Individualised exercise in chronic non-specific low back pain: a systematic review with meta-analysis on the effects of exercise alone or in combination with psychological interventions on pain and disability

Authors:
Johannes Fleckenstein[1] (0000-0003-3442-0380), Philipp Flössel[2], Tilman Engel[3], Laura Klewinghaus[4], Josefine Stoll[3], Martin Behrens[5,6], Daniel Niederer[1]

Affiliations
[1]Goethe-University Frankfurt, Institute of Sports Sciences, Department of Sports Medicine and Exercise Physiology, Ginnheimer Landstr. 39, 60596 Frankfurt am Main
[2]University Center for Orthopaedic and Trauma Surgery, University Hospital Carl Gustav Carus, Fetscherstr. 74, 01307 Dresden, Germany
[3]University of Potsdam, University Outpatient Clinic, Sports Medicine and Sports Orthopedics, Am Neuen Palais 10, 14469 Potsdam, Germany
[4]University of Wuppertal, Department of Clinical Psychology and Psychotherapy, Max-Horkheimer-Str. 20, 42119 Wuppertal, Germany
[5]Department of Sport Science, Institute III, Otto-von-Guericke University Magdeburg, Magdeburg 39104, Germany
[6]Department of Orthopaedics, University Medicine Rostock, Doberaner Str. 142, 18057 Rostock, Germany

Correspondence to:
Johannes Fleckenstein, Goethe-University Frankfurt, Institute of Sports Sciences, Department of Sports Medicine and Exercise Physiology, Ginnheimer Landstr. 39, 60487 Frankfurt am Main, Germany, phone: +49 (0)69 798 24484, mail: johannes.fleckenstein@sport.uni-frankfurt.de

Contributor and guarantor information

All authors contributed to the planning of the study. JF, PF and DN conducted the study, JF wrote and DN analysed data, JF wrote the first draft of the manuscript. All authors read, revised and approved the manuscript.

JF accepts full responsibility for the work and/or the conduct of the study, has access to the data, and controlled the decision to publish.

The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Copyright/license for publication

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above."

Competing interests' declaration

JF received honorary for scientific counselling from the German Medical Acupuncture Association. All other authors declare no conflict of interest.

ABSTRACT

Objective To investigate the effects of individualised exercise interventions consisting with or without combined psychological intervention on pain intensity and disability in patients with chronic non-specific low back pain.

Design Systematic review with meta-analysis and meta-regression.

Data sources Five databases (PubMed, Cochrane Central, EMBASE, Clarivate Web of Science, and Google Scholar) were searched up to 31 March 2021.

Selection criteria Randomised controlled trials were eligible if they included participants with chronic non-specific low back pain, compared at least one individualised/personalised/stratified exercise intervention with or without psychological treatment to any control / comparator group, and if they assessed at least pain intensity or disability as outcome measure.

Results Fifty-eight studies (n = 9099 patients, 44.3 years, 56% female) compared individualised to other types of exercise (n = 44; 62%), usual care (n = 16; 23%), advice to stay active, or true controls. The remaining studies had passive controls.

At short-term follow-up, low-certainty evidence for pain intensity (SMD -0.33 [95%CI -0.47 to -0.18]) and very low-certainty evidence for disability (-0.16 [-0.30 to -0.02]) indicates effects of individualised exercise compared to other exercises. Very low-certainty evidence for pain intensity (-0.35; [-0.53 to -0.17])) and low-certainty evidence for disability (-0.12; [-0.22 to -0.02]) indicates effects compared to passive controls.

At long-term follow-up, moderate-certainty evidence for pain intensity (-0.14 [-0.23 to -0.06]) and disability (-0.23 [-0.33 to -0.12]) indicates effects compared to passive controls exercises.

All findings stayed below the threshold for minimal clinically important difference (MCID). Certainty of evidence was downgraded mainly due to evidence of risk of bias, publication bias and inconsistency that could not be explained. Sensitivity analyses indicated that the effects on pain, but not on disability (always short-term and versus active treatments) were robust. Sub-group analysis of pain outcomes suggested that individualised exercise treatment is probably more effective in combination with psychological interventions (-0.32 [-0.51 to -0.14]), a clinically important difference

Conclusion We found very low to moderate-certainty evidence that individualised exercise is effective for treatment of chronic non-specific low back pain. Individualised exercise seems superior to other active treatments and sub-group analysis suggests that some forms of individualised exercise (especially motor-control based treatments) combined with behavioural therapy interventions enhances the treatment effect. Certainty of evidence was higher for long-term follow-up. In summary, individualised exercise can be recommended from a clinical point of view.

Introduction

Back pain, in particular chronic non-specific low back pain, is one of the most serious public-health problems worldwide.[1] Its global median 1-year period prevalence in the adult population is around 37%.[1] Together with the consequences of low back pain, such as loss of labour productivity or rates of hospital admissions, this high number of involved individuals results in tremendous direct and indirect costs for societies' healthcare systems and economies.[2] For those affected, pain symptoms are often accompanied by functional limitations and decreased performance, substantially deteriorating the quality of their individual and family lives.[3]

Exercise therapy and psychological (especially cognitive behavioural) therapy, both as stand-alone treatments but especially in combination, are considered being effective in the treatment of chronic low back pain.[4] Still, current research reveals only relatively small and heterogeneous effects when combining exercise and cognitive-behavioural therapies.[5-7] This is likely to be attributed to inter-individual heterogeneities in the effects.[8] Stratified or even individualised/personalised therapies may be the potential key to increase the (group and individual) therapeutic effect.[9] On a somatic-biological level, patients with low back pain predominantly present deficits in motor control strategies, though varying in its characteristics between individuals.[10 11] On a psychosocial level, dysfunctional cognition (fear-avoidance-beliefs, catastrophizing) and dysfunctional pain behaviour (kinesiophobia) are known risk factors for chronification or for a worsening of symptoms in chronic low back pain.[12 13]

Functional strength/restoration, coordination (Pilates, McKenzie) and stabilization training are, based on meta-analytic evidence on comparative effectiveness, likely to be the most effective types of exercise in the management of chronic non-specific low back pain.[9 14] Simply and mechanistically expressed, all these training modalities improve motor control.[11] Unfortunately, such effects are only of limited validity for the affected individuals due to the presence of heterogeneous influencing psychosocial factors in group analyses.[14] The latter has been addressed by a variety of validated individualised psychological interventions against chronic low back pain. A prominent approach is the STarT Back Screening Tool, used to classify patients with low back pain into tailored focus therapy groups (exercises and cognitive-behavioural strategies) based on bio-psycho-social factors.[15] Stratification, or even better personalisation/individualisation of exercise and psychological interventions may consequently lead to better therapeutic outcomes (fewer days of illness, improvement of symptomatology) than mono-interventions or, even non-personalised therapies only.[16 17]

Consequently, individualised exercise with or without psychological interventions were investigated by a large number of clinical trials. However, individualisation as an independent variable defining exercise has not been assessed on a meta-analytical level before. There is one recently published review analysing the effects of classification-based exercise to treat chronic low back pain,[18] showing small and no meaningful effects. However, this study focussed on classification systems and not on exercise that was individualised based on the patient's needs. In addition, exercise was often combined with psychological interventions, which may have impacted the overall outcomes. Consequently, the aim of this systematic review with meta-analysis and –regression was to qualify and quantify the effects of individualised exercises as mono-therapies, or in combination with cognitive or

behavioural types of psychological interventions. We will distinguish active and passive controls, and differentiate the type of individualised exercise. This will consolidate previous knowledge and will build a bridge to known effects such as the above mentioned classification systems [18] or classical exercise.[6]

Methods

The systematic review with meta-analysis and meta-regression protocol was developed using guidance from the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) statement,[19] registered in the PROSPERO database (CRD42021247331, Date of registration in PROSPERO: 10 May 2021, Date of first submission: 09 April 2021, available at: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=247331).

Studies were identified by searching multiple databases, including PubMed, Cochrane Central Register of Controlled Trials, EMBASE, Clarivate Web of Science, and Google Scholar from their inception to March 2021. The search terms were identified after preliminary searches of the literature and by comparing them against previous systematic reviews. A sensitive search strategy was used based on the recommendations of the Cochrane Back and Neck Group for 'randomised controlled trials' (RCT) and 'low back pain',[20] combined with search terms for 'exercise'. The full search strategy is presented in the supplementary data appendix B. In addition, reference lists of relevant reviews and included RCTs were manually searched, and citation tracking of all included trials was performed. A grey literature search in google scholar was further performed. The searches and inclusion criteria were restricted to English, German, French, Spanish, and Italian.

We tested the eligibility criteria by piloting a small sample (10 trials). Two independent researchers (JF, PF) screened titles and abstracts of the publications retrieved by the search strategy and assessed the full texts for potential inclusion. Studies not meeting the inclusion criteria were discarded. Disagreements between researchers were resolved by discussion and consultation with a third reviewer (DN), if necessary.

For the sake of brevity, the term (chronic) low back pain in this manuscript refers to chronic non-specific low back pain if not stated otherwise.

Eligibility criteria:

This study followed the participants, interventions, comparisons, outcomes, and study design (PICOS) framework.

Population

The population of interest was patients suffering from chronic non-specific low back pain that have been treated with any approach of stratified, personalised and/or individualised exercise. Chronic non-specific low back pain is defined as pain without a known pathoanatomical cause [21], chronicity of pain lasting longer than 3 month, or recurrent low back pain defined as two episodes per year.[6] All studies had to research populations were the aforementioned criteria were fulfilled for the majority of their included patients to be considered potentially eligible.

Intervention (exposure):

The interventions/exposures were all forms of individualised, tailored, personalised, or stratified exercise of any frequency, intensity, volume or type, with a duration of at least 4 weeks. Interventions that consisted of exercise alone or exercise combined with psychological interventions (i.e. behavioural or cognitive, or combined behavioural-cognitive therapy) were considered.

Comparison:

Eligible comparators were all matched or non-matched conservative (non-surgery) control/comparator groups, that could be either active, passive, or no treatment (real) controls

Outcomes

We reported the following outcomes up to 1 year after randomisation:

Pain intensity, measured with a pain scale (for example, visual analogue scale (VAS), numerical rating scale (NRS), or McGill pain score)

Disability, measured with a back pain-specific scale (for example, the Roland-Morris Disability Questionnaire (RMDQ)[22], or the Oswestry Disability Index (ODI)[23]). The two scales can be matched.[24]

All scores were normalised to VAS-scores (range 0 to 10, with 0 indicating no burden and 10 = maximum suffering from symptom).[25]

Planned secondary outcomes were the number of days of absence from work and a measure of quality of life.

