Exercise treatment effect modifiers in persistent low back pain: an individual participant data meta-analysis of 3514 participants from 27 randomised controlled trials

Jill A Hayden [1], Maria N Wilson,[1] Samuel Stewart,[1] Jennifer L Cartwright,[1] Andrea O Smith,[1] Richard D Riley,[2] Maurits van Tulder,[3] Tom Bendix,[4] Francesca Cecchi,[5] Leonardo O P Costa [6] Ninna Dufour,[7] Manuela L Ferreira,[8] Nadine E Foster,[9] Maruti R Gudavalli,[10] Jan Hartvigsen [11] Pieter Helmhout,[12] Jan Kool,[13] George A Koumantakis [14] Francisco M Kovacs [15] Tiina Kuukkanen,[16] Audrey Long,[17] Luciana G Macedo [18] Luciana A C Machado [19] Chris G Maher [20] Wolf Mehling [21] Giovanni Morone [22] Tom Peterson,[23] Eva Rasmussen-Barr [24] Cormac G Ryan [25] Tuulikki Sjögren [26] Rob Smeets [27] J Bart Staal [28] Monica Unsgaard-Tøndel,[29] Henry Wajswelner,[30] Ella W Yeung [31] On behalf of Chronic Low Back Pain IPD Meta-Analysis Group

ABSTRACT

Background Low back pain is one of the leading causes of disability worldwide. Exercise therapy is widely recommended to treat persistent non-specific low back pain. While evidence suggests exercise is, on average, moderately effective, there remains uncertainty about which individuals might benefit the most from exercise.

Methods In parallel with a Cochrane review update, we requested individual participant data (IPD) from high-quality randomised clinical trials of adults with our two primary outcomes of interest, pain and functional limitations, and calculated global recovery. We compiled a master data set including baseline participant characteristics, exercise and comparison characteristics, and outcomes at short-term, moderate-term and long-term follow-up. We conducted descriptive analyses and one-stage IPD meta-analysis using multilevel mixed-effects regression of the overall treatment effect and prespecified potential treatment effect modifiers.

Results We received IPD for 27 trials (3514 participants). For studies included in this analysis, compared with no treatment/usual care, exercise therapy on average reduced pain (mean effect/100 (95% CI) −10.7 (−14.1 to −7.4)), a result compatible with a clinically important 20% smallest worthwhile effect. Exercise therapy reduced functional limitations with a clinically important 23% improvement (mean effect/100 (95% CI) −10.2 (−13.2 to −7.3)) at short-term follow-up. Not having heavy physical demands at work and medication use for low back pain were potential treatment effect modifiers—these were associated with superior exercise outcomes relative to non-exercise comparisons. Lower body mass index was also associated with better outcomes in exercise compared with no treatment/usual care. This study was limited by inconsistent availability and measurement of participant characteristics.

Conclusions This study provides potentially useful information to help treat patients and design future studies of exercise interventions that are better matched to specific subgroups.

Protocol publication https://doi.org/10.1186/2046-4053-1-64

INTRODUCTION

Low back pain is one of the leading causes of disability worldwide and has a substantial socioeconomic impact.[1–4] The majority of the cost associated with low back pain is generated by individuals with persistent symptoms.[2 5] Both the prevalence and the cost of persistent low back pain are increasing.[2] Exercise therapy is widely recommended to treat persistent low back pain and is one of several interventions that are moderately effective in reducing pain and improving function.[6]

Identifying subgroups of individuals who may benefit more from treatment, and potential treatment effect modifiers (also known as moderators or treatment–covariate interactions), is an important goal in health research. There are difficulties with most existing subgroup/classification systems for patients with low back pain; these include unclear reliability or validity in clinical practice, lack of comprehensive predictor variables, and inclusion of measures or information that is not useful, nor feasibly collected in practice.[7] Furthermore, most low back pain trials are not designed to detect treatment effect modifiers.[8 9]

Our team conducted a Cochrane review where we concluded that exercise therapy appears to be effective in decreasing pain and improving function in adults with persistent low back pain; however, this work was limited by inclusion of only published 'aggregate' data, such as overall treatment effects and average patient characteristics.[10–12] Individual participant data (IPD) meta-analysis standardises analyses across trials, allowing for more powerful

and reliable examination of differential treatment effects across subgroups of individuals,[13 14] since within-trial information can be used to estimate how individual characteristics modify treatment benefit.[15]

Our primary objective in this study was to assess the treatment effect of exercise therapy for reducing pain and functional limitations in adults with persistent low back pain, as well as any modification of treatment effect by considering a set of plausible patient features. We aimed to identify the characteristics of persons with low back pain who are more likely to benefit from exercise therapy.

MATERIALS AND METHODS
We used standard systematic review methods advocated by Cochrane Back and Neck to identify eligible trials.[16] For this IPD meta-analysis, we identified randomised controlled trials from an updated search conducted in 2013 for the Cochrane review, 'Exercise therapy for treatment of chronic low back pain'. Complete descriptions of the full search strategy, screening, selection and trial-level data extraction for the related Cochrane review are reported elsewhere.[10] The protocol for this IPD study has been published.[17]

Eligible trials evaluated all forms of exercise therapy compared with any or no other interventions in adults (>18 years of age) with non-specific persistent (>12 weeks' duration) low back pain (alone or with leg pain). Trials with mixed subacute (>6 weeks' duration) and persistent low back pain populations were eligible for the IPD meta-analysis if it was possible to extract information exclusively for those participants experiencing persistent low back pain.

For retrieval of IPD, we selected trials included in the updated Cochrane review that were rated as moderate to low risk of bias. Following Cochrane Back and Neck Methods Guidelines, we defined moderate to low risk of bias as at least 6 of 12 recommended items rated as having low risk of bias[18]; these risk of bias items align with the current Cochrane Back and Neck Methods Guidelines.[16]

Data collection and management
The original IPD was requested from the authors of 56 eligible trials. We extracted trial-level information about each eligible trial (details can be found in our Cochrane review protocol[10]) and included IPD reported characteristics of the trial sample, variables collected at baseline and at follow-up periods, and subgroup and treatment effect modifiers investigated and presented in the trial report.

For eligible trials that provided their IPD, each raw data set was saved on a secure server at Dalhousie University in its original format before being converted to a common format. We evaluated the IPD received from each trial and compared it with the available related publication(s) based on descriptive summary, range of included variables and missing observations. We attempted to use the IPD from each trial to replicate results reported in the original publications, including baseline characteristic descriptive analyses and advanced analyses of outcome data at each available follow-up period, by reproducing the statistical analyses as reported by the trial authors. We discussed and clarified any discrepancies or missing information between our results and those presented in each original publication with the original trial authors. Ultimately, we included only trial data where we could reproduce published trial findings or explain/clarify discrepancies. Once data checks were complete and satisfactory, individual trial data sets were combined to form a new master data set with a variable added to indicate the original trial.

