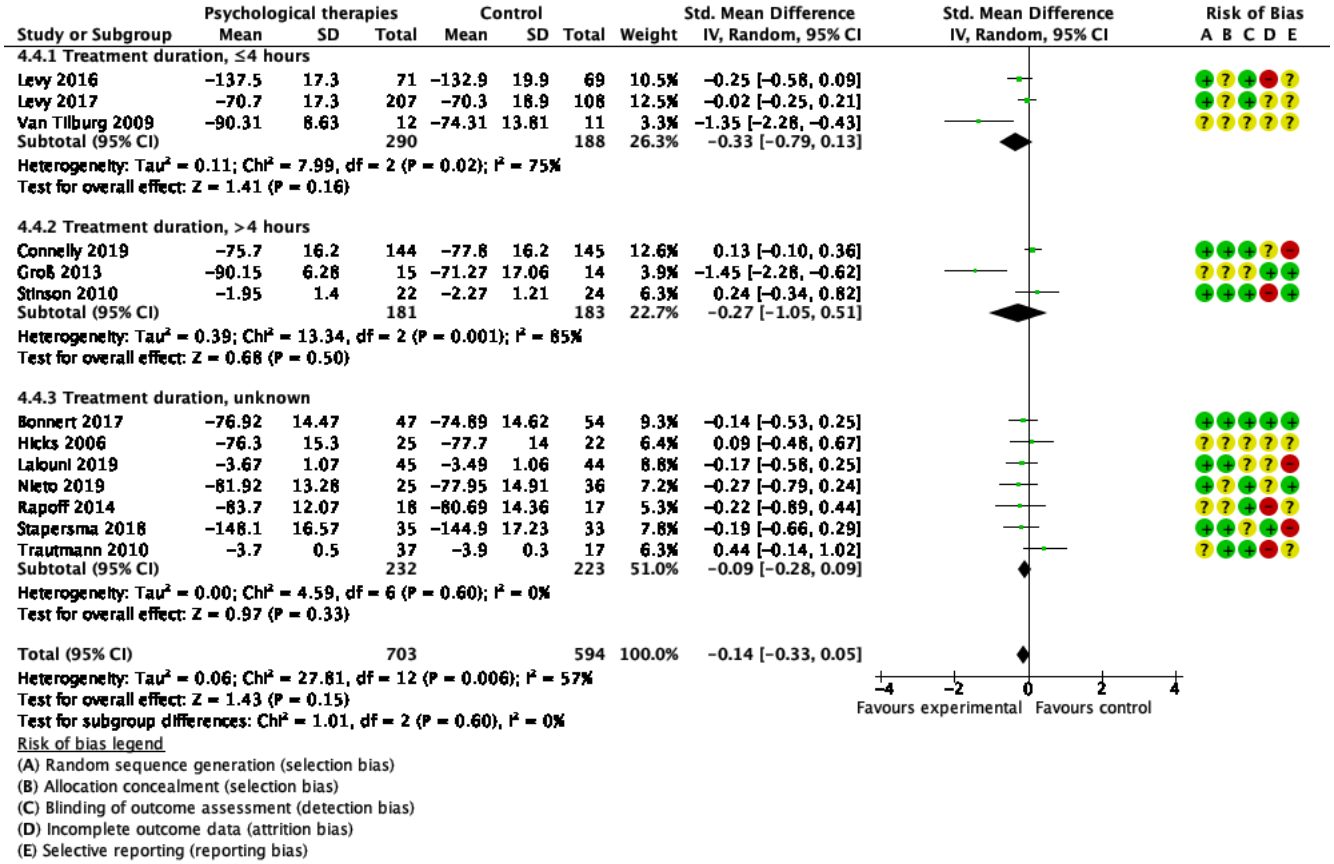


Less than 4 hours
⊕⊕⊕◯
MODERATE

More than 4 hours
⊕◯◯◯
VERY LOW

Health-related quality of life, post-treatment *Lower scores indicate better quality of life*

Health-related quality of life, post-treatment

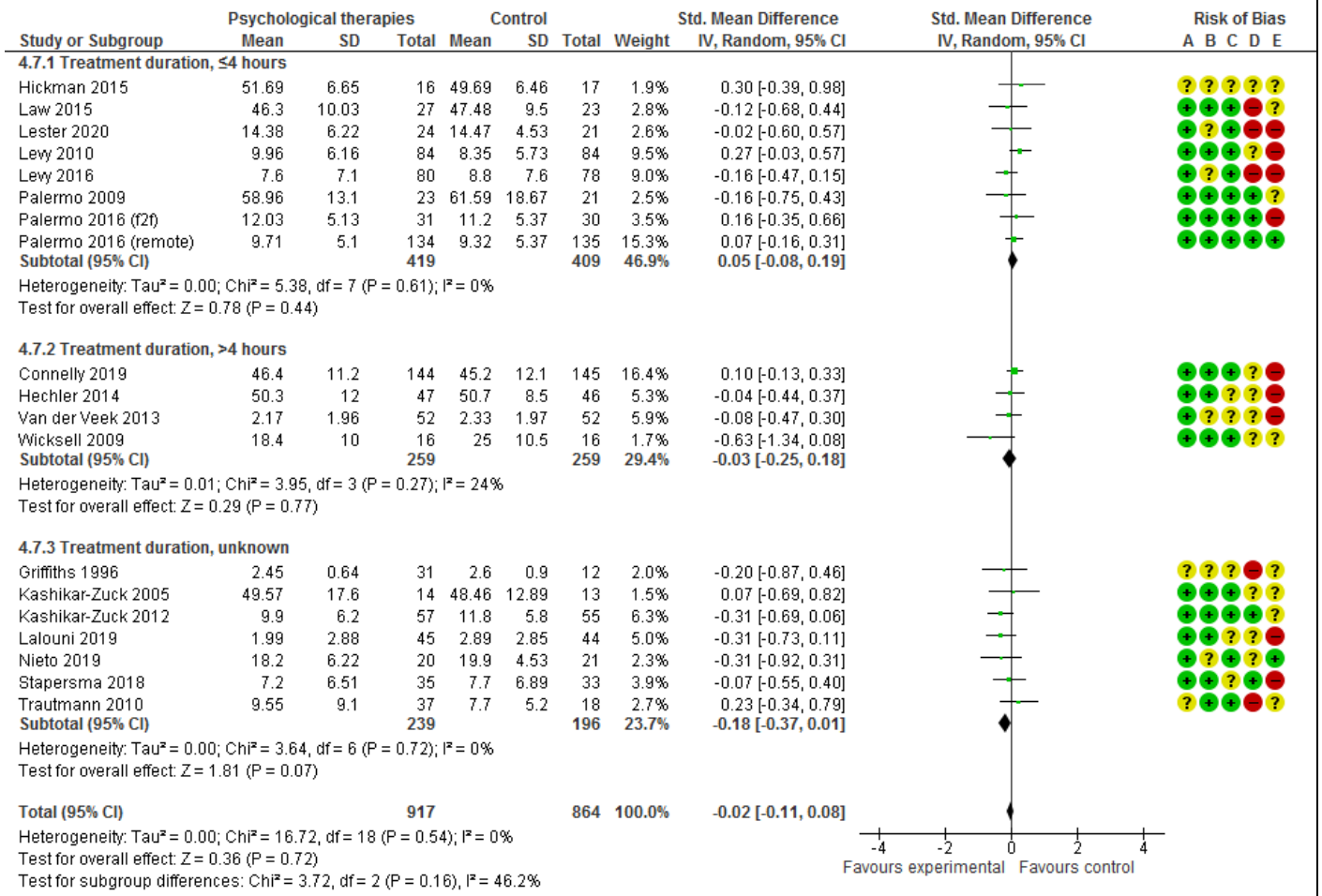


Less than 4 hours
⊕◯◯◯
VERY LOW

More than 4 hours
⊕◯◯◯
VERY LOW

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*

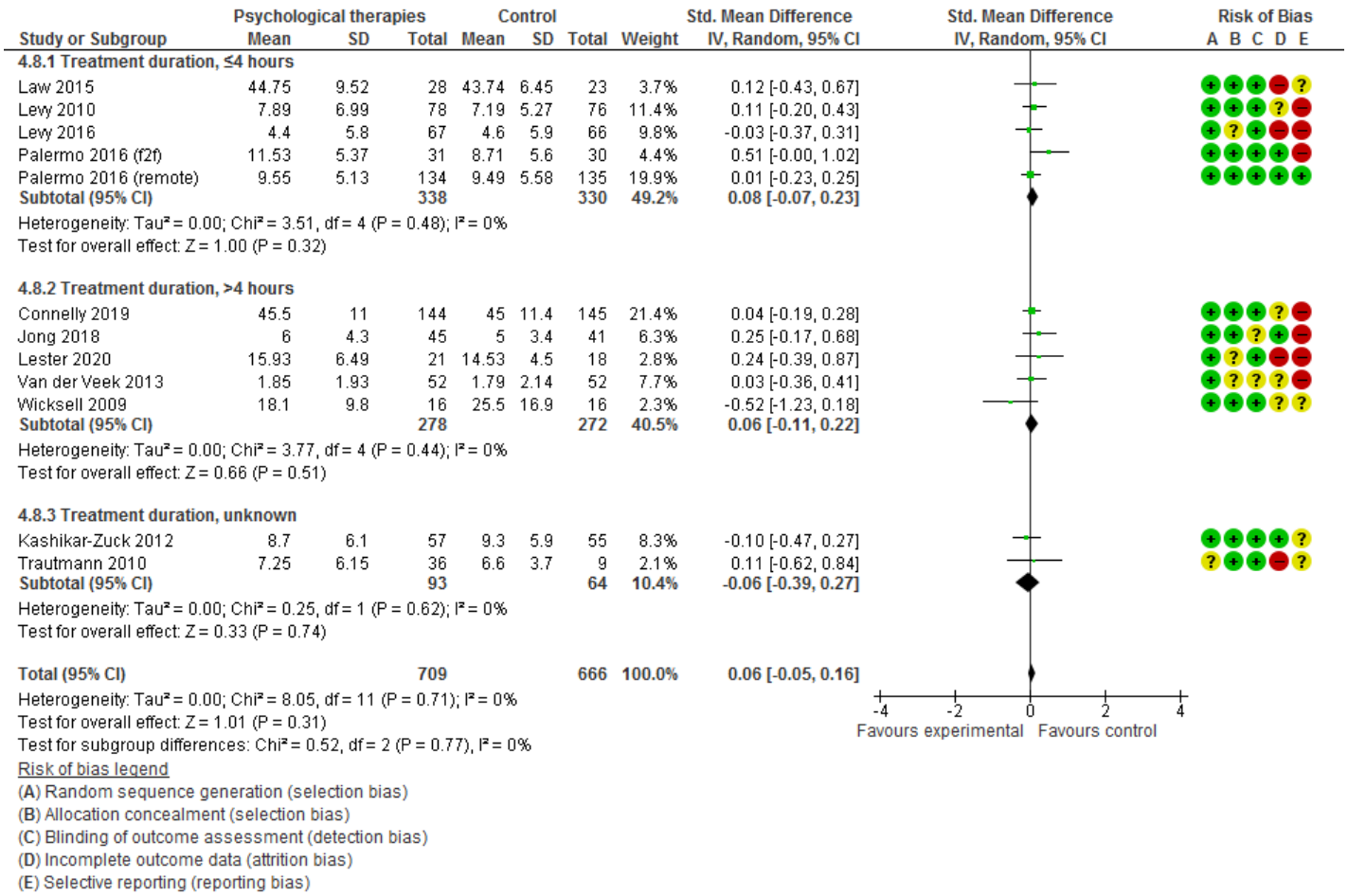
Emotional functioning: Depression, post-treatment



Less than 4 hours
⊕⊕⊕⊕
HIGH

More than 4 hours
⊕⊕⊕◯
MODERATE

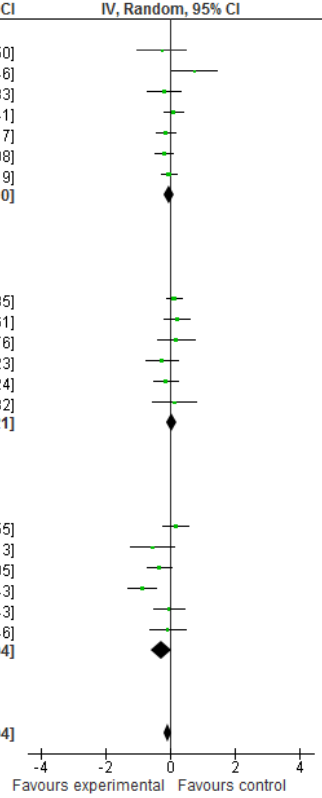
Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow up	Less than 4 hours ⊕⊕⊕⊕ HIGH More than 4 hours ⊕⊕⊕◯ MODERATE



Emotional functioning: Anxiety, post-treatment
Higher scores indicate higher depressive symptomology