Study design: Only RCTs were eligible.

Data extraction: We designed and piloted a data collection form created with Excel (Microsoft). Two researchers (JF, PF) independently extracted the study characteristics and outcome data. Disagreements were resolved through discussion or with assistance from a third researcher (DN), if necessary. From each study we extracted: number of study arms, predetermined duration of chronic non-specific low back pain, real duration of low back pain, number of participants (total and in study arms), age, sex, type of interventions (individualised, exercise, psychological), type of control, length of training period (both intervention and controls, training frequency (sessions per week), single training duration (minutes), number of adverse events.

For the effect size calculations, change scores of the outcomes and standard deviations of the change scores of the outcomes were retrieved. If not directly retrievable from the studies, the required data was calculated/imputed from the available data (baseline, change score, follow-up-values, and their standard deviations) according to the formulas defined in the Cochrane handbook. In addition (or, were not applicable), missing data was requested from the corresponding authors.

Effect sizes were processed as negative values (indicating a decrease in pain, which equals a symptom improvement), ie, higher negative values indicate larger effect sizes.

Study risk of bias assessment: Two researchers independently assessed the risk of bias (RoB) for the outcomes of interest of the included trials. We assessed the RoB for each study/outcomes using the following RoB assessment tools recommended by the Cochrane Collaboration:[26] random sequence generation, allocation concealment, blinding of participants, providers and outcome assessment, incomplete outcome data (dropouts) and selective outcome reporting. In the selective outcome data, we accounted for a broader assessment considering also the selective non-reporting RoB due to missing results in index meta- analyses (e.g., missing or unavailable outcome results crosschecked from method plans) according to published criteria.[27] [28] For each study, the score of the items indicated high, low or some concerns (not enough information reported) regarding RoB.[26] As both pain and disability are considered comparable in terms of the impact of bias on the effect sizes, they were rated cumulatively.

Certainty assessment:
Certainty of evidence was assessed applying the framework as provided by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system of rating quality of evidence and grading strength of recommendations in systematic reviews (see supplementary data appendix F.[29] The five GRADE domains were applied: study limitations, indirectness, inconsistency, imprecision and publication bias.

To interpret the clinical relevance of the meta-analyses results, we considered a 1.5-point difference in pain and a 1.0-point difference on a 10 point VAS in disability to be clinically important for the comparisons of exercise versus any comparator group .[6] We interpreted smaller differences in the effectiveness of exercise treatments compared to other conservative treatments as 'probably meaningful', when the 95% confidence interval was entirely on one side of the no effect line. This is relevant given similar inconveniences and adverse effects for comparison of treatments considered in this review (Qaseem 2017).

Statistics: Meta-analysis and -regression
For data pooling, random-effects meta-analyses were modelled. Main analyses were calculated separated by outcome (pain versus disability), by the combined comparator (active/exercise or passive comparator group), and by the intervention duration (short- and long-term-effects). Weighted standardised mean differences between the intervention and control/comparator groups (Hedge's g) were used as effect size estimators. For pooling analyses, mean effect sizes and their 95% confidence intervals were calculated; summary estimate data were displayed using Forest plots. To test for overall effects, Z-statistics at a 5% alpha-error-probability level were calculated.
Between-effects heterogeneity was assessed using the I^2- and, using restricted maximum-likelihood estimators, Tau^2-statistic. In case of heterogeneity ($tau^2 > 0$) a prediction interval for the true outcomes was also provided.

Sensitivity analyses were performed including only effect sizes from studies without a serious overall RoB and using the exercise type, the type of psychological intervention, and the characteristics of the comparator group.

To detect outliers on study level, Studentised residuals and Cook's distances were used. Outliers were Bonferroni-corrected with two-sided alpha = 0.05.

To check the included studies for a potential publication bias, Funnel plots and Egger's regression test, to check a potential funnel plot asymmetry using the standard error of the observed outcomes as predictor, were used.

All pooled effects analyses were performed using the MAJOR package for jamovi (version 1.6.23, jamovi.org, Sydney, Australia).

Meta-regressions were further calculated. Independent variables were the intervention duration, the type of comparator (active versus passive), known effective versus rather ineffective exercise types, control group matched to the intervention group "yes versus no", behavioural treatment "yes versus no", number of the participants, overall RoB rating, and mean age of the participants. The dependent variables were the effect sizes for pain and disability (continuous). For the regression model, a syntax for IBM SPSS was used (D.B. Wilson; Meta-Analysis Modified Weighted Multiple Regression; MATRIX procedure Version 2005.05.23). Inverse variance-weighted regression models with random intercepts and fixed slopes were calculated. Homogeneity analysis (Q and p-values), meta-regression estimates (95% confidence intervals and p-values), and Z-statistics were calculated.

Results

Our electronic database searches identified 1,172 unique citations. We screened the full-texts of 109 publications, identified 73 as potentially eligible, and included 58 RCTs of individualised exercise treatment for chronic low back pain in this systematic review with meta-analysis and meta-regression. At least, 53 studies contained sufficient data to be included into the qualitative analysis, with five studies lacking sufficient follow-up data. Figure 1 provides detail about the flow of citations and studies through our search and selection process. References to included and excluded studies at the full-text stage are provided within supplementary data appendix D.

Study characteristics: A total of 9099 participants were included in 58 trials (Table 1, supplementary data appendix D). The sample size of trials ranged between 67 and 221 participants (IQR) with a median of 108 participants for each study. Reported inclusion criteria for low back pain in trials was more than 12 (7 to 12) weeks prior to study inclusion, and if reported the median time of perceived burden was 128 (72-459) weeks. Thirteen studies included as well sub-chronic pain duration and sub-acute intervals of low back pain, [30-42] but consisted with a major share of participants with chronic or chronic-recurrent pain.

The median year of publication of RCTs was 2015 (range 1993–2021). The median age of participants was 44.3 years (IQR 41.6-49.2), and the median percentage of females was 55.7% (49.3%–62.5%). The median training period was 8.0 (6.0-12.0) weeks, and the median number of training sessions per week was 2.0 (1.5-3.0) for the individualised exercise and 1.5 (1.0-2.4) for the control exercises. Duration of training was 47.5 (35.8-60.0) minutes for individualised, and 45.0 (30.0-60.0) minutes for control exercise. Twenty-four (41.4%) studies investigated mixed individualised exercises, and 20 studies (34.5%) sensorimotor training (SMT). Other individualised interventions included resistance (n = 1), endurance (n = 3), stretching (n = 1), pilates (n = 1), McKenzie (n = 4), back schools (n = 3), and yoga (n = 1).

Thirty-one (53.4%) studies added some form of psychological intervention to the exercise procedure, i.e., cognitive-behavioural therapy (n = 6, 10.3%), behavioural therapy (n = 4, 6.9%), or other (n = 21, 36.2%).

In total, 71 control groups could be identified, with eleven studies presenting more than one control group. Forty-four (62.0%) of control interventions were non-individualised active exercises, and 16 (22.5%) no treatments with advice to stay active and usual care. Other controls were true controls (no treatment, n = 4), passive controls (n = 6), and medications (n = 1). Eleven of these control groups were matched (i.e., same type of exercise except individualisation).

There was not sufficient data to analyse secondary data.

>>INSERT FIGURE 1<<

>>INSERT TABLE 1<<

RoB assessement:

RoB assessments on study level (as derived from the outcomes) are summarised in Figure 2 and supplementary data appendix E. Of the 58 studies, 26 (44.3%) had an overall unclear RoB (i.e., some concerns), and 32 (55.7%) a high overall RoB. In summary, there was a considerably high overall RoB and this due to deviations from intended interventions (Figure 2, supplementary data appendix C).

The funnel plot (supplementary data appendix E) highlights the RoB across studies (publication bias on meta-level for the main outcomes pain and disability (short-term, in comparison to any comparator group). For the detailed main analyses, both the rank correlation and the regression test indicated potential funnel plot asymmetry for pain in the short-term follow-up (Egger's regression test $p = 0.02$ versus active treatments and $p = 0.04$ versus passive treatments), but not in the long-term follow-ups ($p = 0.9$ versus active and $p = 0.09$ versus passive). No funnel plot asymmetry was detected in any disability model (short-term follow-up: $p = 0.1$ versus active treatments and $p = 0.9$ versus passive treatments; long-term follow-up: $p = 0.9$ versus active and $p = 0.4$ versus passive). For the pain models of short-term follow-up, a publications bias is supposable, whereas it is unlikely for long-term pain and disability.

>>INSERT FIGURE 2<<

Overall GRADE assessment:

The overall certainty of the evidence was rated as very low for meta-analytic outcomes of self-reported pain and disability. The main reasons for downgrading the evidence were RoB, inconsistency, and publication bias. Imprecision and indirectness were no problem, as this systematic review encompassed specific populations, types of interventions, and outcome measures (see certainty of evidence section and supplementary data appendix F).

Results of individual studies:

Pain intensity was assessed in 46 studies at short-term (closest to 12 weeks after randomisation or after 12 weeks of intervention) or at long-term follow-up (closest to 12 months after randomisation).

Mean pain intensity at baseline was 4.89 (standard deviation 1.39) on a VAS in the individualised groups and 4.76 (1.52) in controls. Pain following individualised exercise was reduced to 2.82 (1.23) at 12 weeks and 3.08 (1.08) at 52 weeks. Pain intensity in active controls decreased to 3.40 (1.27) at 12 weeks and 3.38 (1.14) at 52 weeks, and in passive controls to 3.66 (1.14) and 3.23 (1.09).

Disability was assessed in 45 studies at short-term or at long-term follow-up. Mean disability at baseline was 3.39 (1.47) VAS in the individualised groups and 3.28 (1.46) in controls. Disability following individualised exercise was reduced to 2.29 (1.43) at 12 weeks and 2.43 (1.44) at 52 weeks. At 12 and 52 weeks, disability in active controls was reduced to 2.59 (1.41) and 2.43 (1.36), and in passive controls to 2.56 (1.49) and 2.61 (0.98). For individual differences in studies see supplementary data appendix C.

Pain intensity

Individualised exercise significantly reduced pain compared with all active comparators at short-term follow-up (closest to 12 weeks after randomisation; I^2 = 75%, k = 32 SMD −0.33; 95% CI −0.47 to –0.18; Figure 3), and was not significantly different at long-term follow-up (closest to 12 months; I^2 = 35%, k = 19 SMD −0.09; 95% CI −0.20 to 0.01; supplementary data appendix H). Mean clinically important differences did not suggest relevance at short-term (-0.63 cm VAS) or long-term follow-up (-0.20).