Data preparation
We used a prespecified framework for mapping, classifying and renaming sufficiently similar variables (defined following the variable map presented in Hayden *et al*[5]). Potential treatment effect modifiers included variables in the following broad domains: participant sociodemographic characteristics, lifestyle factors, overall health, psychological status, previous low back pain, characteristics of the current episode and physical examination findings. For all variables, we preferentially selected the most valid and reliable measures available in each data set, based on supporting literature. Whenever possible, we maintained variables measured continuously in their continuous data form, while also creating categorical or dichotomous variables, as necessary, for homogeneity across studies. We assessed participant-level missing data on variables and outcomes. Individual subjects with missing data within each trial were excluded from specific analyses, as necessary.

Data analysis
Descriptive analyses
We described trial-level and participant-level characteristics of the included trials. We compared trial-level characteristics from trials included in the IPD analysis with those from eligible trials from the Cochrane review update to determine whether the IPD trials available were a representative sample of the full set of eligible trials.[19]

Meta-analyses of the overall effect of any type of exercise therapy compared with no treatment, non-exercise usual care and other conservative treatments were conducted as part of the associated Cochrane review based on aggregate data presented in the publications of primary studies (in preparation[10]). In this study, we compared published aggregate treatment effect results of exercise therapy versus any non-exercise comparisons from trials available for the IPD analysis with those from eligible studies in the Cochrane review update that did not provide IPD.

Our primary outcomes of interest were pain intensity, back-related functional limitations and a composite measure of global recovery. Pain and back-related functional limitations outcomes were self-reported as continuous measures and mainly analysed on the continuous scale to avoid losing power. To achieve this, as outcome scales varied across studies, we converted each trial's outcome data to a common 0–100 scale to facilitate synthesis across studies and interpretability of the IPD meta-analysis results. We also calculated global recovery as a dichotomous measure of clinically important individual pain or functional limitations response as any improvement in score ≥30% of its baseline value with a minimum value of 20-point (/100) improvement in pain[20] or 10-point (/100) improvement in function.[21–23] We collected existing outcome data for all available time periods. We assessed outcomes at short-term follow-up (post-treatment time period closest to 3 months) for primary analyses, and for moderate-term (time period closest to 6 months) and long-term (time period closest to 12 months) follow-up in sensitivity analyses.

IPD syntheses
For synthesis of IPD, we employed a one-stage meta-analysis approach, as specified in our protocol, using multilevel mixed-effects linear regression for continuous pain and functional limitations outcomes using restricted maximum likelihood estimation,

and multilevel mixed-effects logistic regression for dichotomous global recovery outcome using maximum likelihood estimation, accounting for the clustering of individuals within studies.[24 25] These models specified a random treatment effect (to allow for between-trial heterogeneity in effect), trial-specific intercepts (to account for clustering) and random effects for baseline values of outcome variables (to allow for between-trial heterogeneity). We assessed the effectiveness of exercise therapy compared with no treatment or usual care, and compared with other conservative treatments, including adjustment for the baseline value of the outcome of interest (ie, pain or functional limitations, as appropriate; functional limitations for global recovery outcome).

We extended the one-stage IPD meta-analysis framework to assess potential treatment effect modifiers (ie, differences between subgroups) related to effectiveness of exercise therapy compared with any other studied non-exercise treatments. We identified candidate treatment effect modifiers from available data by considering potential mechanisms for modification of treatment response (biological reasoning and by understanding the mechanism by which response is modified), and from existing prognostic research (treatment effect modification studies[26] and prognostic factor studies[27]). These were age, sex, level of education, current smoker, physical activity, body mass index (BMI), history of low back pain, work status, sick leave for the past 12 months, heavy physical demands at work, general health, general mental health, fear avoidance beliefs, social support, low back pain duration, baseline pain and functional limitations, leg pain symptoms, and medication use for low back pain.

To identify potential treatment effect modifiers, we examined interaction terms between treatment and each variable to assess subgroup effects using unadjusted and adjusted models. Interaction coefficients for dichotomous variables are interpreted as the effect of exercise treatment, relative to non-exercise comparisons, in those with the baseline characteristic compared with those who do not have the characteristic. Interaction coefficients for continuous variables are interpreted as the additional benefit of exercise therapy, relative to non-exercise comparisons, for every one-unit increase in the continuous variable. Unadjusted models included fixed effects at the participant level and random effects at the trial level: baseline pain or functional limitations (corresponding with the outcome; functional limitations variable was included for global recovery outcome), potential treatment effect-modifying variable, treatment group (exercise or comparison), and the variable–treatment group interaction. To assess whether effect modifiers remained after further adjustment, we also adjusted interaction analyses for age and sex (at the random and fixed effects levels). We centred the participant-level covariates about their trial-specific means to remove the contribution of across-trial associations of covariates and treatment effects (removing the impact of potential ecological bias on the effect modifier estimates).[28 29] Online supplementary appendix 1 describes the treatment effect modifier statistical model and sample code. For each model, participants missing data on any of the included variables were not included in the corresponding analysis.

We performed secondary analyses to explore the robustness of our results. We repeated the adjusted treatment effect modifier analyses described above to separately assess potential treatment effect modifiers for exercise therapy compared with no treatment/usual care comparison groups or with other conservative comparison groups. We assessed potential treatment effect modifiers for pain, functional limitations and global recovery outcomes in adjusted analyses at moderate-term (time period closest to 6 months) and long-term (time period closest to 12 months) follow-up.

We considered the smallest worthwhile effect for exercise treatment compared with no treatment/usual care on pain and functional limitation outcomes to be a 20% change from baseline.[30 31] We considered a variable as a potential treatment effect modifier using statistical and clinical importance of results. We report 95% CIs of summary treatment–variable interaction coefficients and exact p values. For easier interpretation of results, we computed and present the observed treatment effect sizes (as mean difference (MD) for continuous outcomes and as OR for dichotomous outcomes, with 95% CIs) from available trial data for subgroups of potential treatment effect modifiers, categorising continuous variables at clinically relevant cut-points when possible or based on the observed median. For this forest plot presentation, we used uncentred values of the potential treatment effect modifier to calculate the mean treatment effect across categories. Models to determine the magnitude of potential treatment effect modifiers were adjusted using the same centred covariates as are described for primary analyses. For our primary analyses, we discuss treatment effect modifiers as potentially important where a level of the participant characteristic changed the direction of the mean effect, with results compatible with a clinically important effect in one group and opposite effect in another group, or if a clinically important difference between covariate groups is greater than the proposed smallest worthwhile effect for exercise treatment (20% change from baseline on pain and functional limitation outcomes, and positive OR of 2.0 or greater)[30 31]; we consider these results exploratory.