Emotional functioning: Anxiety, post-treatment

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
4.9.1 Treatment duration, ≤4 hours									
Bussone 1988	28.1	3.49	20	29.2	5.1	10	2.3%	-0.26 [-1.02, 0.50]	? ? ? + -
Hickman 2015	52.56	7.36	16	47.38	6.1	17	2.6%	0.75 [0.04, 1.46]	? ? ? ? ?
Law 2015	46.33	8.99	30	48.32	10.81	25	4.0%	-0.20 [-0.73, 0.33]	+ + + - ?
Lewy 2010	13.5	4.86	83	13.04	4.04	80	7.5%	0.10 [-0.21, 0.41]	+ + + ? -
Lewy 2016	8.2	2.8	80	8.6	2.9	78	7.4%	-0.14 [-0.45, 0.17]	+ ? + - -
Lewy 2017	1.09	0.94	159	1.28	1.07	81	8.3%	-0.19 [-0.46, 0.08]	+ ? + ? -
Palermo 2016 (remote)	10.56	5.91	134	10.85	6.1	135	9.0%	-0.05 [-0.29, 0.19]	+ + + + +
Subtotal (95% CI)			**522**			**426**	**41.0%**	**-0.05 [-0.21, 0.10]**	
Heterogeneity: Tau² = 0.01; Chi² = 7.80, df = 6 (P = 0.25); I² = 23%									
Test for overall effect: Z = 0.66 (P = 0.51)									
4.9.2 Treatment duration, >4 hours									
Connelly 2019	46.8	11.3	144	45.5	11	145	9.2%	0.12 [-0.11, 0.35]	+ + + ? -
Hechler 2014	52.5	12.1	50	50	11.4	46	5.7%	0.21 [-0.19, 0.61]	+ + ? ? -
Lester 2020	7.08	6.24	24	6.1	4.96	21	3.4%	0.17 [-0.42, 0.76]	+ ? + - -
Palermo 2016 (f2f)	11.42	5.33	31	13	6.03	30	4.3%	-0.27 [-0.78, 0.23]	+ + + + -
Van der Veek 2013	6.83	6	52	7.76	6.33	52	6.0%	-0.15 [-0.53, 0.24]	+ ? ? ? -
Wicksell 2009	13.4	3.9	16	12.8	5.5	16	2.6%	0.12 [-0.57, 0.82]	+ + + ? ?
Subtotal (95% CI)			**317**			**310**	**31.2%**	**0.05 [-0.10, 0.21]**	
Heterogeneity: Tau² = 0.00; Chi² = 3.75, df = 5 (P = 0.59); I² = 0%									
Test for overall effect: Z = 0.66 (P = 0.51)									
4.9.3 Treatment duration, unknown									
Bonnert 2017	25.23	16.32	47	22.62	16.31	54	5.8%	0.16 [-0.23, 0.55]	+ + + + +
Griffiths 1996	9.6	5.9	30	13.6	9.5	12	2.7%	-0.55 [-1.24, 0.13]	? ? ? - ?
Kashikar-Zuck 2012	2.11	0.72	50	2.39	0.9	50	5.8%	-0.34 [-0.74, 0.05]	+ + + + ?
Lalouni 2019	8.59	7.71	45	15.31	7.63	44	5.2%	-0.87 [-1.30, -0.43]	+ + ? ? -
Stapersma 2018	7.1	4.14	35	7.3	4.6	33	4.6%	-0.05 [-0.52, 0.43]	+ + ? + -
Trautmann 2010	30.9	7.95	38	31.7	8.3	18	3.7%	-0.10 [-0.66, 0.46]	? + + - ?
Subtotal (95% CI)			**245**			**211**	**27.8%**	**-0.28 [-0.60, 0.04]**	
Heterogeneity: Tau² = 0.10; Chi² = 13.87, df = 5 (P = 0.02); I² = 64%									
Test for overall effect: Z = 1.70 (P = 0.09)									
Total (95% CI)			**1084**			**947**	**100.0%**	**-0.08 [-0.21, 0.04]**	
Heterogeneity: Tau² = 0.03; Chi² = 31.94, df = 18 (P = 0.02); I² = 44%									
Test for overall effect: Z = 1.28 (P = 0.20)									
Test for subgroup differences: Chi² = 3.46, df = 2 (P = 0.18), I² = 42.3%									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

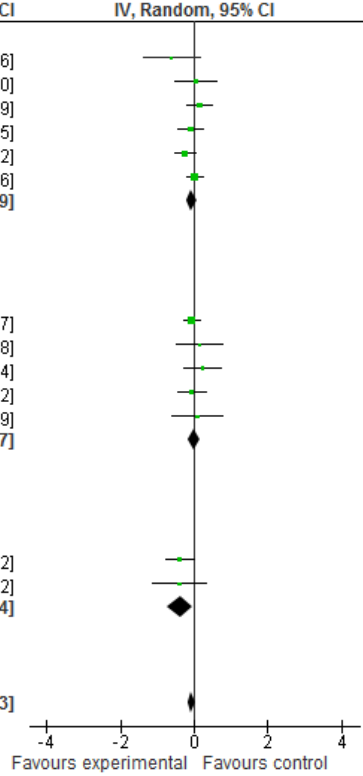
Less than 4 hours
⊕⊕⊕◯
MODERATE

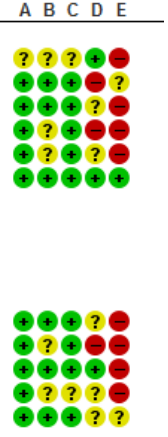
More than 4 hours
⊕⊕⊕◯
MODERATE

Emotional functioning: Anxiety, follow-up
Higher scores indicate higher depressive symptomology

Emotional functioning: Anxiety, follow-up

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.10.1 Treatment duration, ≤4 hours								
Bussone 1988	27.8	2.3	20	29.1	1.4	10	1.7%	-0.62 [-1.39, 0.16]
Law 2015	45.82	10.96	28	45.36	9.9	22	3.4%	0.04 [-0.52, 0.60]
Levy 2010	13.21	3.98	75	12.59	4.14	63	9.3%	0.15 [-0.18, 0.49]
Levy 2016	7.9	3.3	67	8.2	3.2	66	9.0%	-0.09 [-0.43, 0.25]
Levy 2017	0.87	0.88	151	1.1	0.98	78	13.9%	-0.25 [-0.52, 0.02]
Palermo 2016 (remote)	10.35	6.12	134	10.23	5.45	135	18.3%	0.02 [-0.22, 0.26]
Subtotal (95% CI)			**475**			**374**	**55.6%**	**-0.06 [-0.22, 0.09]**
Heterogeneity: Tau² = 0.01; Chi² = 5.95, df = 5 (P = 0.31); I² = 16%								
Test for overall effect: Z = 0.81 (P = 0.42)								
4.10.2 Treatment duration, >4 hours								
Connelly 2019	45.3	12	144	46	11.4	145	19.7%	-0.06 [-0.29, 0.17]
Lester 2020	4.71	5.09	21	4.07	2.99	18	2.6%	0.15 [-0.48, 0.78]
Palermo 2016 (f2f)	12.61	6.05	31	11.21	5.55	30	4.1%	0.24 [-0.27, 0.74]
Van der Veek 2013	5.47	5.22	52	5.82	6.09	52	7.1%	-0.06 [-0.45, 0.32]
Wicksell 2009	12.2	4.6	16	11.7	5.8	16	2.2%	0.09 [-0.60, 0.79]
Subtotal (95% CI)			**264**			**261**	**35.7%**	**-0.00 [-0.17, 0.17]**
Heterogeneity: Tau² = 0.00; Chi² = 1.49, df = 4 (P = 0.83); I² = 0%								
Test for overall effect: Z = 0.01 (P = 0.99)								
4.10.3 Treatment duration, unknown								
Kashikar-Zuck 2012	1.89	0.82	50	2.22	0.91	50	6.7%	-0.38 [-0.77, 0.02]
Trautmann 2010	24.95	7	31	28.1	9.9	10	2.0%	-0.40 [-1.12, 0.32]
Subtotal (95% CI)			**81**			**60**	**8.7%**	**-0.38 [-0.73, -0.04]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.96); I² = 0%								
Test for overall effect: Z = 2.16 (P = 0.03)								
Total (95% CI)			**820**			**695**	**100.0%**	**-0.07 [-0.17, 0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 11.21, df = 12 (P = 0.51); I² = 0%								
Test for overall effect: Z = 1.31 (P = 0.19)								
Test for subgroup differences: Chi² = 3.76, df = 2 (P = 0.15), I² = 46.7%								





Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

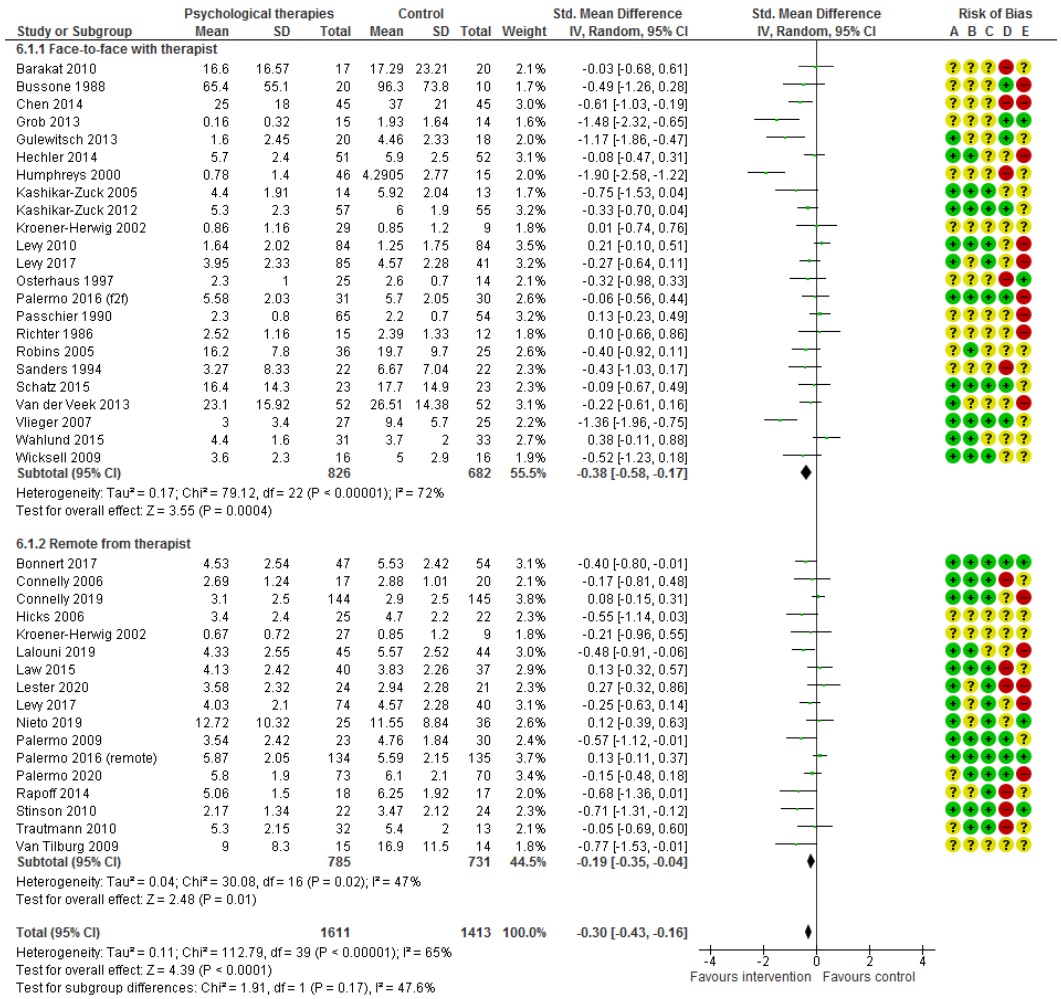
Less than 4 hours
⊕⊕⊕⊕
HIGH

More than 4 hours
⊕⊕⊕⊕
HIGH

Appendix G.6. WHO review: Psychological interventions for children with chronic pain
Subgroup analysis: by route

Comparison: Psychological therapies versus active (non-psychological), standard care or waitlist control; by route of intervention
Population: Children and adolescents with chronic pain
Setting: Any setting
Studies: Randomised controlled trials

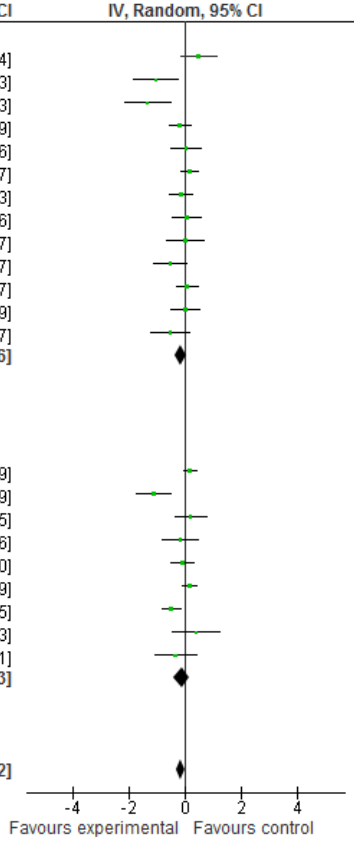
Risk of bias legend
(**A**) Random sequence generation (selection bias)
(**B**) Allocation concealment (selection bias)
(**C**) Blinding of outcome assessment (detection bias)
(**D**) Incomplete outcome data (attrition bias)
(**E**) Selective reporting (reporting bias)

Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment 	Face to Face with therapist: ⊕⊕◯◯ LOW Remote from therapist: ⊕⊕⊕◯ MODERATE