The intensity of pain following the intervention showed that individualised exercise significantly reduced pain compared with passive treatments or true control at short-term follow-up (I^2 = 87%, k = 23 SMD −0.35; 95% CI −0.53 to –0.17; Figure 4), and at long-term follow-up (I^2 = 26%, k = 15 SMD −0.14; 95% CI −0.23 to –0.06; supplementary data appendix H). Mean clinically important differences did not suggest relevance at short-term (-0.75 cm VAS) or long-term follow-up (-0.29).

Disability

Under consistency, individualised exercise did significantly reduce disability compared with all active comparators at short-term follow-up (I^2 = 74%, k = 33 SMD −0.16; 95% CI −0.30 to –0.02, Figure 5), and at long-term follow-up (I^2 = 65%, k = 17 SMD −0.07; 95% CI −0.22 to 0.07, supplementary data appendix H). Mean clinically important differences did not suggest relevance at short-term (-0.27 cm VAS) or long-term follow-up (-0.02).

Individualised exercise significantly reduced disability compared with passive treatments or true control at short-term follow-up I^2 = 51%, k = 20 SMD −0.12; 95% CI −0.22 to –0.02, Figure 6) and at long-term follow-up (I^2 = 48%, k = 14 SMD −0.23; 95% CI −0.33 to –0.12, supplementary data appendix H). Mean clinically important differences did not suggest relevance at short-term (-0.24 VAS) or long-term follow-up (-0.07).

Certainty of evidence:

Main outcomes on pain intensity have been graded with low (vs. active) and very low (vs. passive interventions), and on disability as very low (vs. active) and low (vs. passive). Outcomes at long-term follow-up seemed moderate when comparing individualised exercise to passive treatments regarding both pain and disability. Compared to active treatments evidence was rated low (pain intensity) or very low (disability).

Sensitivity and moderator analyses

All effects at short-term follow-up were calculated in comparison to other active treatments (exercises). The effect on pain, but not on disability, was robust when only studies with utmost moderate overall RoB were pooled (Table 1, *supplementary data appendix E)*

Most further sensitivity analyses, i.e., i) individualised multimodal (mixed) exercise interventions, ii) individualised exercise in combination with psychological intervention, iii) individualised exercise interventions based on SMT, iv) only studies with matched control groups, v) only studies with the inclusion criterion of low back pain lasting for at least 12 weeks, and vi) only studies with an additional behavioural treatment showed individualised exercise to significantly reduce pain intensity at short-term follow-up compared with other active treatments, but not disability (Table 2, supplementary data

appendix I1-I7). In contrast, the comparisons to vi) only true control comparators revealed superior effects by the individualisation in both pain and disability.

>>INSERT TABLE 2<<<

Mean differences when compared to controls ranged from -0.20 to -1.10 VAS (depending on the sensitivity scenario) for pain intensity, with largest clinical effects for combined treatments of individualised exercise and cognitive behavioural therapy (-1.10 VAS), and from -0.07 to -0.49 for disability (see supplementary data appendix G). On an individual level, all effects of individualised exercise on pain were clinically important (> 1.50 VAS), except at long-term follow-up. Effects on disability were clinically important at both short-term and long-term follow-up, except for individualised exercise not including psychological interventions and studies including low back pain episodes shorter than 12 weeks. Controls showed clinical important effects when including psychological interventions, and on long-term disability.

Dose-response relationships have been performed by means of meta-regression analysis (supplementary data appendix I8). The independent variables (intervention duration, type of comparator, distinction between rather effective or ineffective exercise, matched control group, additional psychological intervention, total study size (n), overall risk of bias rating, and mean age) explained 13% of the pain and 8% of the disability effect size heterogeneity. For pain, the only significant contributor was the training duration (in weeks, longer durations mediated larger effect sizes). For disability, no significant regressors could be identified (see supplementary data appendix I8).

Adverse Events

Adverse events have been reported by 14 (24%) of studies, with a total of 553 events, of which 50 (9%) were at least possibly treatment-related. Retrieved data did not allow to distinguish between types of active interventions or perform further analysis.

Discussion

This systematic review and meta-analysis with meta-regression found that individualised exercise interventions versus other active exercise interventions, passive controls, or usual care were associated with reduced self-reported pain intensity and disability. Our analysis revealed medium effect sizes for the reduction of pain intensity by individualised exercise with low levels of certainty when compared to active and very low certainty compared to passive treatments at the follow-ups closest to 12 weeks. Disability was reduced with very low (versus active) and low certainty (versus passive comparators). At long-term follow-up (closest to 12 months) no effect could be observed when comparing individualised exercise to active treatments (low certainty: pain intensity; very low certainty: disability). The reduction of pain and disability compared to passive treatments was rated with moderate certainty. Sensitivity analysis revealed that the combination of individualised exercise in combination with psychological interventions (e.g., cognitive and/or behavioural therapy) showed highest effect sizes on a low level of certainty compared to both active and passive treatments. Effects obtained following individualised exercise were mostly clinically relevant.

These data indicate that general individualised exercise can be recommended in order to achieve improvements in in pain intensity and disability in low back pain patients

Regarding the MCID, our data implicates a favour on pain and disability towards individualised exercise compared to both, active and passive controls. However, MCID mostly failed to be relevant by definition, i.e.,below 1.5 points on the VAS. Anyhow, we consider (based on the confidence intervals) most effects on pain as probably clinically relevant when individualised exercises are compared to passive measures or other exercises. The clinical relevance may be strengthened due to the following reasons: i) adding a stratification and /or personalisation to the exercise intervention is simple and almost time and cost neutral, ii) at 12 weeks pain intensity was reduced by individualised exercise by 40% and at 26 weeks by 38%, which from a patient's perspective is of clear benefit, iii) all results are at least consistent in the point that there is a trend (or, even, mostly a significant between-group difference leading to a probable relevant effect) towards individualised treatments that becomes larger when not comparing to active but passive controls, iv) taking into account the strong known placebo effects in low back pain[43] emphasizes the small effects retrieved in our study. As all studies were patient-examiner blinded -what hands-on trials cannot be blamed for- this observation seems robust.

Compared to other low back pain treatments, the observed effect sizes in our study are superior to, e.g., paracetamol (which has been declared being not effective compared to placebo),[44], somewhat larger than therapeutic ultrasound (small to no effects versus placebo),[45] non-steroidal anti-inflammatory drugs (small effects versus placebo),[46] spinal manipulative therapy (small effects versus standard care),[47] and comparable to acupuncture (low-certainty evidence to reduce pain and improve function compared to sham and no treatment).[48] Hayden et al. recently assessed the effects of exercise on chronic low back pain and disability to be small compared to no treatment, usual care or placebo comparisons, and other conservative treatments.[6] Certainty of evidence is low to moderate,

and studies did not meet the threshold for minimal clinically important difference.[6] The mean difference between groups as assessed by the meta-analysis on classification-based exercise in low back pain,[18] was presented as the difference to all comparators, suggesting small effects. All above mentioned reviews are principally in line with our results, suggesting effects on very low to moderate certainty, which are at first sight not clinically important. The problem of the above made assumptions is that the clinical relevance of the effect cannot just be defined by the MCID.[49] Medical assessment should include other factor as well, such as regression, baseline severity, or even expectations. If the MCID is calculated on the basis of a comparison to no treatment controls, it reflects well the clinical state of art.[6 49] In the case of the present systematic review, patients usually received at least advice, and a larger number of no-treatment controlled studies is missing (see Table 1). Calculating the mean differences compared to no treatment controls (n = 4) or usual care and advice only (n = 16; see supplementary data appendix G) resulted in clinically relevant intragroup differences for the individualised exercise group, but not controls, making our findings more likely. Consequently, the MCID is a context-specific value rather than a fixed number,[50] and gives insight into relative but also absolute effects. The absolute symptom reduction varied between 40% (pain) and 33% (disability) from baseline, which can indeed be considered clinically relevant from a pain physician's perspective. Especially with regard to the combined approaches (adding psychological interventions or exercise based on motor control) the reduction of pain intensity (on a low to moderate certainty level) is promising from a clinical point of view. In this context, the impact of the results as shown by Tagliaferri et al. can be regarded clinically relevant.[18] Summarising our data, the overall low certainty, RoB, MCID, and SMDs, data support the conclusion that individualised exercise interventions versus other active exercise interventions or usual care or passive controls are associated with clinically relevant reduced self-reported pain intensity and disability.

Recent clinical practice guidelines[51-54] are basically in line with our findings. The German national guideline on chronic non-specific low back pain[51] strongly recommends the combination of exercise treatments with education, the latter based on behavioural therapy, to treat sub-acute and chronic low back pain, and to promote physical activity.

The NICE guidelines[52] universally recommend to consider a combined physical and psychological intervention, incorporating a cognitive behavioural approach. Physical programs refers to a group exercise program (biomechanical, aerobic, mind–body or a combination of approaches) for people with specific episodes or flare-up of low back pain with or without sciatica that should take people's specific needs, preferences and capabilities into account when choosing the type of exercise.

The guideline of the North American Spine Association[53] recommends cognitive behavioural therapy in combination with physical therapy, as compared with physical therapy alone, to improve pain levels in patients with low back pain over 12 months on level A. Cognitive behavioural therapy in combination with physical therapy is recommended to improve functional outcomes (disability) and return to work in low back pain patients on a level B.

We belief that our results strengthen the use of exercise therapies in multimodal settings for low back pain therapies. On the basis of our data, the guideline statements in regard of patients' needs and differentiation among exercise modalities seem justified. Personalisation is an important option in

tailoring exercise to patient's needs. Different exercise modalities seem to have different impact on pain and disability. Furthermore, the bio-psycho-social dimension of low back pain should also be addressed using adequate psychological interventions.

Promising research deficits regarding that should be subject to future research include aspects like digitalisation[55] or cost-effectiveness.[56]. From a therapist's point of view, different forms of exercise therapy should be compared with each other to determine a potentially superior mode of exercise training for rehabilitation. However, given the current paucity of literature, pairwise meta-analyses may be limited in drawing such conclusions. Network meta-analysis may be more suitable for determining potentially superior modes of exercise training and have recently gained momentum in the field of sports medicine.[14] Studies should focus on strong methodological rigor and larger sample sizes to reduce the RoB and increase the certainty in observations. To ensure a low risk of bias, placebo- or sham-controlled trials should be considered.[57] Furthermore, the studies should follow current guidelines for intervention description (e.g., the template for intervention description and replication checklist[57]) to enable transparent evaluation and replication of intervention programs and should report factors potentially influencing the findings (e.g. comorbidities and pain management). Despite of a low adverse effects risk being attributed to exercise therapies,[58] the reporting of adverse events in retrieved studies was poor. Our observations mainly rely on one large study,[34] including almost 75% (n = 414) of reported adverse events, of which 98 are treatment-related. The strengths of our study include the overall assessment of individualised exercise intervention to give a concise overview of the different modalities how personalisation in exercise therapy might be performed (Supplementary File Appendix C). We also retrieved studies that included patients with sub-chronic low back pain, and sensitivity analysis suggested these patients not to distort our analysis. This group of patients could also benefit from individualised exercise. Statistical sensitivity analyses were finally conducted for all outcomes to further check the robustness and validity of the results.