Descriptive analyses and one-stage IPD overall treatment effect meta-analyses were conducted using Stata V.15.0.[32] Extension of the one-stage IPD models to include treatment effect modifiers was conducted in R V.3.5.2. Review Manager V.5.3[33] was used for meta-analyses of published aggregate data. IPD data checking and replication were conducted in SPSS V.24.0 and SAS V.9.4.[34 35]

RESULTS

Authors from 27 eligible trials, published between 2000 and 2012,[36–72] provided IPD for this study (table 1). Each trial included between 28 and 264 participants (median, 109), randomised to 1–3 exercise groups and 0–2 comparison groups (the most common design compared two exercise groups; 10 trials). The mean sample size per group in the included trials was 56 participants: four trials included more than 100 participants per group, and seven trials included less than 30 participants per group. Trials were conducted in Australia (four trials), Denmark (four trials), Italy (three trials), the Netherlands (three trials), UK (two trials), USA (two trials), Finland (two trials), one in each of Hong Kong, Spain, Switzerland, Sweden, Norway and Brazil, and one multicountry trial. Most trials were conducted in healthcare settings (10 trials in secondary care settings and 6 in primary care settings); 5 trials were conducted in occupational settings and 6 in general population or mixed settings. The mean age by trial was between 37.0 and 60.1 years, and studies included 0%–78% female participants. Twenty-six of the 27 included trials reported pain (96%) outcomes and 25 reported functional limitation outcomes (93%). All studies reported participant outcomes at short-term follow-up, with 19 and 15 respectively reporting pain or functional limitations at moderate-term and long-term follow-up. Twenty-nine eligible studies were unable to provide their data or did not respond after four attempted requests for data (figure 1); the only characteristic on which these trials differed significantly was design of exercise

Table 1 Characteristics of 27 studies with data available for IPD meta-analysis

Trial author, year	Country	Population source	Participants	% female	Mean age	Trial groups Exercise; comparison (n)	Comparison category	Outcomes available Pain intensity	Functional limitations	Follow-up periods Short	Moderate	Long
Albaladejo, 2010[36]	Spain	Primary care	264	68.2	50.9	1; 2	Education*	VAS	RMDQ†	+	+	–
Cecchi, 2010[37]	Italy	Primary care	210	66.7	58.8	2; 1	Manual therapy	Pain Rating Scale	RMDQ	+	+	+
Chan, 2011[38]	Hong Kong	Primary care	46	78.3	46.6	2; 0	N/A	VAS	Aberdeen‡	+	–	+§
Costa, 2009[39]	Australia	Primary care	154	60.4	53.7	1; 1	Ø/Usual	NRS	RMDQ	+	+	+
Dufour, 2010[40]	Denmark	Secondary care	286	55.9	40.9	2; 0	N/A	VAS	RMDQ	+	+	+
Ferreira, 2007[41]	Australia	Primary care	240	68.8	53.7	2; 1	Manual therapy	VAS	RMDQ	+	+	+
Gudavalli, 2006[43]	USA	Mixed	235	37.5	41.6	1; 1	Manual therapy	VAS	RMDQ	+	–	–
Hartvigsen, 2010[44]	Denmark	Secondary care	136	71.4	46.6	2; 1	Education	LBPRS-P	LBPRS-F	+	+	+
Helmhout, 2004[45]	The Netherlands	Occupational	81	0.0	40.9	2; 0	N/A	–	RMDQ	+	+	+
Kool, 2007[46]	Switzerland	Occupational	151	21.2	42.3	2; 0	N/A	NRS	–	+	–	–
Koumantakis, 2005[47]	UK	Primary care	52	53.9	37.0	2; 0	N/A	VAS	RMDQ	+	–	–
Kuukkanen, 2000[48]	Finland	Occupational	60	50.0	40.5	1; 1	Ø/Usual	Borg CR-10	ODI	+	+	+
Long, 2004[49]	Multisite¶	Secondary care	133	51.1	41.6	3; 0	N/A	VAS	RMDQ	+	–	–
Macedo, 2012[42]	Australia	Mixed	172	59.3	49.2	2; 0	N/A	NRS	RMDQ	+	+	+
Machado, 2007[50]	Brazil	Secondary care	33	69.7	43.5	1; 1	Psychological therapy	VAS	RMDQ**	+	+	–
Mehling, 2005[51]	USA	Mixed	28	64.3	49.3	2; 0	N/A	VAS	RMDQ††	+	+	–
Morone, 2011[52]	Italy	Secondary care	70	64.3	60.1	1; 1	Ø/Usual	VAS	ODI	+	+	–
Morone, 2012[53]	Italy	Secondary care	75	72.0	55.3	1; 2	Multiple‡‡	VAS	ODI	+	+	–
Petersen, 2011[63]	Denmark	Secondary care	224	56.3	37.1	2; 0	N/A	VAS§§	RMDQ	+	–	+
Rasmussen-Barr, 2009[54]	Sweden	Mixed	71	50.7	38.1	2; 0	N/A	VAS	ODI	+	+	+
Ryan, 2010[55]	UK	Secondary care	38	65.8	45.3	1; 1	Education	NRS	RMDQ	+	–	–
Sjögren, 2006[56]	Finland	Occupational	90	73.3	45.7	1; 1	Ø/Usual	Borg CR-10	–	+	+	–
Smeets, 2006[57]	The Netherlands	Secondary care	223	47.1	41.6	3; 1	Ø/Usual	VAS	RMDQ	+	+	+
Sorensen, 2010[58]	Denmark	Secondary care	207	52.2	39.1	1; 1	Education	NRS	RMDQ	+	+	+
Staal, 2004[59]	The Netherlands	Occupational	39	7.7	38.2	1; 1	Ø/Usual	NRS	RMDQ	+	+	+
Unsgaard-Tøndel, 2010[60]	Norway	Mixed	109	69.7	40.1	3; 0	N/A	NRS	ODI	+	–	+¶¶
Wajswelner, 2012[61]	Australia	General	87	55.2	49.6	2; 0	N/A	NRS	PSFS	+	+	–