Pain intensity, follow-up *Higher scores indicate higher pain intensity*

Pain intensity, follow-up

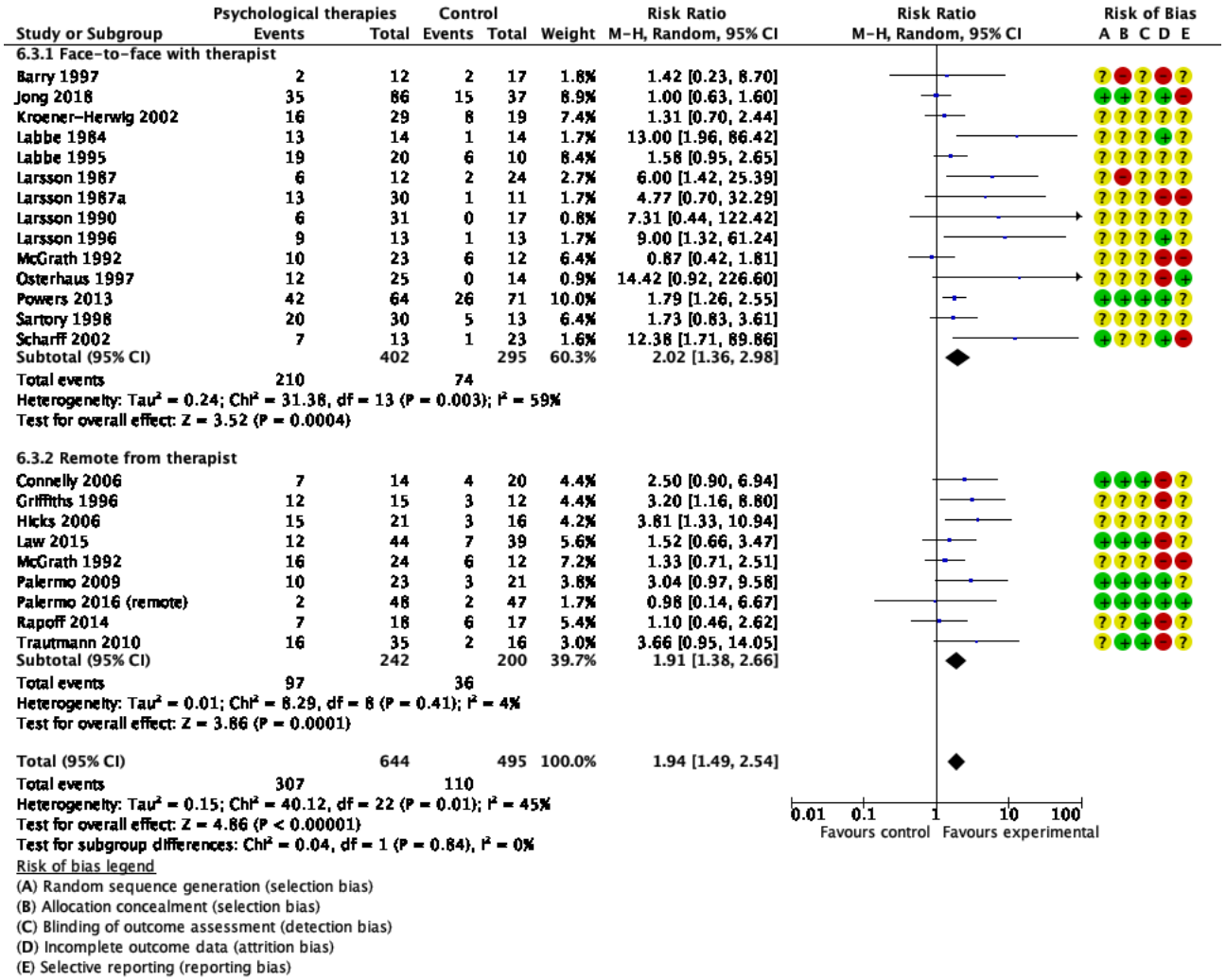
Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
6.2.1 Face-to-face with therapist									
Barakat 2010	16.71	23.03	17	7.84	12.31	20	3.5%	0.48 [-0.18, 1.14]	? ? ? − ?
Bussone 1988	20	18.1	20	88.8	110.3	10	2.6%	-1.04 [-1.85, -0.23]	? ? ? + −
Grob 2013	0.08	0.31	15	1.55	1.49	14	2.6%	-1.35 [-2.17, -0.53]	? ? ? + +
Kashikar-Zuck 2012	4.9	2.2	57	5.3	2.1	55	5.9%	-0.18 [-0.56, 0.19]	+ + + + ?
Kroener-Herwig 2002	0.48	0.65	29	0.46	0.61	27	4.4%	0.03 [-0.49, 0.56]	? ? ? ? ?
Lewy 2010	0.93	1.42	78	0.7	1.53	76	6.5%	0.16 [-0.16, 0.47]	+ + + ? −
Lewy 2017	3.41	2.35	81	3.79	2.48	37	5.7%	-0.16 [-0.55, 0.23]	+ ? + ? −
Palermo 2016 (f2f)	5.42	2.05	31	5.3	2.12	30	4.6%	0.06 [-0.45, 0.56]	+ + + + −
Richter 1986	2.02	1.48	30	2.02	1.39	12	3.4%	0.00 [-0.67, 0.67]	? ? ? ? −
Sanders 1994	0.64	1.38	22	2.11	3.56	22	3.8%	-0.53 [-1.14, 0.07]	? ? ? − ?
Van der Veek 2013	19.03	17.0393	52	17.72	15.19	52	5.7%	0.08 [-0.30, 0.47]	+ ? ? ? −
Wahlund 2015	2.8	1.9	31	2.8	1.6	33	4.7%	0.00 [-0.49, 0.49]	+ + ? ? ?
Wicksell 2009	3.1	2.7	16	4.5	2.4	16	3.2%	-0.53 [-1.24, 0.17]	+ + + ? ?
Subtotal (95% CI)			**479**			**404**	**56.6%**	**-0.15 [-0.35, 0.06]**	
Heterogeneity: Tau² = 0.07; Chi² = 25.09, df = 12 (P = 0.01); I² = 52%									
Test for overall effect: Z = 1.40 (P = 0.16)									
6.2.2 Remote from therapist									
Connelly 2019	3.1	2.5	144	2.7	2.4	145	7.4%	0.16 [-0.07, 0.39]	+ + + ? −
Hicks 2006	2.9	2.1	25	4.9	1.3	22	3.7%	-1.11 [-1.73, -0.49]	? ? ? ? ?
Law 2015	4.19	2.45	28	3.7	2.54	22	4.2%	0.19 [-0.37, 0.75]	+ + + − ?
Lester 2020	2.67	1.9	21	3.07	2.64	18	3.6%	-0.17 [-0.80, 0.46]	+ ? + − −
Lewy 2017	3.54	2.3	70	3.79	2.48	36	5.6%	-0.11 [-0.51, 0.30]	+ ? + ? −
Palermo 2016 (remote)	5.85	1.97	134	5.55	2.02	135	7.3%	0.15 [-0.09, 0.39]	+ + + + +
Palermo 2020	5.3	1.9	73	6.2	1.8	70	6.3%	-0.48 [-0.82, -0.15]	? + + + −
Rapoff 2014	4.46	1.88	11	3.68	2.04	11	2.5%	0.38 [-0.46, 1.23]	? ? + − ?
Trautmann 2010	4.9	1.4	12	5.5	1.9	16	2.9%	-0.34 [-1.10, 0.41]	? + + − ?
Subtotal (95% CI)			**518**			**475**	**43.4%**	**-0.13 [-0.39, 0.13]**	
Heterogeneity: Tau² = 0.09; Chi² = 26.18, df = 8 (P = 0.0010); I² = 69%									
Test for overall effect: Z = 0.97 (P = 0.33)									
Total (95% CI)			**997**			**879**	**100.0%**	**-0.14 [-0.30, 0.02]**	
Heterogeneity: Tau² = 0.07; Chi² = 51.63, df = 21 (P = 0.0002); I² = 59%									
Test for overall effect: Z = 1.71 (P = 0.09)									
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.91), I² = 0%									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

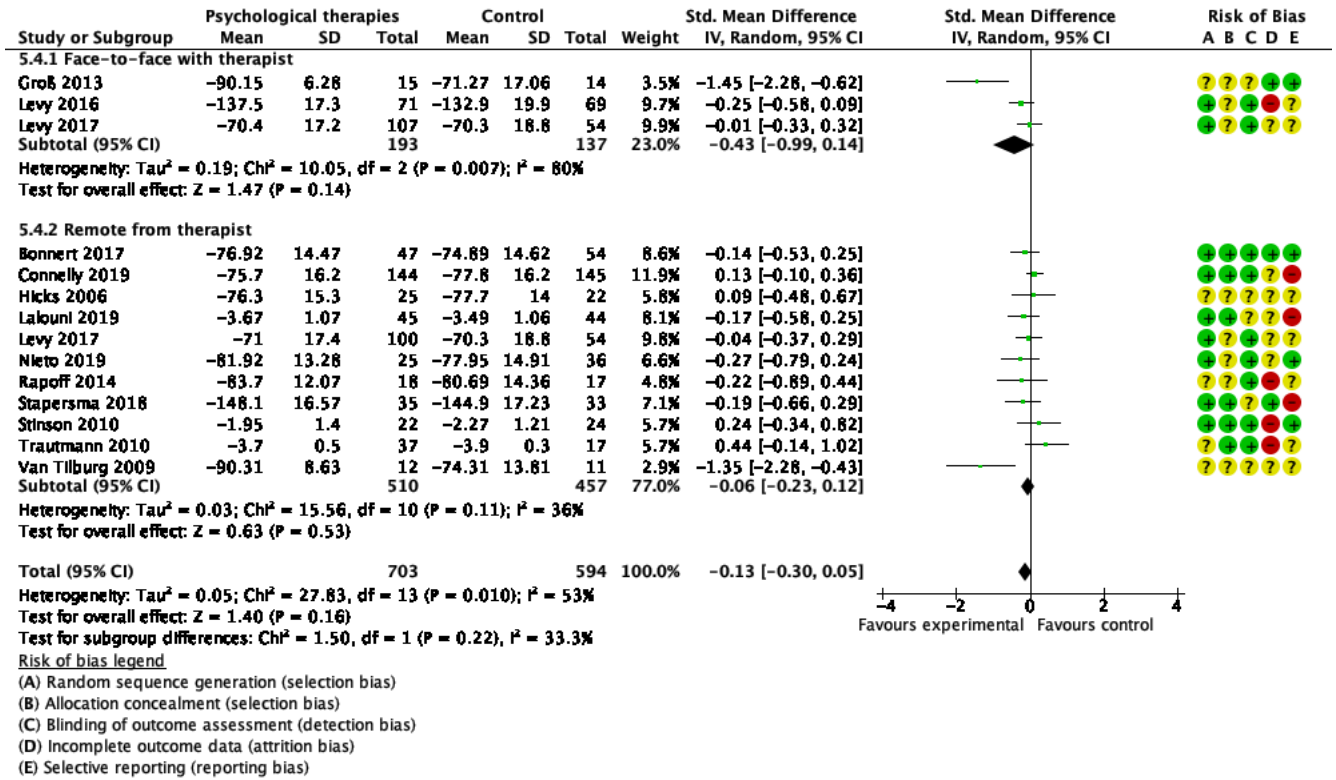
Face to Face with therapist:
⊕⊕◯◯
LOW

Remote from therapist:
⊕⊕◯◯
LOW

50% reduction, post-treatment	50% reduction, post-treatment 	Face to Face with therapist: ⊕◯◯◯ VERY LOW Remote from therapist: ⊕⊕◯◯ LOW

Health-related quality of life, post-treatment
Lower scores indicate better quality of life

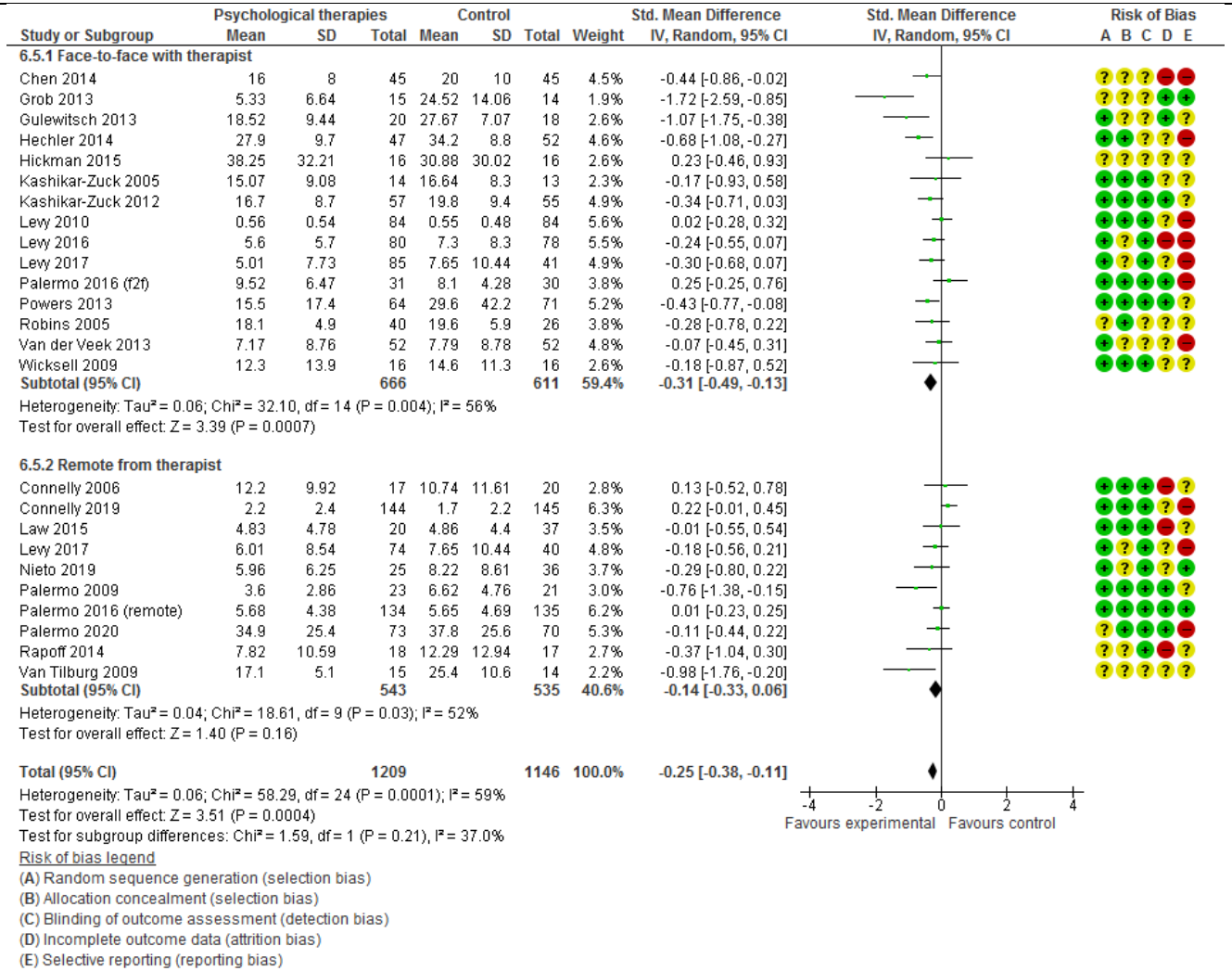
Health-related quality of life, post-treatment