Limitations

There were no included studies with a low RoB. We also could not assess the effect of individualised exercise on the number of days of absence from work and measures of quality of life, because the included studies did not report these important variables consistently. Consequently, our outcome is restricted to pain or disability only. It should also be noted that the results only apply to land-based interventions. In regard to our certainty assessment downgrading was necessary due to supposed inconsistency. When individual study sample sizes are small, point estimates may vary substantially but, because variation may be explained by chance, I^2 may be low. Conversely, when study sample size is large, a relatively small difference in point estimates can yield a large I^2.[59] This might also be a reason for the supposed publication bias regarding pain outcomes.

Finally, although exercise training is considered relatively safe in general,[58] adverse events could not be adequately assessed given insufficient reporting.

Conclusions

This systematic review shows on average low-certainty evidence for the effectiveness of individualised exercise on pain and disability in chronic non-specific low back pain. The effects are of clinical importance when assessing individualised exercise by its own. Given the little extra effort to individualise exercise, practitioners should be encouraged to use this approach. Effects are larger when combining exercise and cognitive-behavioural interventions for the treatment of chronic low back pain. With regard to the bio-psycho-social dimension of low back pain, it is important to address psychosocial characteristics, such as inappropriate beliefs about pain, pain catastrophizing, and/or depression, besides bodily functions.[60]

Legends for Figures

Figure 1: Flow Chart. RCT = randomised controlled trial, LBP = chronic non-specific low back pain

Figure 2. Risk of bias rating on an individual study level, displayed as traffic light plot for each study/outcome with green lights = low, yellow = unclear/moderate, and red = high risk of bias. The aggregate Cochrane Risk-of-bias appraisal results summary plot is displayed in the supplementary data appendix E.

Figure 3. Forest plot for the effect sizes for the short-term follow-up (closest to 12 weeks following randomisation) of individualised exercise versus other active treatments on pain intensity. The plot depicts model fit, individual study and pooled effect size estimates (standardised mean differences (SMD) and corresponding 95% confidence intervals (95% CI)). The size of the boxes corresponds to the respective studies' (inverse variance) weighting.

Figure 4. Forest plot for the effect sizes for the short-term follow-up (closest to 12 weeks following randomisation) of individualised exercise versus passive treatments, true control, and usual care on pain intensity. The plot depicts model fit, individual study and pooled effect size estimates (standardised mean differences (SMD) and corresponding 95% confidence intervals (95% CI). The size of the boxes corresponds to the respective studies' (inverse variance) weighting.

Figure 5. Forest plot for the effect sizes for the short-term follow-up (closest to 12 weeks following randomisation) of individualised exercise versus other active treatments on disability. The plot depicts model fit, individual study and pooled effect size estimates (standardised mean differences (SMD) and corresponding 95% confidence intervals (95% CI). The size of the boxes corresponds to the respective studies' (inverse variance) weighting.

Figure 6. Forest plot for the effect sizes for the short-term follow-up (closest to 12 weeks following randomisation) of individualised exercise versus passive treatments or true control on disability. The plot depicts model fit, individual study and pooled effect size estimates (standardised mean differences and corresponding 95% confidence intervals). The plot depicts model fit, individual study and pooled effect size estimates (standardised mean differences (SMD) and corresponding 95% confidence intervals (95% CI). The size of the boxes corresponds to the respective studies' (inverse variance) weighting.

Table 1. Characteristics of included studies. Inclusion criteria of LBP indicates if the study had a defined onset of the current LBP episode, whereas duration of LBP refers to the overall duration of the complaint if indicated. I = individualised group, C(s) = control group(s), SMT = sensorimotor training, BT = behavioural therapy, CBT = cognitive behavioural therapy, LBP = chronic non-specific low back pain, w = weeks, y = years, n = number, min = minute, RoB = Risk of Bias

Author, Year	Study arms (n)	Inclusion criteria LBP (w)	Duration of LBP (w)	N total	N groups	Age Mean (y)	Females (%)	Individ-ualised inter-vention	Psycho-logical inter-vention	controls	training period (w)	training frequency (sessions per week)	frequency (sessions per week) control	training duration (min)	training duration (min) control	Overall RoB
Aasa, 2015	2	12	330	70	I: 35; C: 35	42	55.7	SMT	no	active control	8	1.5	1.5	25	60	some concerns
Andersen, 2016	3	12		140	I: 46; Cs: 47	45.6	55.7	resistance	other	C1: passive control C2: true control	10	3	1	50	150	high
Apeldoorn, 2012	2	6	4	156	I: 74; C: 82	42	57.1	SMT	no	active control	8					high
Azevedo, 2018	2	12		148	I: 74; C: 74	41.9	61.5	mixed	other	active control	8	1.5	1.5	53	53	high
Brady, 2018	2	12	111	48	I: 24; C: 24	54	81.3	mixed	other	active control	12	1	1			some concerns
Cairns, 2006	2		100	97	I: 47; C: 50	38	51.5	SMT	no	active control	12	1	1	30	30	high
Cherkin, 1998	3	1		321	I: 133; C: 122	40.7	48.0	McKenzie	no	passive control	4	2.25	2.25			high
Cuesta-Vargas, 2011	2		15	49	I: 25; Cs: 24	38.4	55.1	mixed	other	active control	15	3	3	80	60	high

Cuesta-Vargas, 2015	2	12		114	I: 58; C: 56	49	23.7	mixed	no	active control	8	2	3	60	60	high
Descarreaux, 2002	2	1		20	I: 10; C: 10	34	30.0	SMT	no	active control	6	7	7			high
Diaz-Arribas, 2015	3			461	I: 132; C1: 120 C2: 209	47	82.0	SMT	BT	C1: true control C2: active control	8	2	2	50	45	high
Ford, 2016	2	6	15	300	I: 156; C: 144	43	49.0	mixed	no	advice and usual care	5	2	0.4			some concerns
Garcia, 2018	2		43	147	I: 74; C: 73	56	76.2	McKenzie	no	true control	5	2	2	35	35	some concerns
Geisser, 2005	4	12	300	100	I: 26; Cs: 24-25	38	59.0	SMT	other	all active controls						high
Godfrey, 2020	2	12		248	I: 124; C: 124	48	59.3	mixed	no	active control	2	1		60		some concerns
Goertz, 2017	3	4		131	I: 44; C1: 44 C2: 43	73	38.9	mixed	other	C1: active control C2: advice and usual care	12					high
Gudavalli, 2006	2	12		235	I: 112; C: 123	41	37.4	mixed	no	passive control	4	3.5	3.5	38	4.50	some concerns
Hansen, 1993	3	4	144	180	I: 60; C1: 61 C2: 59	41.4	31.7	SMT	no	C1: true control C2: active control	4	2	2	60	60	high
Heinrich, 2009	3	1	9	251	I: 76; C1: 53	45.8	6	mixed	CBT	C1: mixed, no CBT	12	2.5	2.5	205	75	some concerns

					C2: 125					C2: advice and usual care						
Henry, 2014	2	52		102	I: 77; C: 25	42	50.0	SMT	no	active control	6	1	1			some concerns
Highland, 2018	2	12		68	I: 34; C: 34	44.3	63.2	yoga	no	advice and usual care	8	1.5		60		some concerns
Hill, 2011	2			851	I: 568; C: 283	49.5	58.8	mixed	other	advice and usual care	16					some concerns
Hill, 2020	2			524	I: 231; C: 293	61.1	60.7	mixed	other	active control	26					high
Hunter, 2012	2	12		52	I: 28; C: 24	43	63.5	mixed	CBT	active control	6	1	1	60	60	high
Hurley, 2015	3	12		246	I: 83; C1: 82 C2: 81	45.4	67.9	endurance	no	C1:active control C2: advice and usual care	8	5	1	30	60	some concerns
Jay, 2015	2	12		25	I: 12; C: 12	46	0.0	mixed	CBT	advice and usual care	10	4		35		some concerns
Jensen, 2011	2			351	I: 176; C: 175	42	52.1	mixed	other	advice and usual care	12					high
Kim, 2020	2	7	100	39	I: 20; C: 19	24	46.2	SMT	other	advice and usual care	6	5	5			high
Lang, 2021	2	12		174	I: 117; C: 57	46	59.8	stretching	other	advice and usual care	12					some concerns
Lehtola, 2016	2	6		70	I: 35; C: 35	49.5	60.0	SMT	no	active control	8	0.8	0.8	45	45	some concerns
Leibetseder, 2007	2	12		44	I: 22; C: 22	58	63.6	endurance	no	active control	3	6	6	80	180	high
Lomond, 2015	2	48		33	I: 21; C: 12	36	100.0	SMT	other	active control	6			53	53	high

Long, 2004	3			230	I: 80; C1: 70 C2: 80		47.4	McKenzie	other	C1: active control C2: advice and usual care	2	3	3	38	38	high
Macedo, 2012	2	12	87.35	172	I: 86; C: 86	49.2	59.3	SMT	CBT	active control	8	1.5	1.5	60	60	high
Magalhães, 2018	2		67	66	I: 33; C: 33	46.9	74.2	mixed	no	active control	6	2	2	60	60	high
Michaelson, 2016	2	12	326	70	I: 35; C: 35	42	55.7	endurance	no	passive control	8	1.5	1.5			some concerns
Moore, 2000	2			226	I: 113; C: 113	49.45	54.0	back school	other	advice and usual care	12					some concerns
O'Keeffe, 2020	2	26		206	I: 106; C: 100	48	73.8	mixed	CBT	active control	8	1.5	1	45	75	high
Paolucci, 2011	2			50	I: 29; C: 21	58	62.0	back school	BT	medication	3	3		60		high
Petersen, 2011	2	6	95.5	350	I: 175; C: 175	73.5	55.7	McKenzie	no	passive control	12	1.25	1.25			high
Rabiei, 2021	2	12		73	I: 37; C: 36	43	53.4	SMT	other	active control	8	2	2		60	high
Rabin, 2014	2			105	I: 48; C: 57	36	53.3	SMT	BT	active control	8	1.375	1.375	15	30	some concerns
Rasmussen-Barr, 2009	2	8	519	71	I: 36; C: 35	38	50.7	SMT	no	advice and usual care	8	1	1	45	15	some concerns
Salas, 2019	2	13		16	I: 8; C: 8	73.55	62.5	mixed	no	active control	5	3	3			high
Saner, 2015	2	6	540	106	I: 52; C: 54	41.6	37.7	mixed	no	active control	12	2	2	30	30	some concerns
Schaller, 2016	2			412	I: 201;	50	30.6	mixed	other	passive control	8	2	0.25	60	30	high