*Both comparison groups are education.
†Spanish RMDQ.
‡Chinese Aberdeen Scale.
§Functional outcome only.
¶Multisite (Canada, USA, Germany, UK, Kuwait).
**Brazilian/Portuguese RMDQ.
††16-item RMDQ.
‡‡Group 1 is Ø/usual; group 2 is back school.
§§Sum of three VAS scales.
¶¶Pain outcome only.
Ø, 'No treatment' comparison; Borg CR-10, Borg Category-Rating Scale; F, function scale; IPD, individual participant data; LBPRS, Low Back Pain Rating Scale; N/A, not applicable; NRS, Numeric Rating Scale; ODI, Oswestry Disability Index; P, pain scale; PSFS, Patient Specific Functional Scale; QBPDS, Quebec Back Pain Disability Scale; RMDQ, Roland Morris Disability Questionnaire; SNQ, Standardised Nordic Questionnaire for activity restrictions; VAS, Visual Analogue Scale.
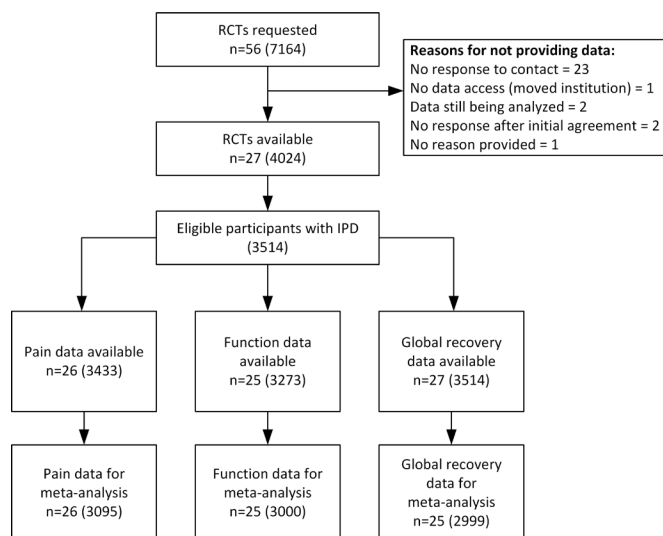
Figure 1 Flow chart of studies and participants included in the IPD meta-analysis. IPD, individual participant data; RCT, randomised controlled trial.

programme (non-included trials had standardised exercises more often) (table 2). Aggregate meta-analysis of published estimates comparing treatment effect results from studies included in the IPD analysis with those from eligible studies not providing data found smaller pain improvement, and similar functional limitation improvement with exercise therapy relative to non-exercise comparisons for included studies, compared with eligible studies not providing data (pain (0–100 scale) MD (95% CI), included: −2.3 (−8.1 to 3.5) vs not included: −10.2 (−16.4 to −4.1) (figure 2); functional limitations (0–100 scale) MD (95% CI), included: −3.6 (−8.7 to 1.4) vs not included: −2.9 (−6.2 to 0.3)) (online supplementary appendix 2).

Data from 3514 trial participants were available for our IPD meta-analyses (510 participants were excluded from two trials due to not having persistent low back pain or having missing data on both outcomes): 2568 participants were randomised to receive exercise therapy, and 946 participants were randomised to receive a non-exercise comparison (142 placebo/no treatment, 125 usual care, 679 other conservative treatments, including manual therapy, education or psychological therapy; see table 2 for a full list). The mean age of participants was 45.7 years (95% CI 45.2 to 46.1) and 55.5% (1945 of 3504) were female. From trials reporting these variables, more than two-thirds of participants reported history of low back pain (598 of 877; 68.2%), and 52.4% had been on sick leave in the previous 12 months (943 of 1801). At baseline, the median low back pain episode duration for participants was 14 months (IQR=6–48 months; n=1692), with a mean pain intensity of 53.5 on a 100-point scale (95% CI 52.6 to 54.3; n=3411); 63.1% of participants reported leg pain with their low back pain (1354 of 2145). Missing data for variables of interest ranged from 0.2% (sex) to 75% (history of low back pain) (table 3). There was heterogeneity in the measurement of potential treatment effect modifiers.

IPD integrity
For all included IPD, we were able to replicate aggregate results that were reported in each of the associated publications. Of the 27 IPD data sets that were received, trial information was fully replicated at the data verification stage for 12 data sets. The remaining

Table 2 Trial-level and group-level characteristics for 27 IPD included studies compared with 29 eligible trials that did not provide data

Trial characteristics	Included IPD trials (n=27) n (%)	Not included eligible trials (n=29) n (%)
Median number of subjects (IQR)	109 (65–209)	90 (60–132)
Population source		
Healthcare	16 (59)	20 (69)
Occupational	5 (19)	2 (7)
General or mixed	6 (22)	7 (24)
Mean age of sample, years (95% CI)	45.7 (45.2 to 46.1)	42.9 (40.1 to 45.6)
Proportion of female sample (95% CI)	55.5 (53.9 to 57.2)	44.3 (36.0 to 52.7)
Mean severity of pain at baseline (95% CI)	53.5 (52.6 to 54.3)	50.9 (44.6 to 57.1)
Number of exercise groups available		
1 exercise group	11 (41)	11 (46)
2 exercise groups	13 (48)	15 (63)
3 exercise groups	3 (11)	3 (13)
Number of comparison groups available		
1 comparison group	13 (48)	15 (63)
2 comparison groups	2 (7)	1 (4)
Exercise group types (46 groups)		
Strengthening	4 (9)	5 (8)
Stretching	1 (2)	2 (3)
Core strengthening	10 (22)	10 (15)
Flexibility/mobilising	1 (2)	0 (0)
Aerobic	4 (9)	2 (3)
McKenzie	2 (4)	0 (0)
Functional restoration	1 (2)	1 (2)
Yoga	0 (0)	5 (8)
Other	4 (9)	0 (0)
Mixed (>2 types)	19 (41)	25 (37)
Design of exercise programme		
Individualised	15 (33)	6 (12)
Partially individualised	21 (46)	18 (36)
Standardised	9 (20)	26 (52)
Not specified	1 (2)	0 (0)
Delivery of exercise programme		
Independent	5 (11)	1 (2)
Independent with FU	2 (4)	1 (2)
Group	18 (39)	30 (60)
One-on-one	17 (37)	16 (32)
Not specified	4 (9)	2 (4)
Comparison group types (17 groups)		
No treatment/usual care comparisons	7 (41)	4 (24)
No treatment	3 (18)	2 (12)
Usual care	4 (24)	2 (12)
Other conservative treatment comparisons	10 (59)	13 (77)
Advice/education	5 (29)	5 (29)
Back school	1 (6)	0 (0)
Electrotherapy	0 (0)	4 (24)
Manual therapy	3 (18)	1 (6)
NSAIDs/analgesics	0 (0)	1 (6)
Physiotherapy	0 (0)	1 (6)
Psychological therapy	1 (6)	0 (0)
Relaxation	0 (0)	1 (6)

Data presented describe the number and percentage for each characteristic unless otherwise noted.
FU, follow-up; IPD, individual participant data; NSAIDs, non-steroidal anti-inflammatory drugs.