Face to Face with therapist:
⊕◯◯◯
VERY LOW

Remote from therapist:
⊕⊕⊕◯
MODERATE

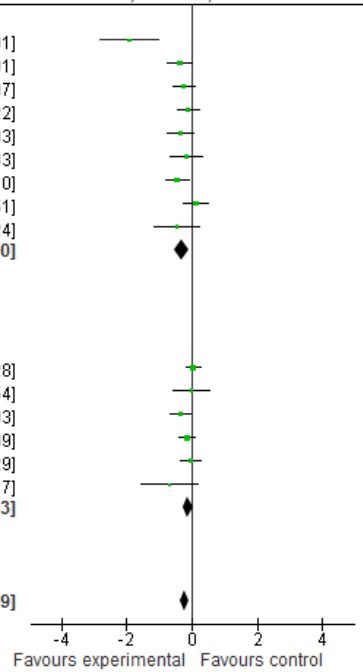
Functional disability, post-treatment *Higher scores indicate lower disability*	Functional disability, post-treatment	Face to Face with therapist: ⊕⊕◯◯ LOW Remote from therapist: ⊕⊕⊕◯ MODERATE


Psychological therapies
Control
Std. Mean Difference
Std. Mean Difference
Risk of Bias
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
A B C D E
6.5.1 Face-to-face with therapist
Chen 2014 16 8 45 20 10 45 4.5% -0.44 [-0.86, -0.02]
Grob 2013 5.33 6.64 15 24.52 14.06 14 1.9% -1.72 [-2.59, -0.85]
Gulewitsch 2013 18.52 9.44 20 27.67 7.07 18 2.6% -1.07 [-1.75, -0.38]
Hechler 2014 27.9 9.7 47 34.2 8.8 52 4.6% -0.68 [-1.08, -0.27]
Hickman 2015 38.25 32.21 16 30.88 30.02 16 2.6% 0.23 [-0.46, 0.93]
Kashikar-Zuck 2005 15.07 9.08 14 16.64 8.3 13 2.3% -0.17 [-0.93, 0.58]
Kashikar-Zuck 2012 16.7 8.7 57 19.8 9.4 55 4.9% -0.34 [-0.71, 0.03]
Lewy 2010 0.56 0.54 84 0.55 0.48 84 5.6% 0.02 [-0.28, 0.32]
Lewy 2016 5.6 5.7 80 7.3 8.3 78 5.5% -0.24 [-0.55, 0.07]
Lewy 2017 5.01 7.73 85 7.65 10.44 41 4.9% -0.30 [-0.68, 0.07]
Palermo 2016 (f2f) 9.52 6.47 31 8.1 4.28 30 3.8% 0.25 [-0.25, 0.76]
Powers 2013 15.5 17.4 64 29.6 42.2 71 5.2% -0.43 [-0.77, -0.08]
Robins 2005 18.1 4.9 40 19.6 5.9 26 3.8% -0.28 [-0.78, 0.22]
Van der Veek 2013 7.17 8.76 52 7.79 8.78 52 4.8% -0.07 [-0.45, 0.31]
Wicksell 2009 12.3 13.9 16 14.6 11.3 16 2.6% -0.18 [-0.87, 0.52]
Subtotal (95% CI) 666 611 59.4% -0.31 [-0.49, -0.13]
Heterogeneity: Tau² = 0.06; Chi² = 32.10, df = 14 (P = 0.004); I² = 56%
Test for overall effect: Z = 3.39 (P = 0.0007)
6.5.2 Remote from therapist
Connelly 2006 12.2 9.92 17 10.74 11.61 20 2.8% 0.13 [-0.52, 0.78]
Connelly 2019 2.2 2.4 144 1.7 2.2 145 6.3% 0.22 [-0.01, 0.45]
Law 2015 4.83 4.78 20 4.86 4.4 37 3.5% -0.01 [-0.55, 0.54]
Lewy 2017 6.01 8.54 74 7.65 10.44 40 4.8% -0.18 [-0.56, 0.21]
Nieto 2019 5.96 6.25 25 8.22 8.61 36 3.7% -0.29 [-0.80, 0.22]
Palermo 2009 3.6 2.86 23 6.62 4.76 21 3.0% -0.76 [-1.38, -0.15]
Palermo 2016 (remote) 5.68 4.38 134 5.65 4.69 135 6.2% 0.01 [-0.23, 0.25]
Palermo 2020 34.9 25.4 73 37.8 25.6 70 5.3% -0.11 [-0.44, 0.22]
Rapoff 2014 7.82 10.59 18 12.29 12.94 17 2.7% -0.37 [-1.04, 0.30]
Van Tilburg 2009 17.1 5.1 15 25.4 10.6 14 2.2% -0.98 [-1.76, -0.20]
Subtotal (95% CI) 543 535 40.6% -0.14 [-0.33, 0.06]
Heterogeneity: Tau² = 0.04; Chi² = 18.61, df = 9 (P = 0.03); I² = 52%
Test for overall effect: Z = 1.40 (P = 0.16)
Total (95% CI) 1209 1146 100.0% -0.25 [-0.38, -0.11]
Heterogeneity: Tau² = 0.06; Chi² = 58.29, df = 24 (P = 0.0001); I² = 59%
Test for overall effect: Z = 3.51 (P = 0.0004)
Test for subgroup differences: Chi² = 1.59, df = 1 (P = 0.21), I² = 37.0%
-4 -2 0 2 4
Favours experimental Favours control
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Functional disability, follow-up *Higher scores indicate lower Disability*

Functional disability, follow-up

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
6.6.1 Face-to-face with therapist									
Grob 2013	4.22	5.26	15	24.76	14	14	2.2%	-1.91 [-2.82, -1.01]	? ? ? + +
Kashikar-Zuck 2012	13.4	8.9	57	17	10.5	55	7.3%	-0.37 [-0.74, 0.01]	+ + + + ?
Levy 2010	0.36	0.39	78	0.48	0.56	76	8.5%	-0.25 [-0.57, 0.07]	+ + + ? −
Levy 2016	5.1	6.4	67	5.9	6.8	66	8.0%	-0.12 [-0.46, 0.22]	+ ? + − −
Levy 2017	4.71	6.13	81	7.6	10.85	37	7.0%	-0.36 [-0.76, 0.03]	+ ? + ? −
Palermo 2016 (f2f)	7.84	5.5	31	8.75	4.64	30	5.2%	-0.18 [-0.68, 0.33]	+ + + + −
Powers 2013	7.6	16.9	57	19	30	67	7.6%	-0.46 [-0.81, -0.10]	+ + + + ?
Van der Veek 2013	5.8	8.2	52	4.87	6.6	52	7.1%	0.12 [-0.26, 0.51]	+ ? ? ? −
Wicksell 2009	8.8	12.9	16	14.7	12.1	16	3.2%	-0.46 [-1.16, 0.24]	+ + + ? ?
Subtotal (95% CI)			**454**			**413**	**55.9%**	**-0.33 [-0.55, -0.10]**	
Heterogeneity: Tau² = 0.06; Chi² = 19.50, df = 8 (P = 0.01); I² = 59%									
Test for overall effect: Z = 2.87 (P = 0.004)									
6.6.2 Remote from therapist									
Connelly 2019	2	2.2	144	1.9	2.2	145	10.5%	0.05 [-0.19, 0.28]	+ + + ? −
Law 2015	5.19	5.02	28	5.27	4.61	22	4.5%	-0.02 [-0.57, 0.54]	+ + + − ?
Levy 2017	4.3	7.15	70	7.6	10.85	76	8.2%	-0.35 [-0.68, -0.03]	+ ? + ? −
Palermo 2016 (remote)	5.46	4.32	134	6.16	5.05	135	10.3%	-0.15 [-0.39, 0.09]	+ + + + +
Palermo 2020	34.1	21.8	73	35.1	27.7	70	8.2%	-0.04 [-0.37, 0.29]	? + + + −
Rapoff 2014	0.91	1.45	11	3.5	4.86	11	2.3%	-0.69 [-1.56, 0.17]	? ? + − ?
Subtotal (95% CI)			**460**			**459**	**44.1%**	**-0.12 [-0.26, 0.03]**	
Heterogeneity: Tau² = 0.01; Chi² = 6.02, df = 5 (P = 0.30); I² = 17%									
Test for overall effect: Z = 1.53 (P = 0.13)									
Total (95% CI)			**914**			**872**	**100.0%**	**-0.23 [-0.38, -0.09]**	
Heterogeneity: Tau² = 0.04; Chi² = 28.80, df = 14 (P = 0.01); I² = 51%									
Test for overall effect: Z = 3.18 (P = 0.001)									
Test for subgroup differences: Chi² = 2.38, df = 1 (P = 0.12), I² = 58.0%									



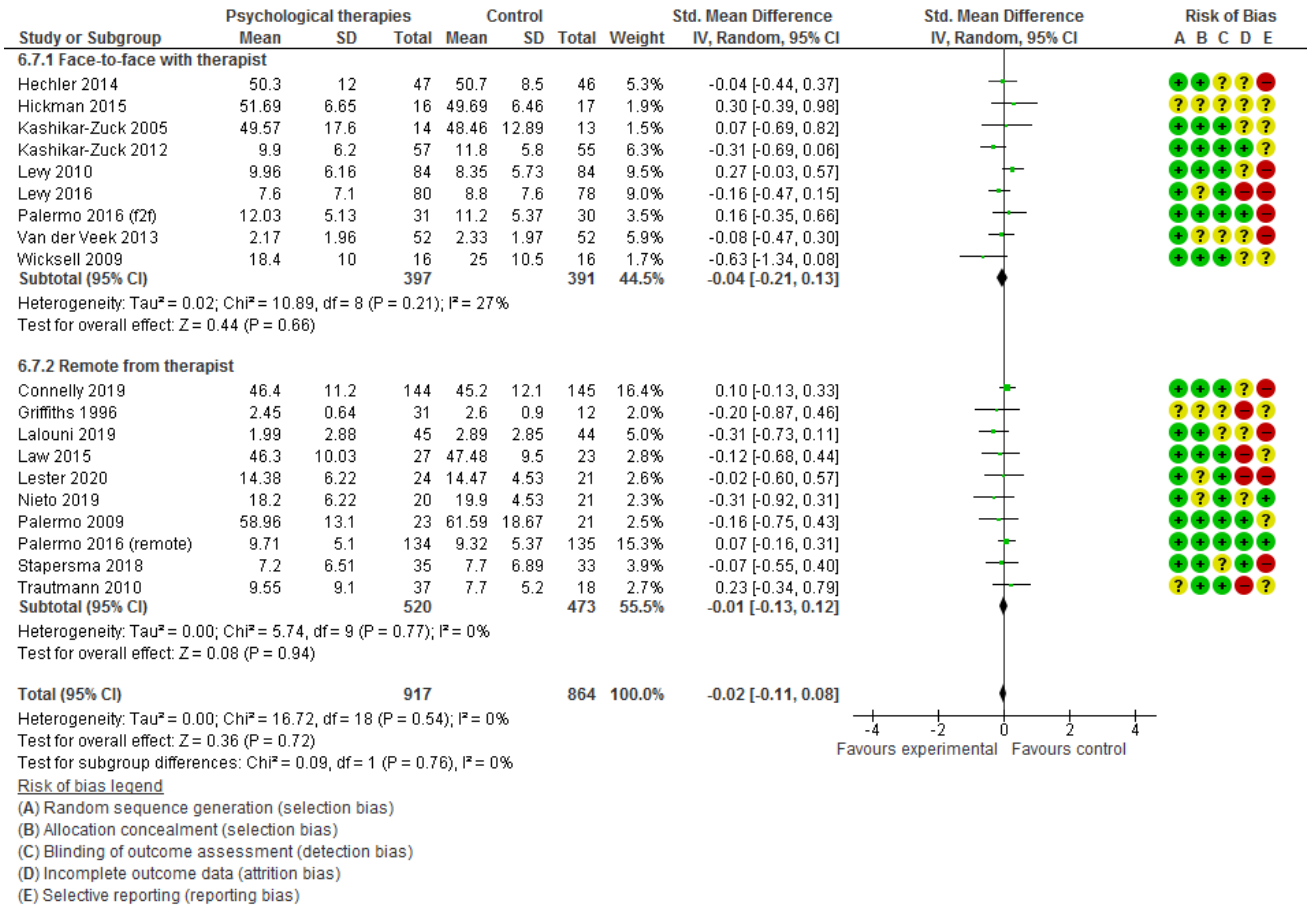
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Face to Face with therapist: ⊕⊕⊕◯ MODERATE