					C: 211											
Soukup, 1999	2		564	69	I: 34; C: 35	39	59.4	mixed	other	advice and usual care	13	1.5	0	60	0	some concerns
Suh, 2019	4	12		60	I: 15; Cs: 15	54.81	55.0	SMT	other	all active controls	6	5	5	45	45	high
Thanawat, 2017	2	12		126	I: 62; C: 64	44	83.3	mixed	other	active control	8	3	3	35	35	high
Tsauo, 2009	2	12		37	I: 20; C: 17	47.46	32.4	mixed	no	active control	12			180		some concerns
Van Baal, 2020	2	12	432	34	I: 18; C: 16	44	76.5	SMT	no	active control	7	1.5	1	40	40	high
Van Dillen, 2016	2	6	540	101	I: 47; C: 54	43	47.5	SMT	no	active control	26	1.5	1	40	40	high
Vasseljen, 2012	3	12		109	I: 36; C1: 36 C2: 37	39	65.1	SMT	no	all active controls	8	1	1	40	50	high
Verra, 2018	2	26	468	139	I: 107; C: 32	46	100.0	mixed	BT	active control	4	4				some concerns
Vibe Fersum, 2013	2	12		121	I: 62; C: 59	42	52.1	mixed	CBT	active control	12	1	1	28	30	some concerns
Von Korff, 2005	2			240	I: 119; C: 121	49.8	62.5	back school	other	advice and usual care	8	0.5	0	60		some concerns
Wajswelner, 2012	2	12	722	87	I: 44; C: 43	49	55.2	pilates	no	active control	6	2	2	60	60	some concerns
Wälti, 2015	2	12		28	I: 14; C: 14	41.5	53.6	SMT	other	active control	8	2	2	30	30	some concerns

Table 2. Sensitivity analysis of outcomes pain and disability at short-term follow-up comparing individualised exercise with active controls. Bold standardised mean differences (SMD) indicate significant effects towards the individualised group. SMT = sensorimotor training

Sensitivity analysis	Outcome	k (included studies)	SMD [95 % confidence interval]	I²-statistics
Studies with utmost moderate overall risk of bias	pain	11	**-0.35 [-0.56 to -0.14]**	66%
	disability	11	-0.17 [-0.44 to 0.10]	80%
Individualised multimodal (mixed) exercise interventions without any form of psychological therapy	pain	11	**-0.34 [-0.58 to -0.09]**	67%
	disability	13	-0.16 [-0.41 to 0.09]	73%
Individualised exercise including psychological interventions	pain	21	**-0.32 [-0.51 to -0.14]**	47%
	disability	20	-0.16 [-0.33 to 0.01]	73%
Exercise based on motor control principles	pain	17	**-0.37 [-0.55 to -0.20]**	59%
	disability	17	-0.19 [-0.39 to 0.00]	51%
Only matched comparator groups (same intervention except individualisation)	pain	8	-0.19 [-0.39 to 0.00]	83%
	disability	8	-0.16 [-0.45 to 0.13]	80%
Only LBP at inclusion > 12 weeks	pain	22	**-0.36 [-0.54 to -0.19]**	70%
	disability	23	-0.12 [-0.28 to 0.05]	69%
Only studies with an additional psychological treatment	pain	21	**-0.32 [-0.51 to -0.14]**	47%
	disability	20	-0.16 [-0.33 to 0.01]	51%
Only studies with a true control comparator group	pain	14	**-0.43 [-0.64 to -0.21]**	67%
	disability	12	-0.21 [-0.32 to 0.09]	73%

REFERENCES

1. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. *Lancet (London, England)* 2018;391(10137):2356-67. doi: 10.1016/S0140-6736(18)30480-X [published Online First: 2018/03/27]
2. Zemedikun DT, Kigozi J, Wynne-Jones G, et al. Methodological considerations in the assessment of direct and indirect costs of back pain: A systematic scoping review. *PloS one* 2021;16(5):e0251406. doi: 10.1371/journal.pone.0251406 [published Online First: 2021/05/12]
3. Ulger O, Demirel A, Oz M, et al. The effect of manual therapy and exercise in patients with chronic low back pain: Double blind randomized controlled trial. *Journal of back and musculoskeletal rehabilitation* 2017;30(6):1303-09. doi: 10.3233/BMR-169673
4. Pergolizzi JV, Jr., LeQuang JA. Rehabilitation for Low Back Pain: A Narrative Review for Managing Pain and Improving Function in Acute and Chronic Conditions. *Pain and therapy* 2020;9(1):83-96. doi: 10.1007/s40122-020-00149-5 [published Online First: 2020/02/02]
5. Hajihasani A, Rouhani M, Salavati M, et al. The Influence of Cognitive Behavioral Therapy on Pain, Quality of Life, and Depression in Patients Receiving Physical Therapy for Chronic Low Back Pain: A Systematic Review. *PM & R : the journal of injury, function, and rehabilitation* 2019;11(2):167-76. doi: 10.1016/j.pmrj.2018.09.029 [published Online First: 2018/09/30]
6. Hayden JA, Ellis J, Ogilvie R, et al. Exercise therapy for chronic low back pain. *The Cochrane database of systematic reviews* 2021;9(9):CD009790. doi: 10.1002/14651858.CD009790.pub2
7. Henschke N, Ostelo RW, van Tulder MW, et al. Behavioural treatment for chronic low-back pain. *The Cochrane database of systematic reviews* 2010;2010(7):CD002014. doi: 10.1002/14651858.CD002014.pub3 [published Online First: 2010/07/09]
8. Macedo LG, Hodges PW, Bostick G, et al. Which Exercise for Low Back Pain? (WELBack) trial predicting response to exercise treatments for patients with low back pain: a validation randomised controlled trial protocol. *BMJ open* 2021;11(1):e042792. doi: 10.1136/bmjopen-2020-042792 [published Online First: 2021/01/22]
9. Hayden JA, Ellis J, Ogilvie R, et al. Some types of exercise are more effective than others in people with chronic low back pain: a network meta-analysis. *Journal of physiotherapy* 2021;67(4):252-62. doi: 10.1016/j.jphys.2021.09.004 [published Online First: 2021/09/21]
10. Niederer D, Engel T, Vogt L, et al. Motor Control Stabilisation Exercise for Patients with Non-Specific Low Back Pain: A Prospective Meta-Analysis with Multilevel Meta-Regressions on Intervention Effects. *Journal of clinical medicine* 2020;9(9) doi: 10.3390/jcm9093058 [published Online First: 2020/09/26]
11. van Dieen JH, Reeves NP, Kawchuk G, et al. Analysis of Motor Control in Patients With Low Back Pain: A Key to Personalized Care? *The Journal of orthopaedic and sports physical therapy* 2019;49(6):380-88. doi: 10.2519/jospt.2019.7916 [published Online First: 2018/06/14]
12. Nicholas MK, Linton SJ, Watson PJ, et al. Early identification and management of psychological risk factors ("yellow flags") in patients with low back pain: a reappraisal. *Physical therapy* 2011;91(5):737-53. doi: 10.2522/ptj.20100224 [published Online First: 2011/04/01]

13. Tagliaferri SD, Miller CT, Owen PJ, et al. Domains of Chronic Low Back Pain and Assessing Treatment Effectiveness: A Clinical Perspective. *Pain practice : the official journal of World Institute of Pain* 2020;20(2):211-25. doi: 10.1111/papr.12846 [published Online First: 2019/10/15]

14. Owen PJ, Miller CT, Mundell NL, et al. Which specific modes of exercise training are most effective for treating low back pain? Network meta-analysis. *British journal of sports medicine* 2020;54(21):1279-87. doi: 10.1136/bjsports-2019-100886

15. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet (London, England)* 2011;378(9802):1560-71. doi: 10.1016/S0140-6736(11)60937-9 [published Online First: 2011/10/04]

16. George SI. What is the effectiveness of a biopsychosocial approach to individual physiotherapy care for chronic low back pain? . *Internet Journal of Allied Health Sciences and Practice* 2008;6(1):1-10.

17. Lin CW, Haas M, Maher CG, et al. Cost-effectiveness of guideline-endorsed treatments for low back pain: a systematic review. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* 2011;20(7):1024-38. doi: 10.1007/s00586-010-1676-3 [published Online First: 2011/01/14]

18. Tagliaferri SD, Mitchell UH, Saueressig T, et al. Classification Approaches for Treating Low Back Pain Have Small Effects That Are Not Clinically Meaningful: A Systematic Review With Meta-analysis. *The Journal of orthopaedic and sports physical therapy* 2021:1-49. doi: 10.2519/jospt.2022.10761 [published Online First: 2021/11/16]

19. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4(1):1. doi: 10.1186/2046-4053-4-1 [published Online First: 2015/01/03]

20. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. *Spine* 2015;40(21):1660-73. doi: 10.1097/BRS.0000000000001061

21. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. *The Lancet* 2017;389(10070):736-47. doi: https://doi.org/10.1016/S0140-6736(16)30970-9

22. Roland M, Fairbank J. The Roland–Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 2000;25(24)

23. Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66(8):271-3. [published Online First: 1980/08/01]

24. Chiarotto A, Maxwell LJ, Terwee CB, et al. Roland-Morris Disability Questionnaire and Oswestry Disability Index: Which Has Better Measurement Properties for Measuring Physical Functioning in Nonspecific Low Back Pain? Systematic Review and Meta-Analysis. *Physical therapy* 2016;96(10):1620-37. doi: 10.2522/ptj.20150420 [published Online First: 2016/04/16]

25. Shafshak TS, Elnemr R. The Visual Analogue Scale Versus Numerical Rating Scale in Measuring Pain Severity and Predicting Disability in Low Back Pain. *J Clin Rheumatol* 2021;27(7):282-85. doi: 10.1097/RHU.0000000000001320

26. Higgins JPT, Savović J, Page MJ, et al. Chapter 8: Assessing risk of bias in a randomized trial [Available from: https://training.cochrane.org/handbook/current/chapter-08 accessed 11 November 2021.