15 trials were partially replicated and required further author contact to confirm the data.

Pain and functional limitation outcomes
One-stage random-effects IPD meta-analysis of the included trials found that on average exercise therapy was more effective
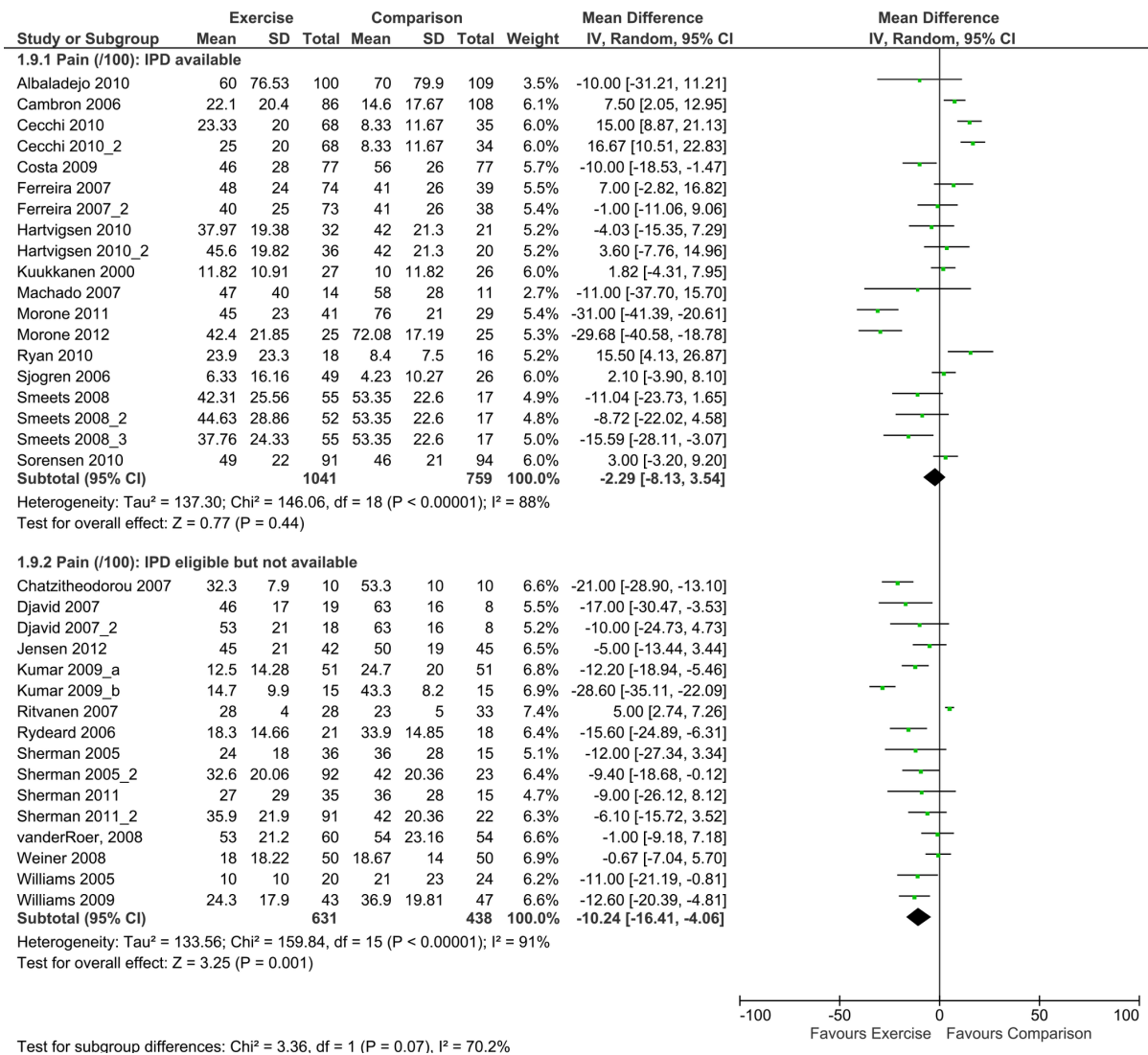
Figure 2 Aggregate meta-analysis results, from published estimates, for studies with IPD available and eligible studies that did not provide data for comparison of exercise therapy with any non-exercise comparisons on pain intensity outcome (/100) at the earliest follow-up period. Syntheses only trial group data with non-exercise comparison (ie, all other conservative and no treatment/usual care comparisons), similar to primary analysis in the related Cochrane review (14 studies/19 groups for IPD and 13 studies/16 groups for IPD eligible, but not available); pain intensity outcome (/100). Trials that include only comparisons of different types of exercise are not included. IPD, individual participant data; IV, inverse variance.

than no treatment or usual care on pain outcome (MD (95% CI) −10.7 (−14.1 to –7.4); 26 studies, 2466 participants) at short-term follow-up, a result most compatible with a clinically important 20% smallest worthwhile effect. Other conservative treatments were found to be more effective than exercise therapy, although again not by a clinically important amount (MD (95% CI) 3.7 (1.3 to 6.0); 26 studies, 2850 participants).

For functional limitations outcome, exercise therapy was associated with a clinically important 23% improvement compared with no treatment or usual care (MD (95% CI) −10.2 (−13.1 to –7.3); 25 studies, 2366 participants) at short-term follow-up. Other conservative treatments were found to be more effective than exercise therapy, although not by clinically important amount (MD (95% CI) 1.9 (0.03 to 3.8); 25 studies, 2778 participants) at short-term follow-up. We observed similar results in meta-analyses of global recovery, where exercise therapy was more effective compared with usual care or no treatment (OR (95% CI) 3.8 (2.6 to 5.7); 25 studies, 2366 participants) and similarly effective as other conservative treatments (OR

(95% CI) 0.9 (0.7 to 1.1); 25 studies, 2777 participants) at short-term follow-up.

Exploring individual characteristics that modify exercise therapy treatment effect

We identified several variables that may modify the treatment effectiveness of exercise therapy relative to non-exercise comparison groups. Observed results were consistent between unadjusted and adjusted models for each of pain (table 4), functional limitations (table 5) and global recovery outcomes (table 6).