Remote from therapist: ⊕⊕⊕⊕ HIGH

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*

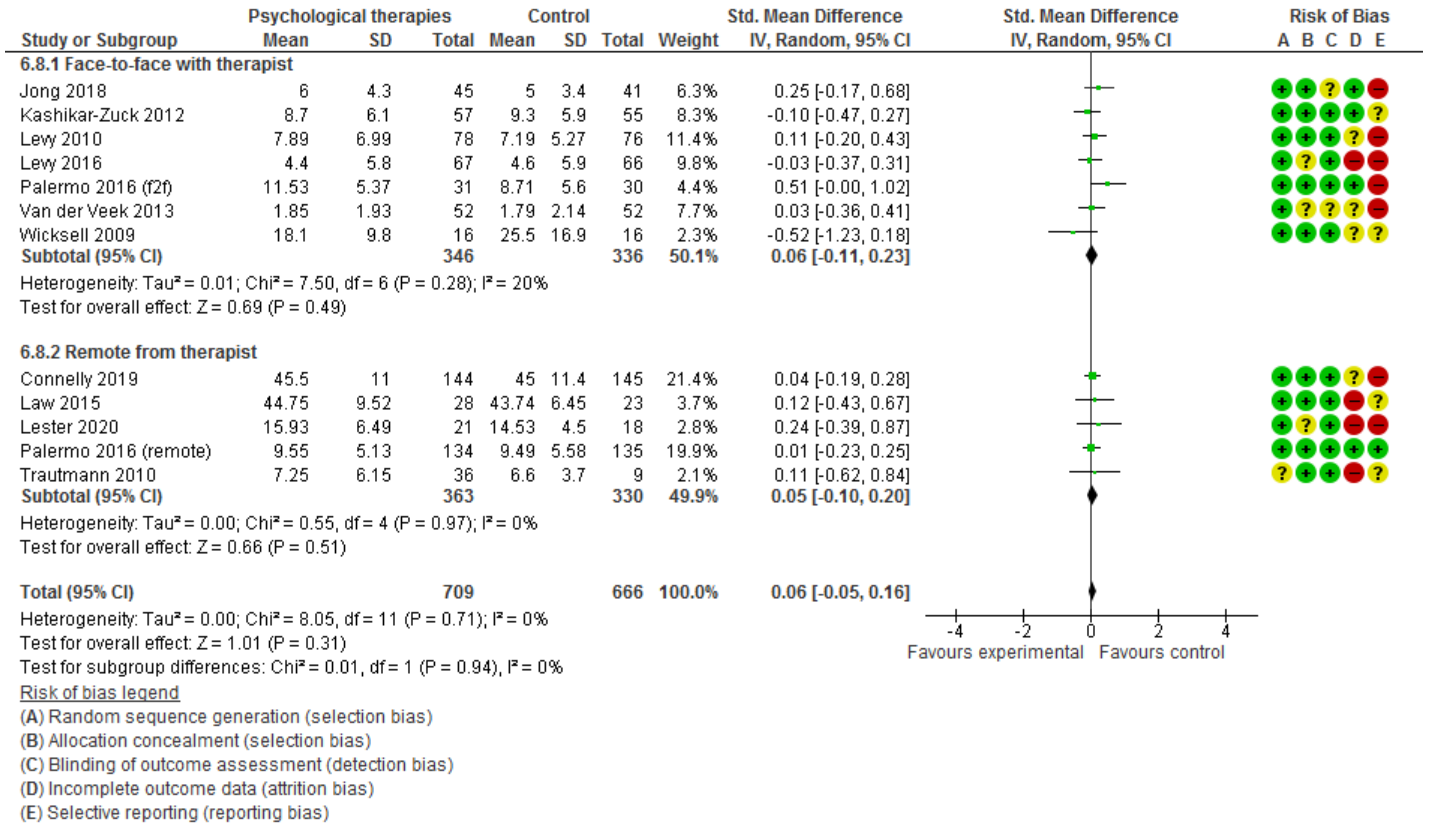
Emotional functioning: Depression, post-treatment



Face to Face with therapist: ⊕⊕⊕◯ MODERATE

Remote from therapist: ⊕⊕⊕⊕ HIGH

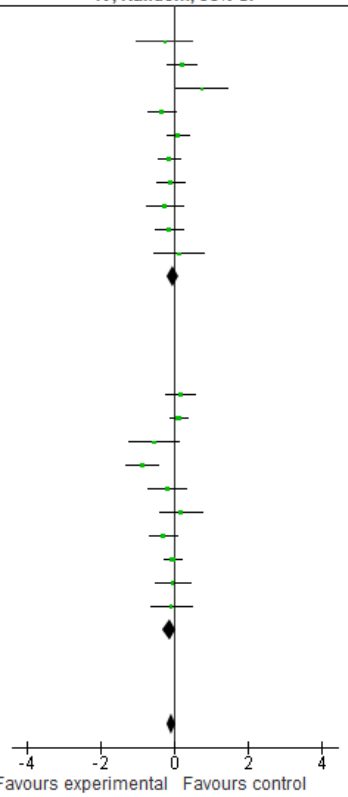
Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Depression, follow up	Face to Face with therapist: ⊕⊕⊕⊕ HIGH Remote from therapist: ⊕⊕⊕⊕ HIGH



Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher depressive symptomology*

Emotional functioning: Anxiety, post-treatment

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E
6.9.1 Face-to-face with therapist									
Bussone 1988	28.1	3.49	20	29.2	5.1	10	2.2%	-0.26 [-1.02, 0.50]	? ? ? + −
Hechler 2014	52.5	12.1	50	50	11.4	46	5.5%	0.21 [-0.19, 0.61]	+ + ? ? −
Hickman 2015	52.56	7.36	16	47.38	6.1	17	2.4%	0.75 [0.04, 1.46]	? ? ? ? ?
Kashikar-Zuck 2012	2.11	0.72	50	2.39	0.9	50	5.6%	-0.34 [-0.74, 0.05]	+ + + + ?
Lewy 2010	13.5	4.86	83	13.04	4.04	80	7.2%	0.10 [-0.21, 0.41]	+ + + ? −
Lewy 2016	8.2	2.8	80	8.6	2.9	78	7.1%	-0.14 [-0.45, 0.17]	+ ? + − −
Lewy 2017	1.18	0.95	85	1.28	1.07	41	5.9%	-0.10 [-0.47, 0.27]	+ ? + ? −
Palermo 2016 (f2f)	11.42	5.33	31	13	6.03	30	4.1%	-0.27 [-0.78, 0.23]	+ + + + −
Van der Veek 2013	6.83	6	52	7.76	6.33	52	5.7%	-0.15 [-0.53, 0.24]	+ ? ? ? −
Wicksell 2009	13.4	3.9	16	12.8	5.5	16	2.5%	0.12 [-0.57, 0.82]	+ + + ? ?
Subtotal (95% CI)			**483**			**420**	**48.3%**	**-0.05 [-0.20, 0.11]**	
Heterogeneity: Tau² = 0.01; Chi² = 11.47, df = 9 (P = 0.24); I² = 22%									
Test for overall effect: Z = 0.58 (P = 0.56)									
6.9.2 Remote from therapist									
Bonnert 2017	25.23	16.32	47	22.62	16.31	54	5.6%	0.16 [-0.23, 0.55]	+ + + + +
Connelly 2019	46.8	11.3	144	45.5	11	145	9.0%	0.12 [-0.11, 0.35]	+ + + ? −
Griffiths 1996	9.6	5.9	30	13.6	9.5	12	2.6%	-0.55 [-1.24, 0.13]	? ? ? − ?
Lalouni 2019	8.59	7.71	45	15.31	7.63	44	5.0%	-0.87 [-1.30, -0.43]	+ + ? ? −
Law 2015	46.33	8.99	30	48.32	10.81	25	3.8%	-0.20 [-0.73, 0.33]	+ + + − ?
Lester 2020	7.08	6.24	24	6.1	4.96	21	3.3%	0.17 [-0.42, 0.76]	+ ? + − −
Lewy 2017	0.99	0.93	74	1.28	1.07	40	5.7%	-0.29 [-0.68, 0.09]	+ ? + ? −
Palermo 2016 (remote)	10.56	5.91	134	10.85	6.1	135	8.8%	-0.05 [-0.29, 0.19]	+ + + + +
Stapersma 2018	7.1	4.14	35	7.3	4.6	33	4.4%	-0.05 [-0.52, 0.43]	+ + ? + −
Trautmann 2010	30.9	7.95	38	31.7	8.3	18	3.5%	-0.10 [-0.66, 0.46]	? + + − ?
Subtotal (95% CI)			**601**			**527**	**51.7%**	**-0.14 [-0.34, 0.06]**	
Heterogeneity: Tau² = 0.05; Chi² = 20.81, df = 9 (P = 0.01); I² = 57%									
Test for overall effect: Z = 1.38 (P = 0.17)									
Total (95% CI)			**1084**			**947**	**100.0%**	**-0.09 [-0.21, 0.04]**	
Heterogeneity: Tau² = 0.03; Chi² = 32.45, df = 19 (P = 0.03); I² = 41%									
Test for overall effect: Z = 1.38 (P = 0.17)									
Test for subgroup differences: Chi² = 0.54, df = 1 (P = 0.46), I² = 0%									

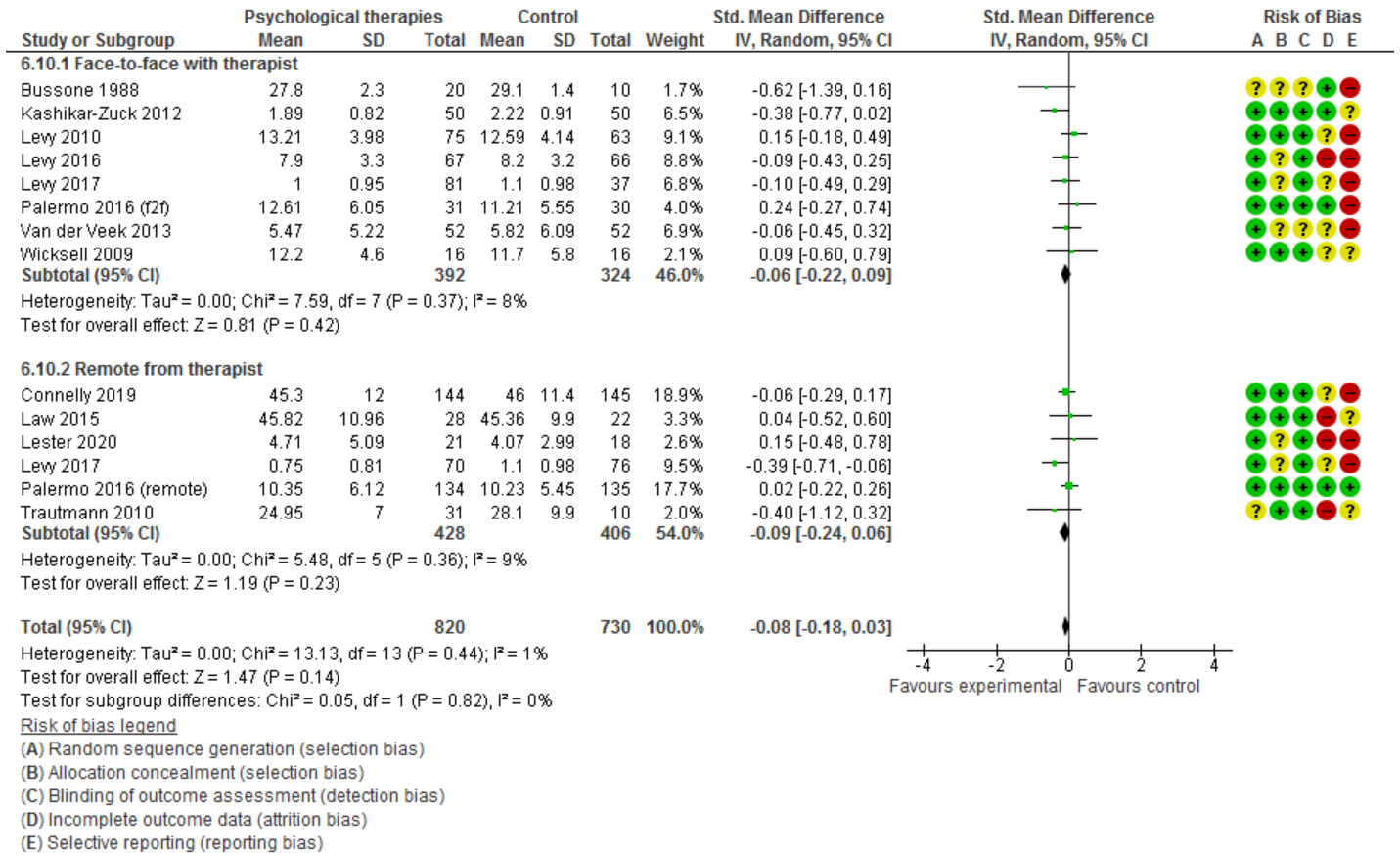


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Face to Face with therapist: ⊕⊕⊕◯ MODERATE

Remote from therapist: ⊕⊕⊕◯ MODERATE

Emotional functioning: Anxiety, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Anxiety, follow-up	Face to Face with therapist: ⊕⊕⊕◯ MODERATE Remote from therapist: ⊕⊕⊕⊕ HIGH



Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
6.10.1 Face-to-face with therapist								
Bussone 1988	27.8	2.3	20	29.1	1.4	10	1.7%	-0.62 [-1.39, 0.16]
Kashikar-Zuck 2012	1.89	0.82	50	2.22	0.91	50	6.5%	-0.38 [-0.77, 0.02]
Levy 2010	13.21	3.98	75	12.59	4.14	63	9.1%	0.15 [-0.18, 0.49]
Levy 2016	7.9	3.3	67	8.2	3.2	66	8.8%	-0.09 [-0.43, 0.25]
Levy 2017	1	0.95	81	1.1	0.98	37	6.8%	-0.10 [-0.49, 0.29]
Palermo 2016 (f2f)	12.61	6.05	31	11.21	5.55	30	4.0%	0.24 [-0.27, 0.74]
Van der Veek 2013	5.47	5.22	52	5.82	6.09	52	6.9%	-0.06 [-0.45, 0.32]
Wicksell 2009	12.2	4.6	16	11.7	5.8	16	2.1%	0.09 [-0.60, 0.79]
Subtotal (95% CI)			**392**			**324**	**46.0%**	**-0.06 [-0.22, 0.09]**
Heterogeneity: Tau² = 0.00; Chi² = 7.59, df = 7 (P = 0.37); I² = 8%								
Test for overall effect: Z = 0.81 (P = 0.42)								
6.10.2 Remote from therapist								
Connelly 2019	45.3	12	144	46	11.4	145	18.9%	-0.06 [-0.29, 0.17]
Law 2015	45.82	10.96	28	45.36	9.9	22	3.3%	0.04 [-0.52, 0.60]
Lester 2020	4.71	5.09	21	4.07	2.99	18	2.6%	0.15 [-0.48, 0.78]
Levy 2017	0.75	0.81	70	1.1	0.98	76	9.5%	-0.39 [-0.71, -0.06]
Palermo 2016 (remote)	10.35	6.12	134	10.23	5.45	135	17.7%	0.02 [-0.22, 0.26]
Trautmann 2010	24.95	7	31	28.1	9.9	10	2.0%	-0.40 [-1.12, 0.32]
Subtotal (95% CI)			**428**			**406**	**54.0%**	**-0.09 [-0.24, 0.06]**
Heterogeneity: Tau² = 0.00; Chi² = 5.48, df = 5 (P = 0.36); I² = 9%								
Test for overall effect: Z = 1.19 (P = 0.23)								
Total (95% CI)			**820**			**730**	**100.0%**	**-0.08 [-0.18, 0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 13.13, df = 13 (P = 0.44); I² = 1%								
Test for overall effect: Z = 1.47 (P = 0.14)								
Test for subgroup differences: Chi² = 0.05, df = 1 (P = 0.82), I² = 0%								

Std. Mean Difference IV, Random, 95% CI

-4 -2 0 2 4

Favours experimental Favours control

Risk of Bias A B C D E

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of outcome assessment (detection bias)

(D) Incomplete outcome data (attrition bias)

(E) Selective reporting (reporting bias)

Appendix G. 7. WHO review: Psychological interventions for children with chronic pain
Subgroup analysis: by therapy type

Comparison: Psychological therapies versus active (non-psychological), standard care or waitlist control, by therapy type
Population: children and adolescents with chronic pain
Setting: Any setting
Studies: Randomised controlled trials

Subgroup analysis: by therapy classification

We analysed studies by the type of therapy they delivered, using classifications of cognitive behavioural therapy, acceptance commitment therapy, hypnosis, and relaxation. Due to the small number of studies using certain types of therapy types, we could not conduct meta-analyses for all therapy types. CBT was the most commonly delivered therapy type. We could not draw any conclusions for ACT, hypnosis, or problem-solving therapy as only one study could be included in any given analysis. More evidence was available for relaxation training and behaviour therapy, although very low-certainty, mainly due to imprecision and the small number of participants that could be included in the subgroup analyses. We have excluded all analyses from the GRADE profiles (Table 12) where only one study or less could be included in the analysis, but these can still be found in the forest plots (Appendix F.5). All certainty of evidence for single study analyses was very low, downgraded twice for imprecision and once for indirectness.