27. Page MJ, McKenzie JE, Higgins JPT. Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review. *BMJ open* 2018;8(3):e019703. doi: 10.1136/bmjopen-2017-019703

28. Page MJ, McKenzie JE, Kirkham J, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *The Cochrane database of systematic reviews* 2014;2014(10):MR000035. doi: 10.1002/14651858.MR000035.pub2 [published Online First: 2014/10/02]

29. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *Journal of clinical epidemiology* 2011;64(4):383-94. doi: 10.1016/j.jclinepi.2010.04.026

30. Apeldoorn AT, Ostelo RW, van Helvoirt H, et al. A randomized controlled trial on the effectiveness of a classification-based system for subacute and chronic low back pain. *Spine* 2012;37(16):1347-56. doi: 10.1097/BRS.0b013e31824d9f2b [published Online First: 2012/02/16]

31. Cherkin DC, Deyo RA, Battié M, et al. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. *The New England journal of medicine* 1998;339(15):1021-9. doi: 10.1056/nejm199810083391502 [published Online First: 1998/10/08]

32. Descarreaux M, Normand MC, Laurencelle L, et al. Evaluation of a specific home exercise program for low back pain. *Journal of manipulative and physiological therapeutics* 2002;25(8):497-503. doi: 10.1067/mmt.2002.127078 [published Online First: 2002/10/17]

33. Ford JJ, Hahne AJ, Surkitt LD, et al. Individualised physiotherapy as an adjunct to guideline-based advice for low back disorders in primary care: a randomised controlled trial. *British journal of sports medicine* 2016;50(4):237-45. doi: 10.1136/bjsports-2015-095058 [published Online First: 2015/10/22]

34. Goertz CM, Salsbury SA, Long CR, et al. Patient-centered professional practice models for managing low back pain in older adults: a pilot randomized controlled trial. *BMC Geriatr* 2017;17(1):235. doi: 10.1186/s12877-017-0624-z

35. Hansen FR, Bendix T, Skov P, et al. Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back pain. A randomized, observer-blind trial. *Spine* 1993;18(1):98-108. doi: 10.1097/00007632-199301000-00015 [published Online First: 1993/01/01]

36. Kim SH, Park KN, Kwon OY. Classification-Specific Treatment Improves Pain, Disability, Fear-Avoidance Beliefs, and Erector Spinae Muscle Activity During Walking in Patients With Low Back Pain Exhibiting Lumbar Extension-Rotation Pattern: A Randomized Controlled Trial.

Journal of manipulative and physiological therapeutics 2020;43(2):123-33. doi: 10.1016/j.jmpt.2019.04.004 [published Online First: 2020/04/22]

37. Lehtola V, Luomajoki H, Leinonen V, et al. Sub-classification based specific movement control exercises are superior to general exercise in sub-acute low back pain when both are combined with manual therapy: A randomized controlled trial. *BMC musculoskeletal disorders* 2016;17:135. doi: 10.1186/s12891-016-0986-y

38. Petersen T, Larsen K, Nordsteen J, et al. The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: a randomized controlled trial. *Spine* 2011;36(24):1999-2010. doi: 10.1097/BRS.0b013e318201ee8e [published Online First: 2011/03/02]

39. Rasmussen-Barr E, Ang B, Arvidsson I, et al. Graded exercise for recurrent low-back pain: a randomized, controlled trial with 6-, 12-, and 36-month follow-ups. *Spine* 2009;34(3):221-8. doi: 10.1097/BRS.0b013e318191e7cb [published Online First: 2009/01/31]

40. Saner J, Kool J, Sieben JM, et al. A tailored exercise program versus general exercise for a subgroup of patients with low back pain and movement control impairment: A randomised controlled trial with one-year follow-up. *Manual therapy* 2015;20(5):672-9. doi: 10.1016/j.math.2015.02.005 [published Online First: 2015/03/17]

41. Van Dillen LR, Norton BJ, Sahrmann SA, et al. Efficacy of classification-specific treatment and adherence on outcomes in people with chronic low back pain. A one-year follow-up, prospective, randomized, controlled clinical trial. *Manual therapy* 2016;24:52-64. doi: 10.1016/j.math.2016.04.003

42. Heinrich J, Anema JR, de Vroome EM, et al. Effectiveness of physical training for self-employed persons with musculoskeletal disorders: a randomized controlled trial. *BMC public health* 2009;9:200. doi: 10.1186/1471-2458-9-200

43. van Lennep J, Trossel F, Perez R, et al. Placebo effects in low back pain: A systematic review and meta-analysis of the literature. *European journal of pain (London, England)* 2021;25(9):1876-97. doi: 10.1002/ejp.1811

44. Saragiotto BT, Machado GC, Ferreira ML, et al. Paracetamol for low back pain. *The Cochrane database of systematic reviews* 2016(6):CD012230. doi: 10.1002/14651858.CD012230

45. Ebadi S, Henschke N, Forogh B, et al. Therapeutic ultrasound for chronic low back pain. *The Cochrane database of systematic reviews* 2020;7(7):CD009169. doi: 10.1002/14651858.CD009169.pub3

46. van der Gaag WH, Roelofs PD, Enthoven WT, et al. Non-steroidal anti-inflammatory drugs for acute low back pain. *The Cochrane database of systematic reviews* 2020;4(4):CD013581. doi: 10.1002/14651858.CD013581

47. de Zoete A, Rubinstein SM, de Boer MR, et al. The effect of spinal manipulative therapy on pain relief and function in patients with chronic low back pain: an individual participant data meta-analysis. *Physiotherapy* 2021;112:121-34. doi: 10.1016/j.physio.2021.03.006 [published Online First: 2021/05/29]

48. Mu J, Furlan AD, Lam WY, et al. Acupuncture for chronic nonspecific low back pain. *The Cochrane database of systematic reviews* 2020;12(12):CD013814. doi: 10.1002/14651858.CD013814

49. Cook CE. Clinimetrics Corner: The Minimal Clinically Important Change Score (MCID): A Necessary Pretense. *The Journal of manual & manipulative therapy* 2008;16(4):E82-3. doi: 10.1179/jmt.2008.16.4.82E

50. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Current opinion in rheumatology* 2002;14(2):109-14. doi: 10.1097/00002281-200203000-00006

51. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Nicht-spezifischer Kreuzschmerz. 2017 [cited 2021 11. november]. 2:[Available from: www.kreuzschmerz.versorgungsleitlinien.de. .

52. National Guideline C. National Institute for Health and Care Excellence: Clinical Guidelines. Low Back Pain and Sciatica in Over 16s: Assessment and Management. London: National Institute for Health and Care Excellence (UK)

Copyright © NICE, 2016. 2016.

53. Kreiner DS, Matz P, Bono CM, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of low back pain. *The spine journal : official journal of the North American Spine Society* 2020;20(7):998-1024. doi: 10.1016/j.spinee.2020.04.006 [published Online First: 2020/04/26]

54. Qaseem A, Wilt TJ, McLean RM, et al. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. *Annals of internal medicine* 2017;166(7):514-30. doi: 10.7326/M16-2367 [published Online First: 2017/02/14]

55. Du S, Hu L, Dong J, et al. Self-management program for chronic low back pain: A systematic review and meta-analysis. *Patient education and counseling* 2017;100(1):37-49. doi: 10.1016/j.pec.2016.07.029 [published Online First: 2016/08/25]

56. Miyamoto GC, Lin CC, Cabral CMN, et al. Cost-effectiveness of exercise therapy in the treatment of non-specific neck pain and low back pain: a systematic review with meta-analysis. *British journal of sports medicine* 2019;53(3):172-81. doi: 10.1136/bjsports-2017-098765 [published Online First: 2018/04/22]

57. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ (Clinical research ed)* 2014;348:g1687. doi: 10.1136/bmj.g1687

58. Niemeijer A, Lund H, Stafne SN, et al. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. *British journal of sports medicine* 2020;54(18):1073-80. doi: 10.1136/bjsports-2018-100461

59. Rucker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008;8(1):79. doi: 10.1186/1471-2288-8-79

60. Roussel NA, Nijs J, Meeus M, et al. Central sensitization and altered central pain processing in chronic low back pain: fact or myth? *The Clinical journal of pain* 2013;29(7):625-38. doi: 10.1097/AJP.0b013e31826f9a71 [published Online First: 2013/06/07]