For pain outcomes adjusted analyses, figure 3 depicts the MD in pain outcomes of exercise compared with non-exercise comparisons at short-term follow-up for subgroups of participants based on levels of characteristics of potential treatment effect modifiers. We observed clinically important outcome improvement with exercise therapy in participants with the following characteristics (exploratory results): normal BMI (compared with

Table 3 Characteristics of participants with persistent LBP with available follow-up data for IPD analyses (N=3514)

Participant characteristics at baseline	Participants, n	%
Age (mean, 95% CI) (missing=0.4%)	3499	45.7 (45.2 to 46.1)
Sex (missing=0.3%)		
Male	1559	44.5
Female	1945	55.5
High school education (missing=56%)		
High school completed or less	691	45.0
Beyond high school	843	55.0
Current smoker (missing=57%)		
No	980	65.0
Yes	527	35.0
Regularly physically active (missing=47%)		
No	916	49.0
Yes	954	51.0
BMI (mean, 95% CI) (missing=33%)	2352	26.3 (26.1 to 26.5)
History of LBP (missing=75%)		
No	279	31.8
Yes	598	68.2
Sick leave (past 12 months) (missing=49%)		
No	858	47.6
Yes	943	52.4
Work status (missing=22%)		
Unemployed	1076	39.2
Employed	1669	60.8
Heavy physical demands at work (missing=57%)		
No	976	63.9
Yes	552	36.1
General health (0–100; mean, 95% CI) (missing=56%)	1529	57.3 (56.3 to 58.2)
Mental health (0–100; mean, 95% CI) (missing=71%)	1034	63.4 (62.1 to 64.7)
Fear avoidance beliefs (0–100; mean; 95% CI) (missing=66%)	1182	39.4 (38.1 to 40.8)
Social support available (missing=56%)		
No	480	30.8
Yes	1080	69.2
Episode duration (months; median, IQR) (missing=52%)	1692	14.0 (6.0 to 48.0)
Pain intensity (0–100; mean, 95% CI) (missing=0.3%)	3411	53.5 (52.6 to 54.3)
Functional limitations (0–100; mean, 95% CI) (missing=0.8%)	3247	44.6 (43.8 to 45.4)
Any leg pain (missing=39%)		
No	791	36.9
Yes	1354	63.1
Any LBP medication use (missing=44%)		
No	927	47.3
Yes	1032	52.7

Percentage of missing data for each participant characteristic is noted.
Data presented describe the number and percentage for each characteristic unless otherwise noted.
BMI, body mass index; IPD, individual participant data; LBP, low back pain.

underweight and obese), sick leave in the previous 12 months and any medication use for low back pain at baseline.

For functional limitation outcomes adjusted analyses, we found two variables most compatible with lower (improved) functional limitations with receiving exercise treatment relative to non-exercise comparisons for treatment–variable interactions: not having heavy physical demands at work (adjusted MD (aMD) (95% CI) 6.0 (1.0 to 11.0), p=0.019) and any medication use for low back pain (aMD (95% CI) −4.8 (−8.7 to −0.9), p=0.016). Figure 4 presents the MD in functional limitations with exercise therapy compared with other/no treatments for subgroups of participants based on potential treatment effect modifiers. Participants with the following characteristics had better functional limitation outcomes with exercise (exploratory results): no heavy physical demands and any medication use for low back pain at baseline.

For global recovery outcome with adjusted analyses, any medication use for low back pain at baseline was most compatible with improved outcomes from exercise relative to non-exercise comparisons (adjusted OR (95% CI) 1.7 (1.0 to 2.8), p=0.046). Figure 5 displays the size of global recovery treatment effects for subgroups of participants. Participants with the following characteristics reported greater global recovery with exercise compared with other/no treatments (exploratory results): normal BMI (compared with underweight and obese), sick leave in the previous 12 months, no heavy physical demands, longer chronic episode duration of back pain and any medication use for low back pain at baseline.

The mean size and direction of the interaction effect for potential treatment effect modifiers were consistent across moderate-term and long-term follow-up time periods for the following characteristics: no physical demands at work and any medication use for low back pain. Other potential treatment effect modifiers that were important in size and most compatible with positive treatment effect modification, at moderate-term or long-term follow-up, were the following: for pain outcome, beyond high school education, not having a history of low back pain, lower fear avoidance beliefs and shorter episode duration; for functional limitations outcome, any medication use for low back pain; and for global recovery outcome, female sex (online supplementary appendices 3–5).

Secondary analyses to explore potential treatment effect modifiers for exercise therapy compared with no treatment/usual care comparison groups or with other non-exercise conservative comparisons found individuals with lower BMI on average experienced more improvement on all three outcomes with exercise compared with no treatment/usual care comparison groups. Individuals with worse baseline function and any medication use for low back pain had improved functional limitations with exercise compared with either no treatment/usual care comparison groups or other non-exercise conservative comparisons. Compared with non-exercise comparisons, individuals with the following characteristics had better outcomes with exercise treatment: for functional limitation outcome, lower fear avoidance beliefs; and for global recovery outcome, any medication use for low back pain (online supplementary appendices 6–8).

DISCUSSION

In this study, we used original data from 27 randomised controlled trials of exercise therapy to explore individual characteristics and identify subgroups based on participants' likely response to exercise treatment. One-stage random-effects meta-analysis of data from included trials found exercise therapy to be more effective than no treatment or usual care on pain, functional limitations and global recovery outcomes at short-term follow-up, most compatible with a clinically important improvement. Exercise therapy was observed to be similarly effective to

Table 4 Size and statistical significance of treatment–variable interactions (unadjusted and adjusted results) on continuous pain outcome (0–100) at follow-up closest to 3 months*