We found small beneficial effects for CBT on the following outcomes; pain intensity post-treatment (low-certainty), 50% pain reduction post-treatment (low-certainty), functional disability post-treatment (low-certainty) and at follow-up (moderate-certainty). We did not find beneficial effects of CBT for pain at follow-up (low-certainty) and emotional functioning (depression: moderate-certainty post-treatment, high-certainty follow-up; anxiety: low-certainty post-treatment, moderate-certainty follow-up). We could analyse relaxation training for pain intensity post-treatment and at follow-up, and 50% reduction post-treatment; behaviour therapy for pain intensity post-treatment and 50% reduction in pain post-treatment. We did not find any benefits of these therapies on the outcomes. For the remaining therapy types and outcomes, we could only include a single studies in the analyses and therefore cannot draw any conclusions; we rated all evidence as very low-certainty, primarily due to imprecision and indirectness.

Risk of bias legend

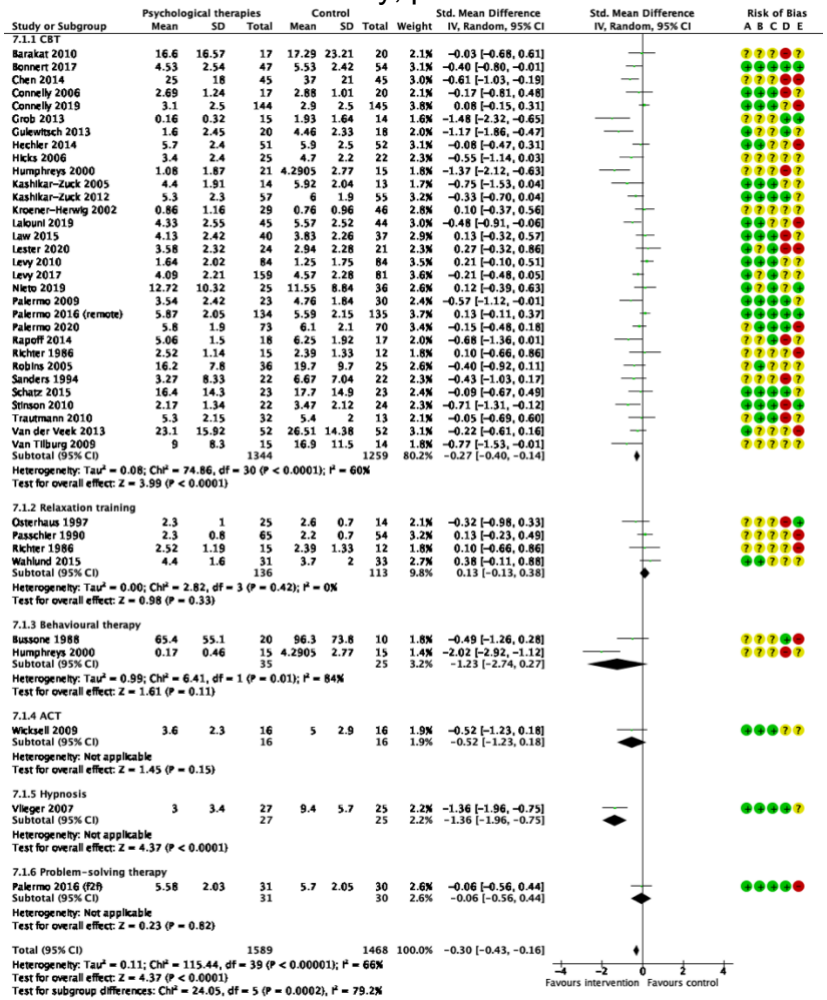
(**A**) Random sequence generation (selection bias)

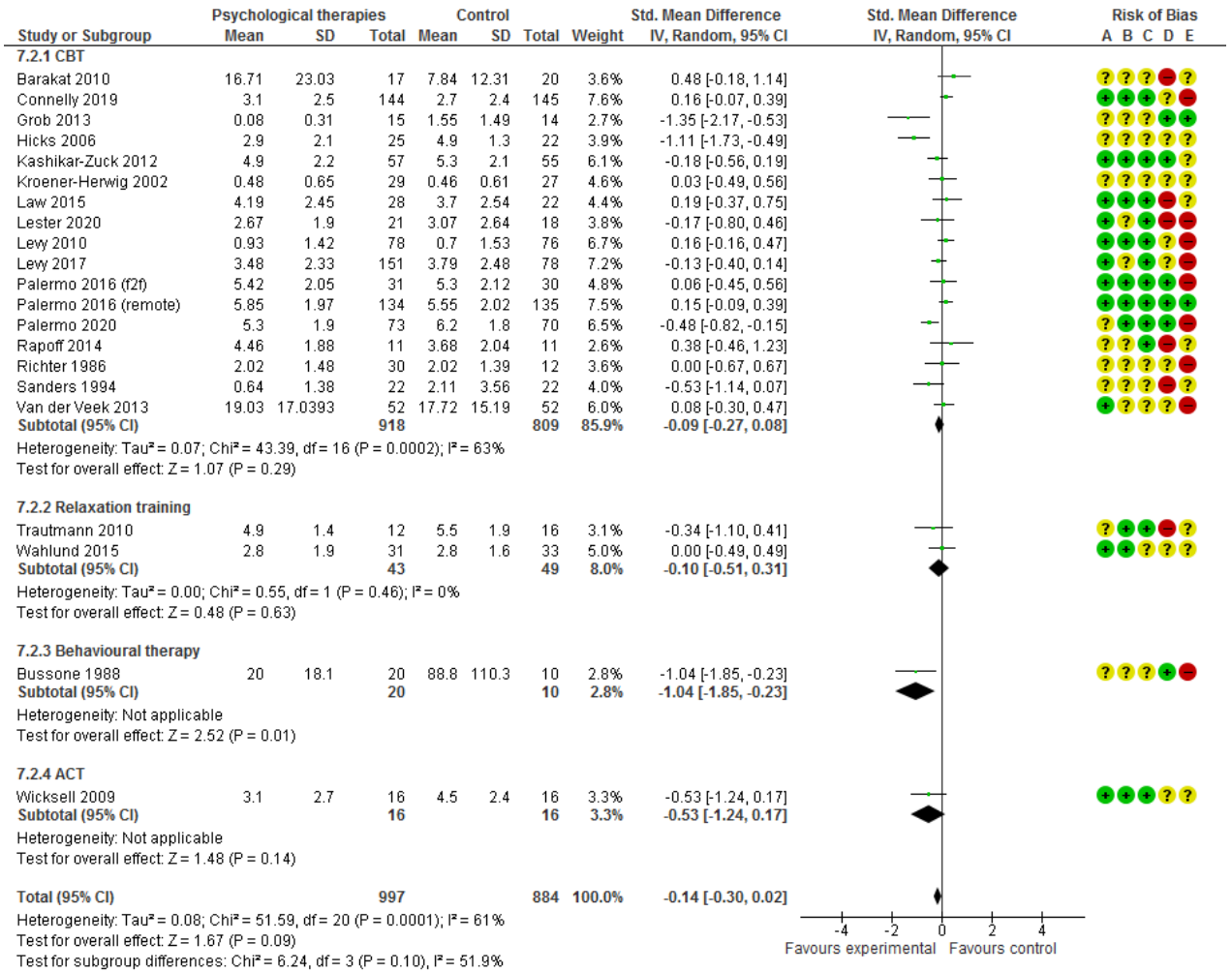
(**B**) Allocation concealment (selection bias)

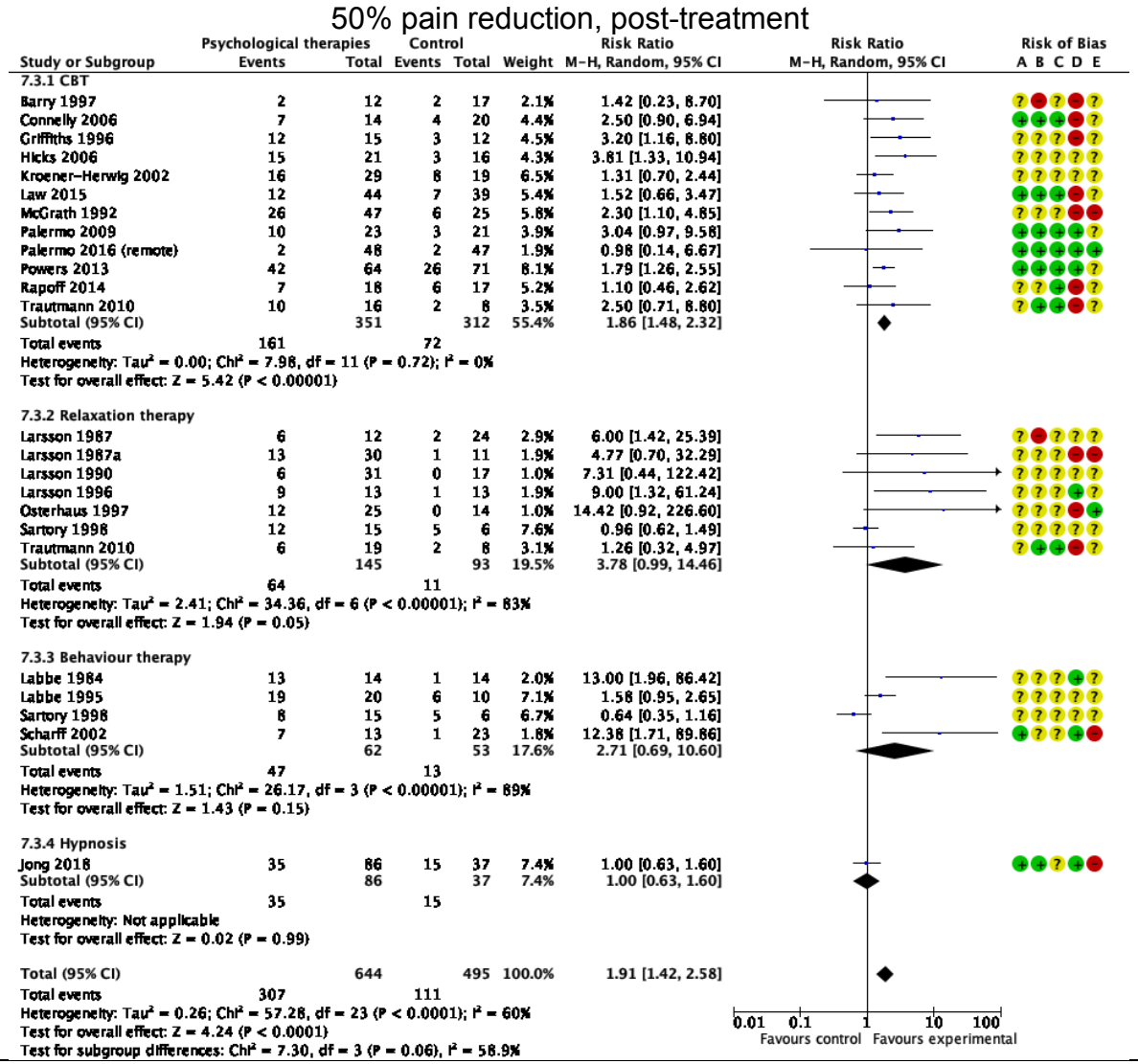
(**C**) Blinding of outcome assessment (detection bias)