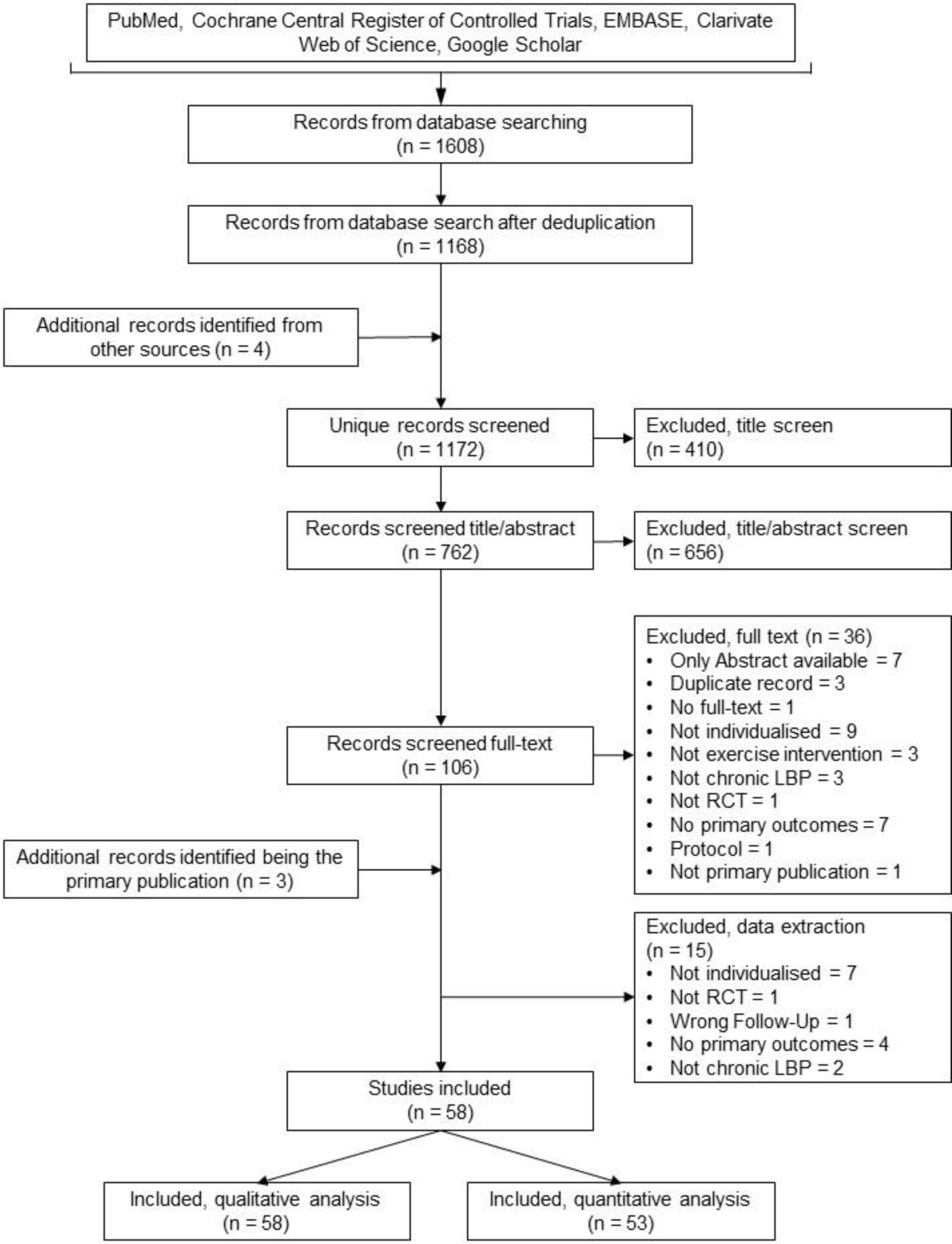
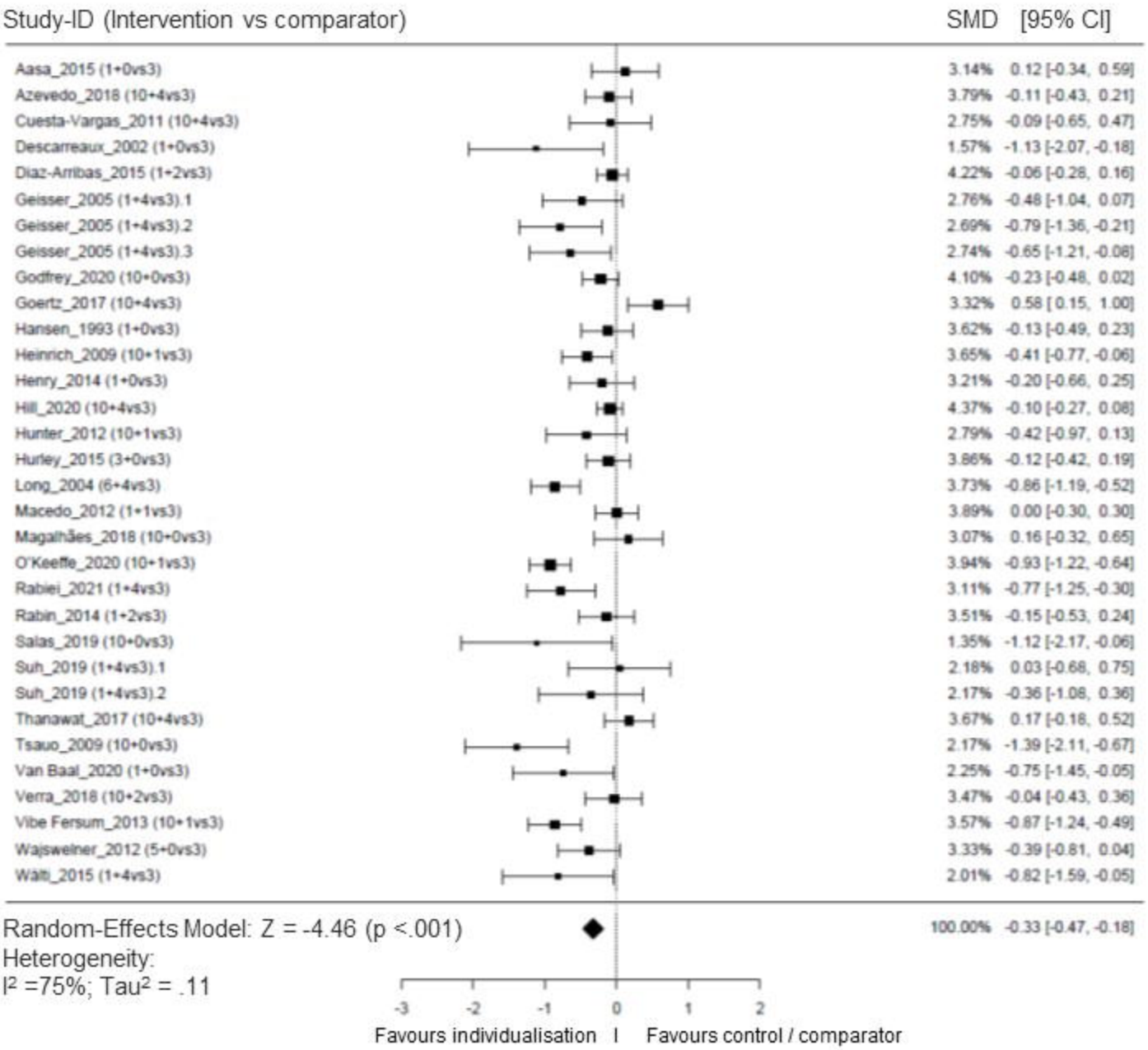

PubMed, Cochrane Central Register of Controlled Trials, EMBASE, Clarivate Web of Science, Google Scholar
Records from database searching (n = 1608)
Records from database search after deduplication (n = 1168)
Additional records identified from other sources (n = 4)
Unique records screened (n = 1172)
Excluded, title screen (n = 410)
Records screened title/abstract (n = 762)
Excluded, title/abstract screen (n = 656)
Excluded, full text (n = 36)
• Only Abstract available = 7
• Duplicate record = 3
• No full-text = 1
• Not individualised = 9
• Not exercise intervention = 3
• Not chronic LBP = 3
• Not RCT = 1
• No primary outcomes = 7
• Protocol = 1
• Not primary publication = 1
Records screened full-text (n = 106)
Additional records identified being the primary publication (n = 3)
Excluded, data extraction (n = 15)
• Not individualised = 7
• Not RCT = 1
• Wrong Follow-Up = 1
• No primary outcomes = 4
• Not chronic LBP = 2
Studies included (n = 58)
Included, qualitative analysis (n = 58)
Included, quantitative analysis (n = 53)

Figure 2

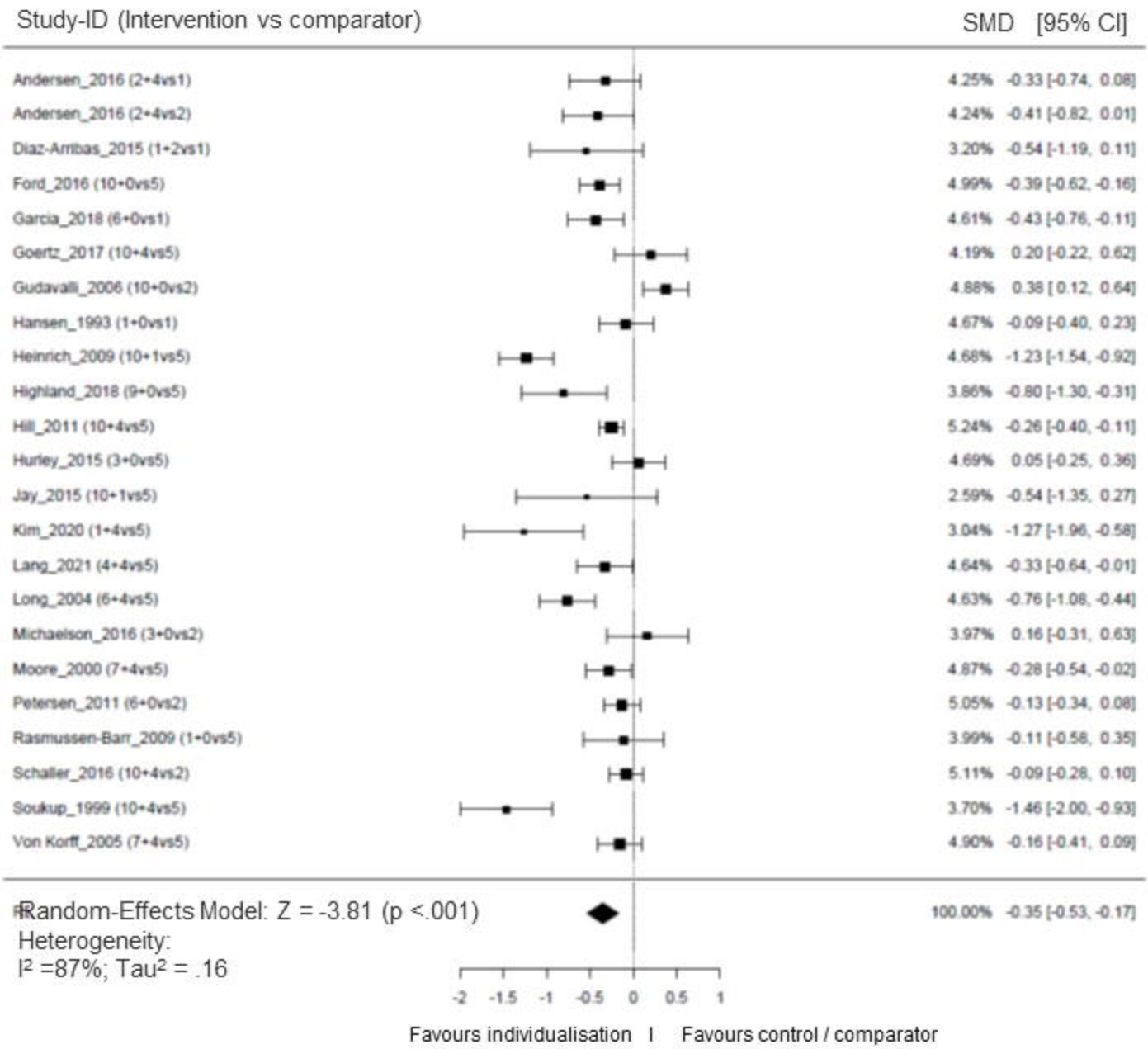
Risk of bias domains
D1
D2
D3
D4
D5
Overall
Study
Aasa_2015
Andersen_2016
Apeldoorn_2012
Azevedo_2018
Brady_2018
Cairns_2006
Cherkin_1998
Cuesta-Vargas_2011
Cuesta-Vargas_2015
Descarreaux_2002
Díaz-Arribas_2015
Ford_2016
Garcia_2018
Geisser_2005
Godfrey_2020
Goertz_2017
Gudavalli_2006
Hansen_1993
Heinrich_2009
Henry_2014
Highland_2018
Hill_2011
Hill_2020
Hunter_2012
Hurley_2015
Jay_2015
Jensen_2011
Kim_2020
Lang_2021
Lehtola_2016
Leibetseder_2007
Lomond_2015
Long_2004
Macedo_2012
Magalhães_2018
Michaelson_2016
Moore_2000
O'Keeffe_2020
Paolucci_2011
Petersen_2011
Rabiei_2021
Rabin_2014
Rasmussen-Barr_2009
Salas_2019
Saner_2015
Schaller_2016
Soukup_1999
Suh_2019
Thiswald_2017
Tsauo_2009
Van Baal_2020
Van Dillen_2016
Vasseljen_2012
Verra_2018
Vibe Fersum_2013
Von Korff_2005
Wajswelner_2012
Walli_2015
Domains:
D1: Bias arising from the randomization process.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing outcome data.
D4: Bias in measurement of the outcome.
D5: Bias in selection of the reported result.
Judgement
High
Some concerns
Low