Variable	Unadjusted				Adjusted			
	Studies	Participants	Interaction MD (95% CI)	P value	Studies	Participants	Interaction MD (95% CI)	P value
Age (years)	26	3091	0.09 (−0.05 to 0.23)	0.22	26	3091	0.09 (−0.05 to 0.23)	0.22
Sex	26	3095			26	3091		
Male			–	–			–	–
Female			−0.98 (−4.49 to 2.53)	0.59			−0.96 (−4.39 to 2.46)	0.58
High school education	12	1401			12	1399		
High school or less			–	–			–	–
Beyond high school			−3.94 (−8.90 to 1.02)	0.12			−3.69 (−8.65 to 1.27)	0.15
Current smoker	9	1411			9	1407		
No			–	–			–	–
Yes			2.07 (−3.15 to 7.29)	0.44			1.38 (−3.86 to 6.63)	0.61
Regular physical activity	11	1695			11	1693		
No			–	–			–	–
Yes			−1.05 (−5.79 to 3.68)	0.66			−0.67 (−5.4 to 4.06)	0.78
BMI	17	2173	0.23 (−0.27 to 0.74)	0.37	17	2170	0.22 (−0.27 to 0.71)	0.38
History of LBP	8	827			8	827		
No			–	–			–	–
Yes			−4.03 (−13.88 to 5.82)	0.42			−3.92 (−13.76 to 5.92)	0.44
Sick leave (past 12 months)	14	1578			14	1575		
No			–	–			–	–
Yes			−2.27 (−9.15 to 4.61)	0.52			−2.15 (−8.92 to 4.62)	0.53
Work status	18	2446			18	2444		
Unemployed			–	–			–	–
Employed			−3.81 (−9.07 to 1.45)	0.16			−4.22 (−9.26 to 0.82)	0.10
Heavy physical demands	8	1388			8	1386		
No			–	–			–	–
Yes			4.45 (−1.84 to 10.74)	0.17			4.92 (−1.35 to 11.19)	0.12
General health (0–100)	10	1319	0.05 (−0.11 to 0.20)	0.54	10	1316	0.05 (−0.11 to 0.21)	0.55
Mental health (0–100)	6	877	−0.05 (−0.19 to 0.10)	0.54	6	874	−0.04 (−0.18 to 0.11)	0.61
Fear avoidance beliefs (0–100)	8	1012	0.08 (−0.03 to 0.20)	0.16	8	1011	0.09 (−0.03 to 0.20)	0.15
Social support	10	1451			10	1450		
No			–	–			–	–
Yes			−1.05 (−5.65 to 3.55)	0.66			−1.38 (−5.97 to 3.20)	0.55
Episode duration (months)	13	1572	0.02 (−0.03 to 0.07)	0.37	13	1569	0.02 (−0.04 to 0.08)	0.58
Functional limitations (0–100)	24	2880	0.01 (−0.08 to 0.11)	0.81	24	2876	0.02 (−0.07 to 0.12)	0.66
Pain intensity (0–100)	26	3095	−0.06 (−0.15 to 0.03)	0.23	26	3091	−0.05 (−0.14 to 0.04)	0.26
Any leg pain	14	1837			14	1833		
No			–	–			–	–
Yes			1.60 (−3.57 to 6.77)	0.54			1.46 (−3.73 to 6.64)	0.58
Any LBP medication use	13	1776			13	1774		
No			–	–			–	–
Yes			−4.04 (−9.66 to 1.58)	0.16			−3.90 (−9.40 to 1.60)	0.16

A negative interaction coefficient indicates decreased pain for the variable level (eg, female) with exercise vs any non-exercise comparison (ie, all other conservative and no treatment/usual care comparisons).

*Unadjusted models include baseline pain, potential treatment effect-modifying variable, treatment group (exercise or comparison treatment), and trial at the random-effects level and the variable–treatment group interaction at the fixed-effects level. Adjusted models add age and sex (at the random-effects level). All participant-level covariates are centred around their trial-specific means.

BMI, body mass index; LBP, low back pain; MD, mean difference.

other included comparison treatments (here including manual therapy, education or psychological therapy) for all outcomes. However, these results should be interpreted cautiously as the trials included in this study may underestimate exercise treatment effect and represent fewer than 10% of the randomised controlled trials now available.

Comparisons of trial characteristics were not noticeably different from eligible trials not providing data. However,

analysis limited to other treatment group comparisons suggested that the average treatment effect for exercise therapy was smaller and not clinically important for studies providing IPD (MD = -2.3 for pain outcome) compared to the effect of treatment in eligible trials not providing IPD (MD = -10.2 for pain outcome).

Our study has provided exploratory evidence that not having heavy physical demands at work and using pain medication are potential treatment effect modifiers for exercise therapy

Table 5 Size and statistical significance of treatment–variable interactions (unadjusted and adjusted results) on continuous functional outcome (0–100) at follow-up closest to 3 months*

Variable	Unadjusted				Adjusted			
	Studies	Participants	Interaction MD (95% CI)	P value	Studies	Participants	Interaction MD (95% CI)	P value
Age (years)	25	2996	0.07 (−0.04 to 0.19)	0.21	25	2996	0.07 (−0.04 to 0.19)	0.21
Sex	25	3000			25	2996		
Male			–	–			–	–
Female			−1.62 (−4.38 to 1.15)	0.25			−1.71 (−4.44 to 1.02)	0.22
High school education	10	1204			10	1202		
High school or less			–	–			–	–
Beyond high school			−0.53 (−4.83 to 3.77)	0.81			−0.24 (−4.54 to 4.05)	0.91
Current smoker	9	1420			9	1416		
No			–	–			–	–
Yes			0.16 (−3.99 to 4.30)	0.94			−0.36 (−4.52 to 3.80)	0.87
Regular physical activity	12	1758			12	1756		
No			–	–			–	–
Yes			−0.96 (−4.59 to 2.67)	0.60			−0.75 (−4.38 to 2.89)	0.69
BMI	15	1984	0.11 (−0.31 to 0.52)	0.61	15	1981	0.11 (−0.30 to 0.52)	0.60
History of LBP	8	830			8	830		
No			–	–			–	–
Yes			0.16 (−8.23 to 8.56)	0.97			0.39 (−7.99 to 8.76)	0.93
Sick leave (past 12 months)	13	1451			13	1448		
No			–	–			–	–
Yes			−0.35 (−6.65 to 5.94)	0.91			−0.03 (−6.24 to 6.17)	0.99
Work status	17	2321			17	2319		
Unemployed			–	–			–	–
Employed			0.96 (−2.93 to 4.85)	0.63			0.58 (−3.24 to 4.40)	0.76
Heavy physical demands	8	1391			8	1389		
No			–	–			–	–
Yes			5.81 (0.73 to 10.88)	0.025			6.02 (1.00 to 11.04)	0.019
General health (0–100)	10	1337	−0.06 (−0.25 to 0.14)	0.56	10	1334	−0.06 (−0.25 to 0.14)	0.56
Mental health (0–100)	7	956	−0.03 (−0.20 to 0.15)	0.75	7	953	−0.02 (−0.19 to 0.15)	0.82
Fear avoidance beliefs (0–100)	9	1110	0.09 (−0.003 to 0.18)	0.059	9	1109	0.09 (−0.002 to 0.18)	0.054
Social support	9	1381			9	1380		
No			–	–			–	–
Yes			2.2 (−2.08 to 6.48)	0.31			1.94 (−2.40 to 6.29)	0.38
Episode duration (months)	13	1573	0.01 (−0.01 to 0.04)	0.29	13	1570	0.02 (−0.01 to 0.04)	0.25
Functional limitations (0–100)	25	3000	−0.08 (−0.18 to 0.02)	0.14	25	2996	−0.06 (−0.16 to 0.03)	0.21
Pain intensity (0–100)	24	2923	0.03 (−0.04 to 0.10)	0.38	24	2919	0.03 (−0.04 to 0.10)	0.36
Any leg pain	15	1950			15	1946		
No			–	–			–	–
Yes			0.86 (−3.21 to 4.94)	0.68			0.54 (−3.53 to 4.61)	0.80
Any LBP medication use	13	1795			13	1793		
No			–	–			–	–
Yes			−5.11 (−9.03 to 1.19)	0.011			−4.81 (−8.72 to 0.90)	0.016

A negative interaction coefficient indicates decreased function for the variable level (eg, female) with exercise vs any non-exercise comparison (ie, all other conservative and no treatment/usual care comparisons).