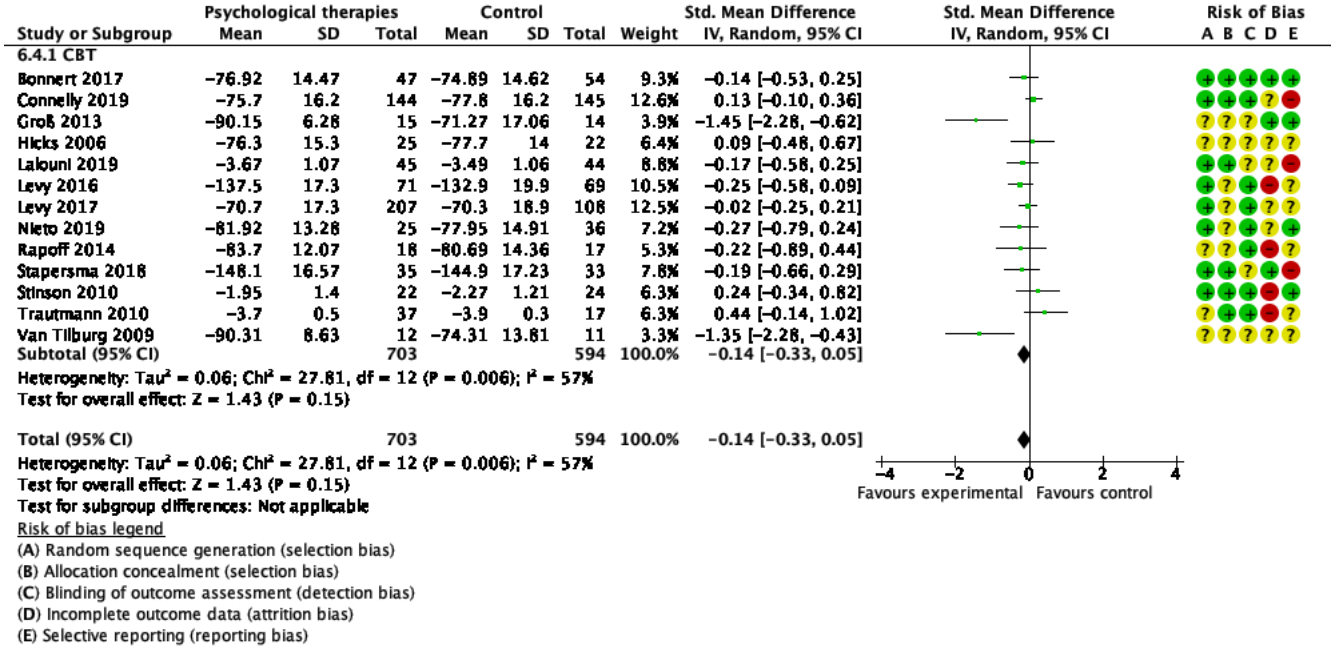
(**D**) Incomplete outcome data (attrition bias)

(**E**) Selective reporting (reporting bias)

Outcome	Forest plot	Quality of evidence (GRADE)
Pain intensity, post-treatment *Higher scores indicate higher pain intensity*	Pain intensity, post-treatment 	CBT: ⊕⊕◯◯ LOW Relaxation training: ⊕◯◯◯ VERY LOW Behavioural Therapy: ⊕◯◯◯ VERY LOW

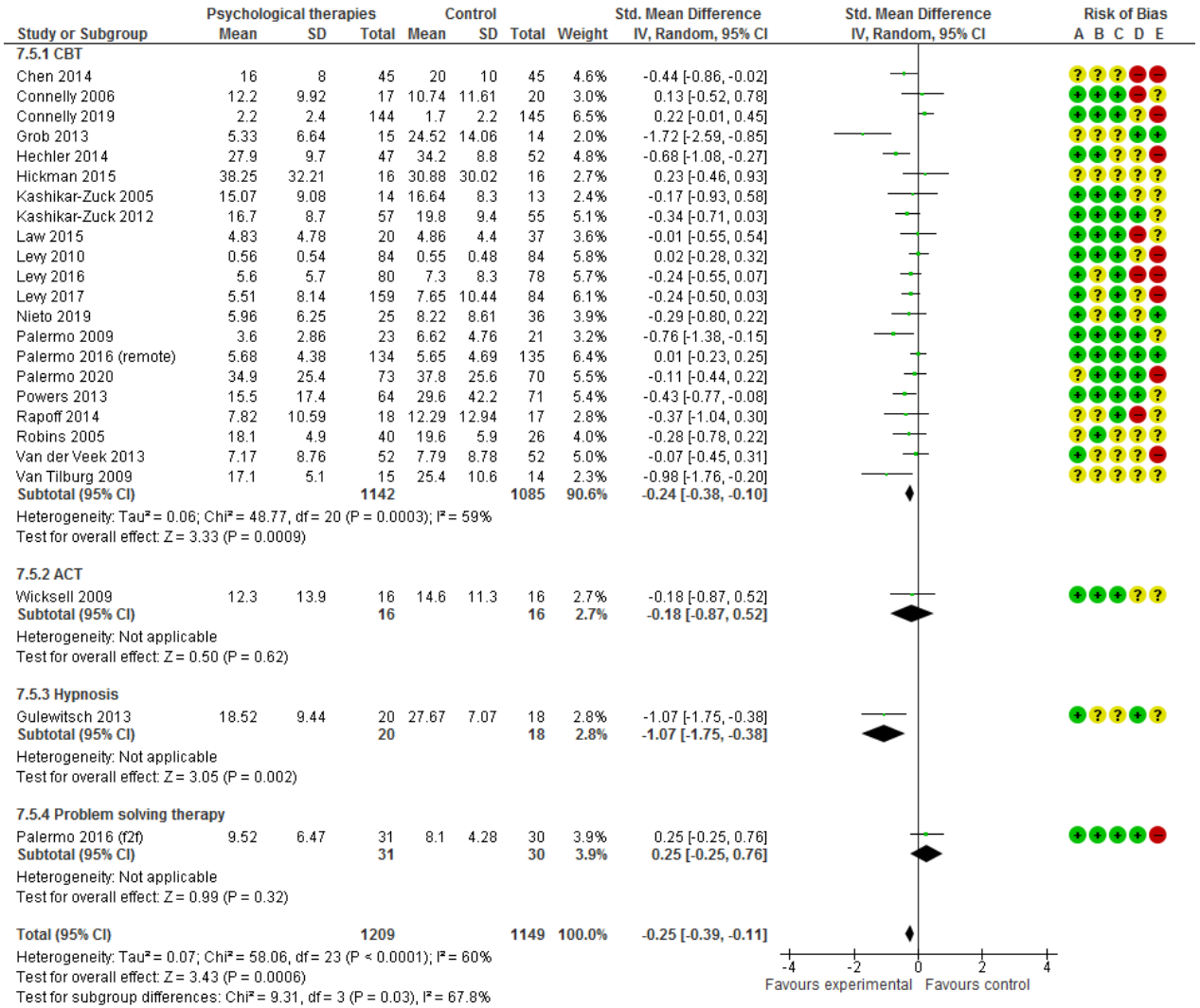
Pain intensity, follow-up *Higher scores indicate higher pain intensity*	Pain intensity, follow-up 	CBT: ⊕⊕◯◯ LOW Relaxation training: ⊕◯◯◯ VERY LOW Behavioural Therapy: ⊕◯◯◯ VERY LOW ACT: ⊕◯◯◯ VERY LOW

50% reduction, post-treatment		CBT: ⊕⊕◯◯ LOW Relaxation training: ⊕◯◯◯ VERY LOW Behavioural Therapy: ⊕◯◯◯ VERY LOW Hypnosis: ⊕◯◯◯ VERY LOW

Health-related quality of life, post-treatment	Health-related quality of life, post-treatment  	CBT: ⊕⊕◯◯ LOW

Functional disability, post-treatment *Higher scores indicate lower disability*

Functional disability, post-treatment



CBT:
⊕⊕◯◯
LOW

ACT:
⊕◯◯◯
VERY LOW

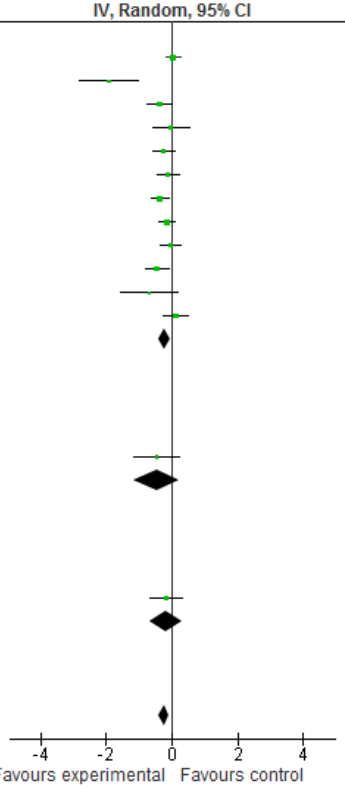
Hypnosis:
⊕◯◯◯
VERY LOW

Problem-solving therapy:
⊕◯◯◯
VERY LOW

Functional disability, follow-up *Higher scores indicate lower disability*

Functional disability, follow-up

Study or Subgroup	Psychological therapies Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
7.6.1 CBT								
Connelly 2019	2	2.2	144	1.9	2.2	145	10.9%	0.05 [-0.19, 0.28]
Grob 2013	4.22	5.26	15	24.76	14	14	2.4%	-1.91 [-2.82, -1.01]
Kashikar-Zuck 2012	13.4	8.9	57	17	10.5	55	7.8%	-0.37 [-0.74, 0.01]
Law 2015	5.19	5.02	28	5.27	4.61	22	4.9%	-0.02 [-0.57, 0.54]
Lewy 2010	0.36	0.39	78	0.48	0.56	76	8.9%	-0.25 [-0.57, 0.07]
Lewy 2016	5.1	6.4	67	5.9	6.8	66	8.4%	-0.12 [-0.46, 0.22]
Lewy 2017	4.5	6.6	151	7.6	10.85	82	10.0%	-0.37 [-0.64, -0.10]
Palermo 2016 (remote)	5.46	4.32	134	6.16	5.05	135	10.7%	-0.15 [-0.39, 0.09]
Palermo 2020	34.1	21.8	73	35.1	27.7	70	8.7%	-0.04 [-0.37, 0.29]
Powers 2013	7.6	16.9	57	19	30	67	8.1%	-0.46 [-0.81, -0.10]
Rapoff 2014	0.91	1.45	11	3.5	4.86	11	2.5%	-0.69 [-1.56, 0.17]
Van der Veek 2013	5.8	8.2	52	4.87	6.6	52	7.5%	0.12 [-0.26, 0.51]
Subtotal (95% CI)			**867**			**795**	**90.8%**	**-0.23 [-0.40, -0.06]**
Heterogeneity: Tau² = 0.05; Chi² = 28.20, df = 11 (P = 0.003); I² = 61%								
Test for overall effect: Z = 2.68 (P = 0.007)								
7.6.2 ACT								
Wicksell 2009	8.8	12.9	16	14.7	12.1	16	3.5%	-0.46 [-1.16, 0.24]
Subtotal (95% CI)			**16**			**16**	**3.5%**	**-0.46 [-1.16, 0.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.28 (P = 0.20)								
7.6.3 Problem solving therapy								
Palermo 2016 (f2f)	7.84	5.5	31	8.75	4.64	30	5.6%	-0.18 [-0.68, 0.33]
Subtotal (95% CI)			**31**			**30**	**5.6%**	**-0.18 [-0.68, 0.33]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.69 (P = 0.49)								
Total (95% CI)			**914**			**841**	**100.0%**	**-0.23 [-0.38, -0.08]**
Heterogeneity: Tau² = 0.04; Chi² = 28.79, df = 13 (P = 0.007); I² = 55%								
Test for overall effect: Z = 2.96 (P = 0.003)								
Test for subgroup differences: Chi² = 0.45, df = 2 (P = 0.80), I² = 0%								



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

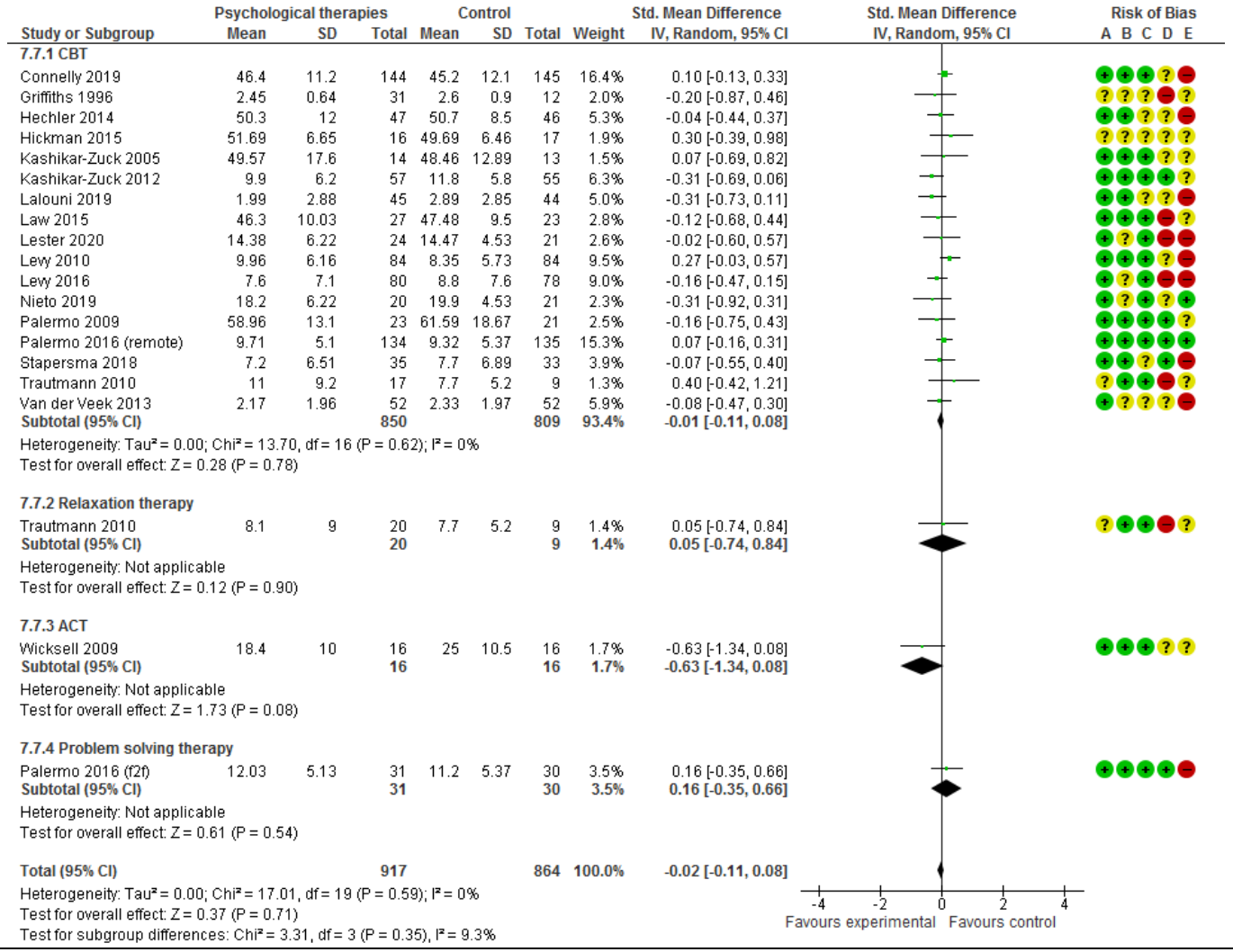
CBT: ⊕⊕⊕◯ MODERATE

ACT: ⊕◯◯◯ VERY LOW

Problem-solving therapy: ⊕◯◯◯ VERY LOW

Emotional functioning: Depression, post-treatment *Higher scores indicate higher depressive symptomology*