Figure 3



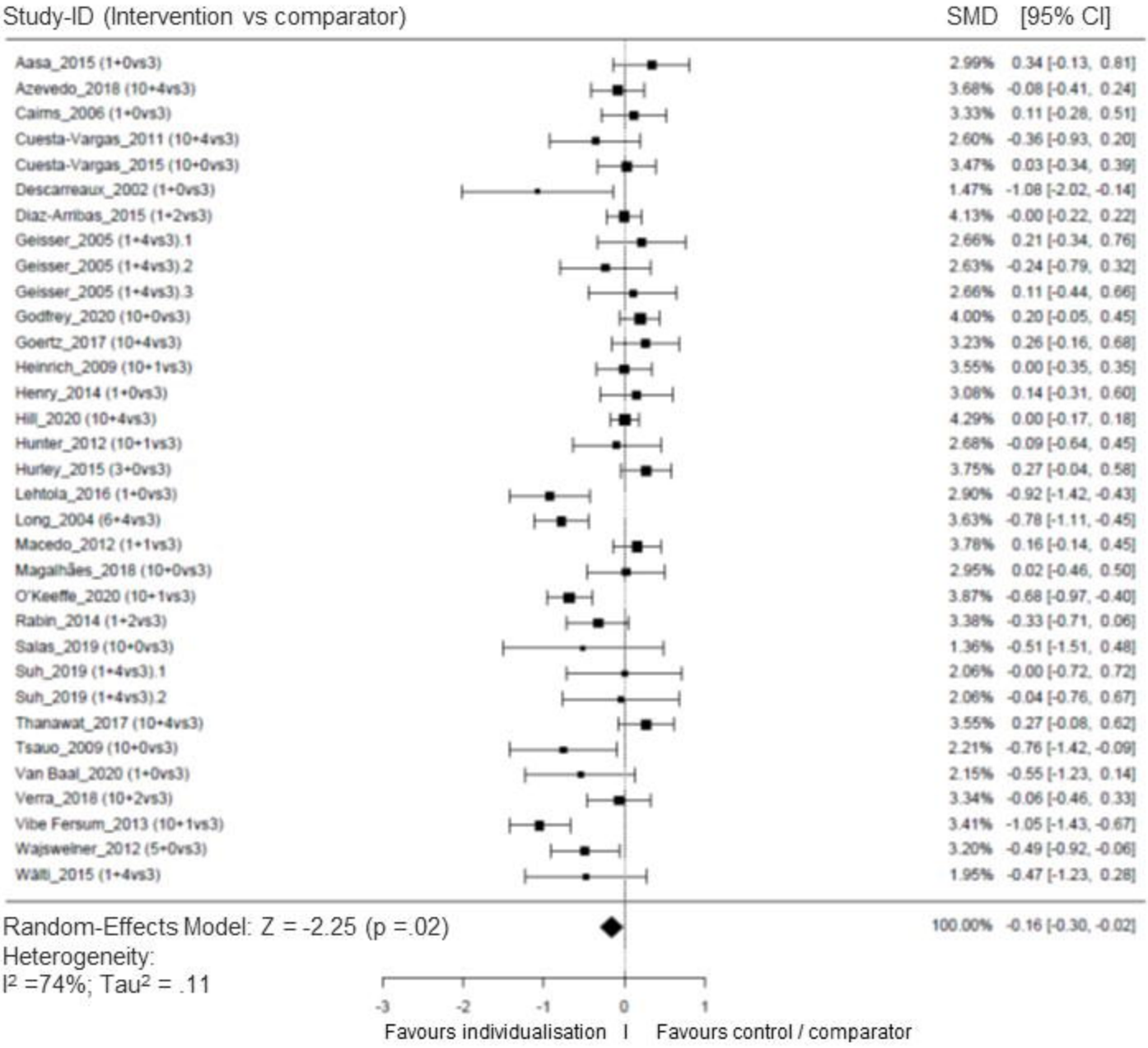
Exercise intervention		Psychological therapy		Comparator
1 Stabilisation 2 Strength 3 Aerobic 4 Stretching 5 Pilates 6 McKenzie 7 Back school 8 Mind-body 9 Yoga 10 Mixed	+	0 None 1 Cognitive behavioural 2 Behavioural 4 Other	VS	1 True control 2 Passive measures 3 Active (Exercise) 4 Medication 5 Advice or Usual care

Figure 4



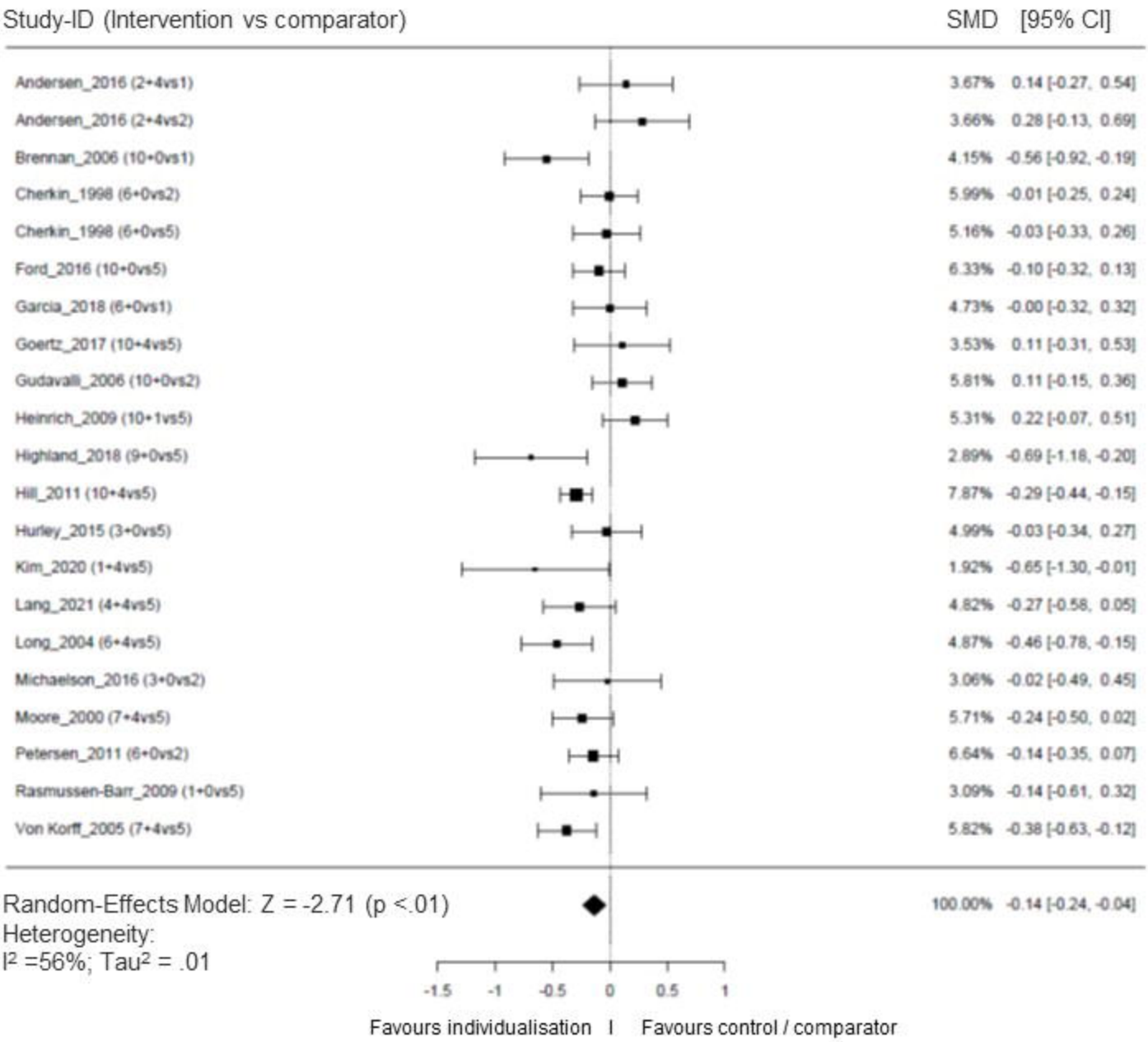
Exercise intervention		Psychological therapy		Comparator
1 Stabilisation 2 Strength 3 Aerobic 4 Stretching 5 Pilates 6 McKenzie 7 Back school 8 Mind-body 9 Yoga 10 Mixed	+	0 None 1 Cognitive behavioural 2 Behavioural 4 Other	VS	1 True control 2 Passive measures 3 Active (Exercise) 4 Medication 5 Advice or Usual care

Figure 5



Exercise intervention		Psychological therapy		Comparator
1 Stabilisation 2 Strength 3 Aerobic 4 Stretching 5 Pilates 6 McKenzie 7 Back school 8 Mind-body 9 Yoga 10 Mixed	+	0 None 1 Cognitive behavioural 2 Behavioural 4 Other	VS	1 True control 2 Passive measures 3 Active (Exercise) 4 Medication 5 Advice or Usual care

Figure 6



Exercise intervention		Psychological therapy		Comparator
1 Stabilisation 2 Strength 3 Aerobic 4 Stretching 5 Pilates 6 McKenzie 7 Back school 8 Mind-body 9 Yoga 10 Mixed	+	0 None 1 Cognitive behavioural 2 Behavioural 4 Other	VS	1 True control 2 Passive measures 3 Active (Exercise) 4 Medication 5 Advice or Usual care