Emotional functioning: Depression, post-treatment



CBT:
⊕⊕⊕◯
MODERATE

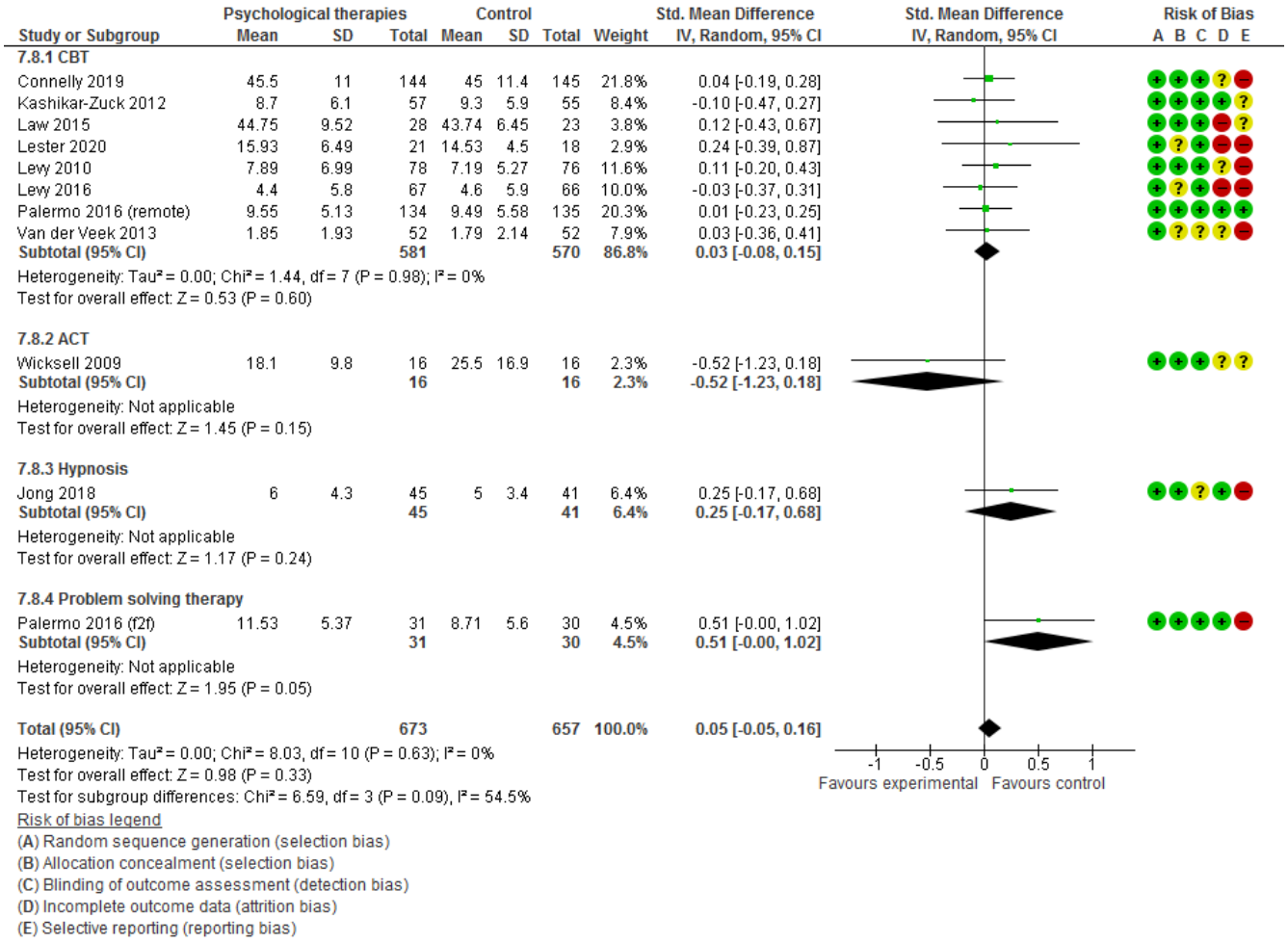
Relaxation therapy:
⊕◯◯◯
VERY LOW

ACT:
⊕◯◯◯
VERY LOW

Problem-solving therapy:
⊕◯◯◯
VERY LOW

Emotional functioning: Depression, follow up *Higher scores indicate higher depressive symptomology*

Emotional functioning: Depression, follow up



CBT: ⊕⊕⊕⊕ HIGH

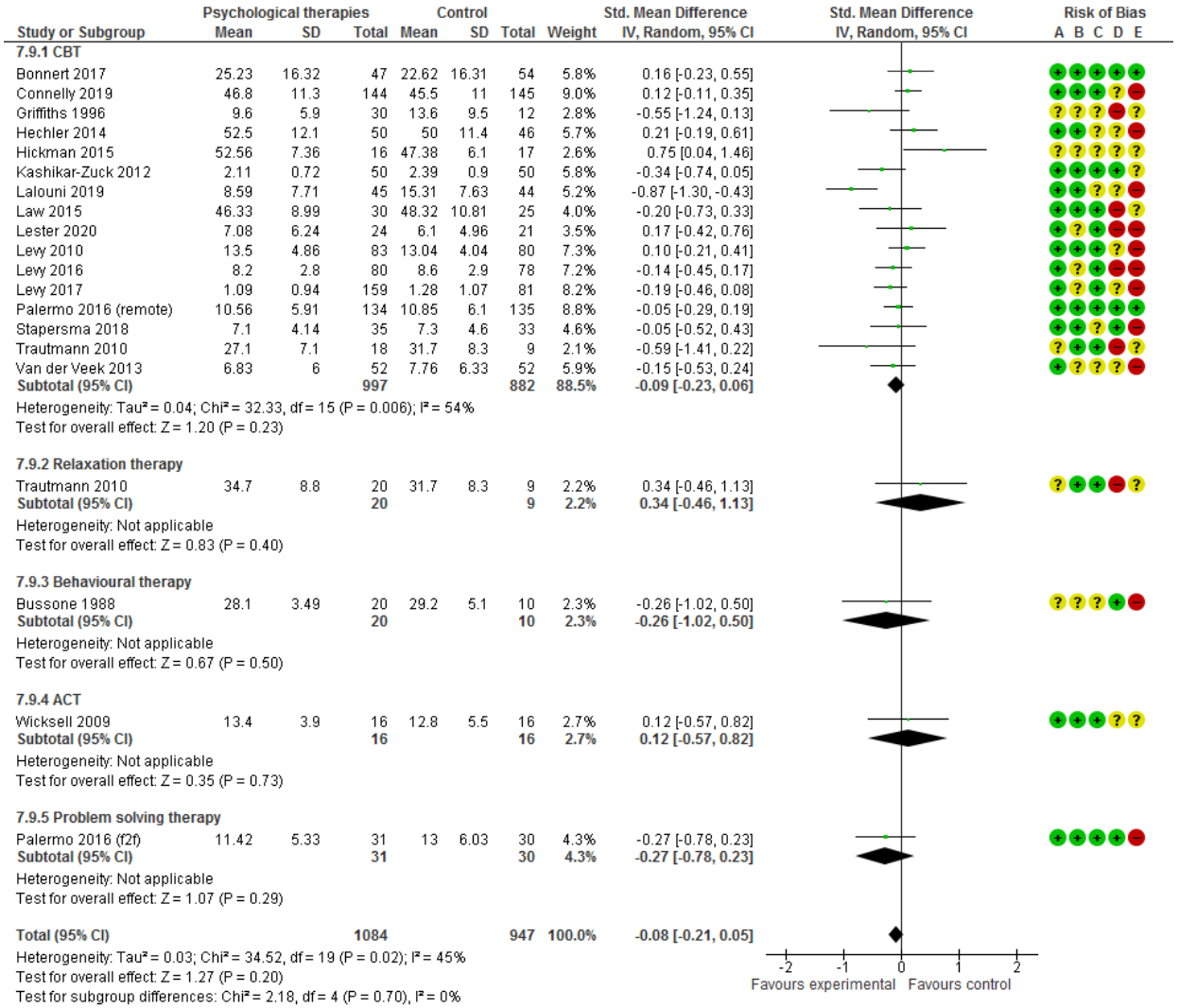
ACT: ⊕◯◯◯ VERY LOW

Hypnosis: ⊕◯◯◯ VERY LOW

Problem-solving therapy: ⊕◯◯◯ VERY LOW

Emotional functioning: Anxiety, post-treatment *Higher scores indicate higher depressive symptomology*

Emotional functioning: Anxiety, post-treatment



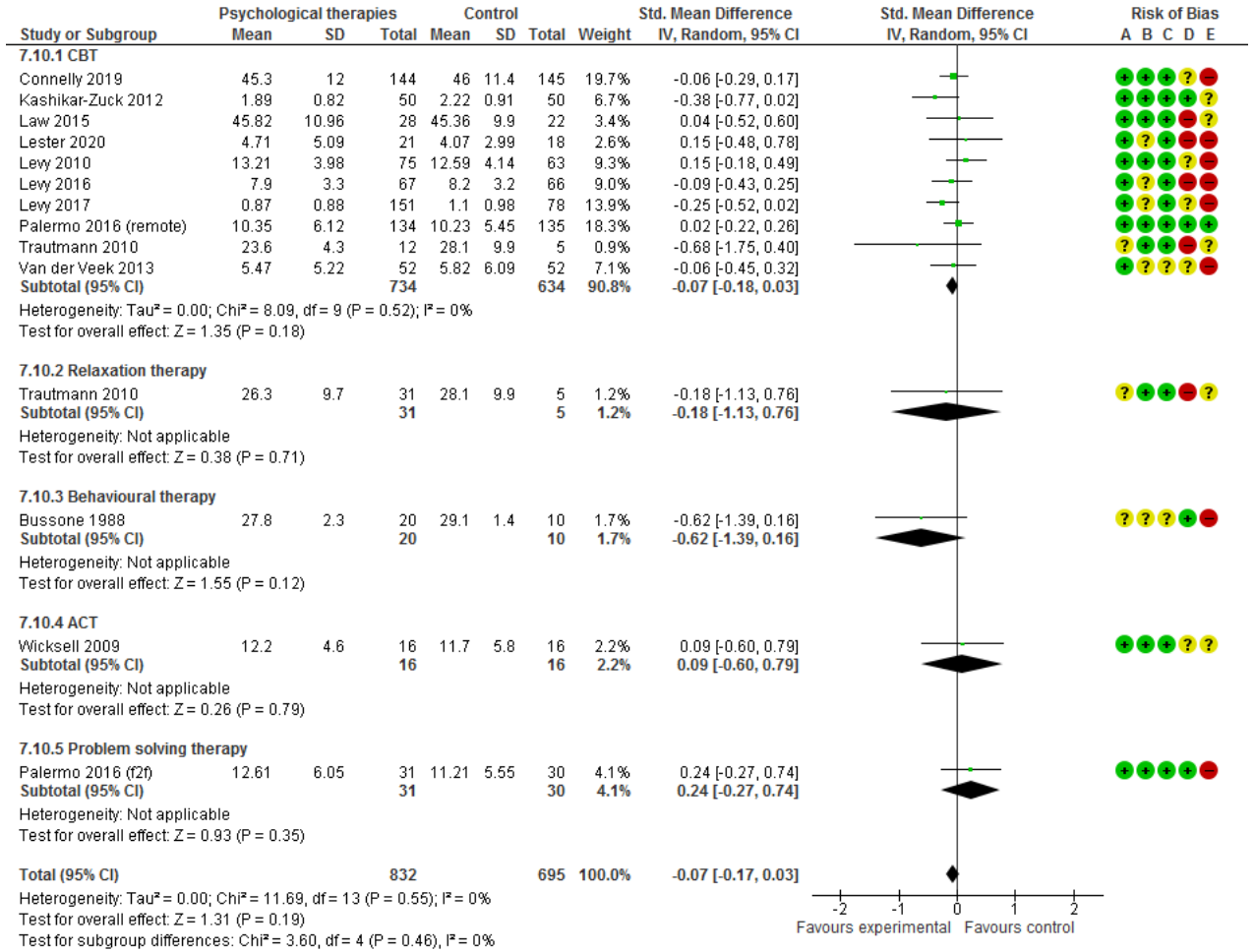
CBT:
⊕⊕◯◯
LOW

Relaxation therapy:
⊕◯◯◯
VERY LOW

Behavioural therapy:
⊕◯◯◯
VERY LOW

ACT:
⊕◯◯◯
VERY LOW

Problem-solving therapy:
⊕◯◯◯
VERY LOW

Emotional functioning: Anxiety, follow-up *Higher scores indicate higher depressive symptomology*	Emotional functioning: Anxiety, follow-up 	CBT: ⊕⊕⊕◯ MODERATE Relaxation therapy: ⊕◯◯◯ VERY LOW Behavioural therapy: ⊕◯◯◯ VERY LOW ACT: ⊕◯◯◯ VERY LOW Problem-solving therapy: ⊕◯◯◯ VERY LOW