*Unadjusted models include baseline functional limitations, potential treatment effect-modifying variable, treatment group (exercise or comparison treatment), and trial at the random-effects level and the variable–treatment group interaction at the fixed-effects level. Adjusted models add age and sex (at the random-effects level). All participant-level covariates are centred around their trial-specific means.

BMI, body mass index; LBP, low back pain; MD, mean difference.

outcomes compared with other treatments at short-term follow-up. This indicates that individuals with these characteristics may benefit more from exercise therapy. One could hypothesise that characteristics that may facilitate compliance with an active treatment programme (eg, using medication to alleviate low back pain symptoms, and not having physical demands at work which could lead to strain and/or a flare up of symptoms) may be associated with improved outcomes with exercise compared with other treatments. Lower BMI was consistently associated with improved outcomes from exercise interventions compared with usual care or no treatment at the follow-up period closest to 3 months.

These results suggest two directions for future research to advance management of persistent low back pain. First, further research is needed to validate and extend our findings. We tested many potential effect modifiers, so our findings may be

Table 6 Size and statistical significance of treatment–variable interactions (unadjusted and adjusted results) on important global recovery in pain or functional outcome (dichotomous) at follow-up closest to 3 months*

Variable	Unadjusted				Adjusted			
	Studies (n)	Participants (n)	Interaction OR (95% CI)	P value	Studies (n)	Participants (n)	Interaction OR (95% CI)	P value
Age (years)	25	2995	1.01 (0.99 to 1.02)	0.43	25	2995	1.01 (0.99 to 1.02)	0.43
Sex	25	2995			25	2995		
Male			–	–			–	–
Female			1.33 (0.91 to 1.94)	0.15			1.32 (0.90 to 1.93)	0.16
High school education	10	1204			10	1202		
High school or less			–	–			–	–
Beyond high school			1.32 (0.73 to 2.40)	0.36			1.32 (0.72 to 2.40)	0.37
Current smoker	9	1421			9	1417		
No			–	–			–	–
Yes			0.98 (0.54 to 1.80)	0.96			1.04 (0.56 to 1.93)	0.90
Regular physical activity	12	1758			12	1756		
No			–	–			–	–
Yes			0.94 (0.56 to 1.57)	0.80			0.90 (0.54 to 1.52)	0.70
BMI	15	1986	0.95 (0.89 to 1.00)	0.067	15	1983	0.95 (0.89 to 1.00)	0.054
History of LBP	8	830			8	830		
No			–	–			–	–
Yes			1.05 (0.40 to 2.79)	0.92			1.05 (0.39 to 2.78)	0.93
Sick leave (past 12 months)	13	1450			13	1447		
No			–	–			–	–
Yes			1.33 (0.61 to 2.92)	0.48			1.26 (0.57 to 2.80)	0.56
Work status	17	2319			17	2317		
Unemployed			–	–			–	–
Employed			0.93 (0.56 to 1.55)	0.78			0.94 (0.56 to 1.58)	0.82
Heavy physical demands	8	1393			8	1391		
No			–	–			–	–
Yes			0.56 (0.26 to 1.21)	0.14			0.55 (0.26 to 1.19)	0.13
General health (0–100)	10	1335	1.00 (0.98 to 1.03)	0.88	10	1332	1.00 (0.98 to 1.03)	0.92
Mental health (0–100)	7	954	0.99 (0.97 to 1.02)	0.55	7	951	0.99 (0.97 to 1.01)	0.46
Fear avoidance beliefs (0–100)	9	1106	1.00 (0.98 to 1.01)	0.69	9	1105	1.00 (0.98 to 1.01)	0.71
Social support	9	1380			9	1379		
No			–	–			–	–
Yes			1.14 (0.66 to 1.97)	0.64			1.17 (0.67 to 2.02)	0.58
Episode duration (months)	13	1575	1.00 (1.00 to 1.01)	0.41	13	1572	1.00 (1.00 to 1.01)	0.43
Functional limitations (0–100)	25	2999	1.01 (1.00 to 1.02)	0.28	25	2995	1.01 (0.99 to 1.02)	0.32
Pain intensity (0–100)	24	2925	1.00 (0.99 to 1.01)	0.79	24	2921	1.00 (0.99 to 1.01)	0.81
Any leg pain	15	1950			15	1946		
No			–	–			–	–
Yes			0.98 (0.54 to 1.78)	0.95			1.01 (0.55 to 1.86)	0.97
Any LBP medication use	13	1795			13	1793		
No			–	–			–	–
Yes			1.72 (1.03 to 2.85)	0.037			1.68 (1.01 to 2.81)	0.046

An interaction coefficient above 1 indicates increased likelihood of recovery for the variable level (eg, female) with exercise vs any non-exercise comparison (ie, all other conservative and no treatment/usual care comparisons).

*Unadjusted models include baseline functional limitations, potential treatment effect-modifying variable, treatment group (exercise or comparison treatment), and trial at the random-effects level and the variable–treatment group interaction at the fixed-effects level. Adjusted models add age and sex (at the random-effects level). All participant-level covariates are centred around their trial-specific means.

BMI, body mass index; LBP, low back pain.

coincidental. Future trials of exercise therapy, including prospectively planned multicentre trials, should consistently measure and test these and other theoretically driven potential treatment effect modifiers. Second, future studies may test incorporation of these characteristics into prediction models to select individuals for exercise treatment. If prediction models are confirmed accurate in future studies, and with alternate strategies for subgroups who do poorly with exercise, then persistent low back

pain outcomes could be improved with more tailoring of treatments received.

IPD meta-analysis is the gold standard for systematic review,[24] and we followed current recommendations for robust analyses. IPD meta-analysis has three key advantages which have benefited this study. First, the availability of data from 27 trials identified through a systematic review and rated as moderate to low risk of bias resulted in a large sample size available to
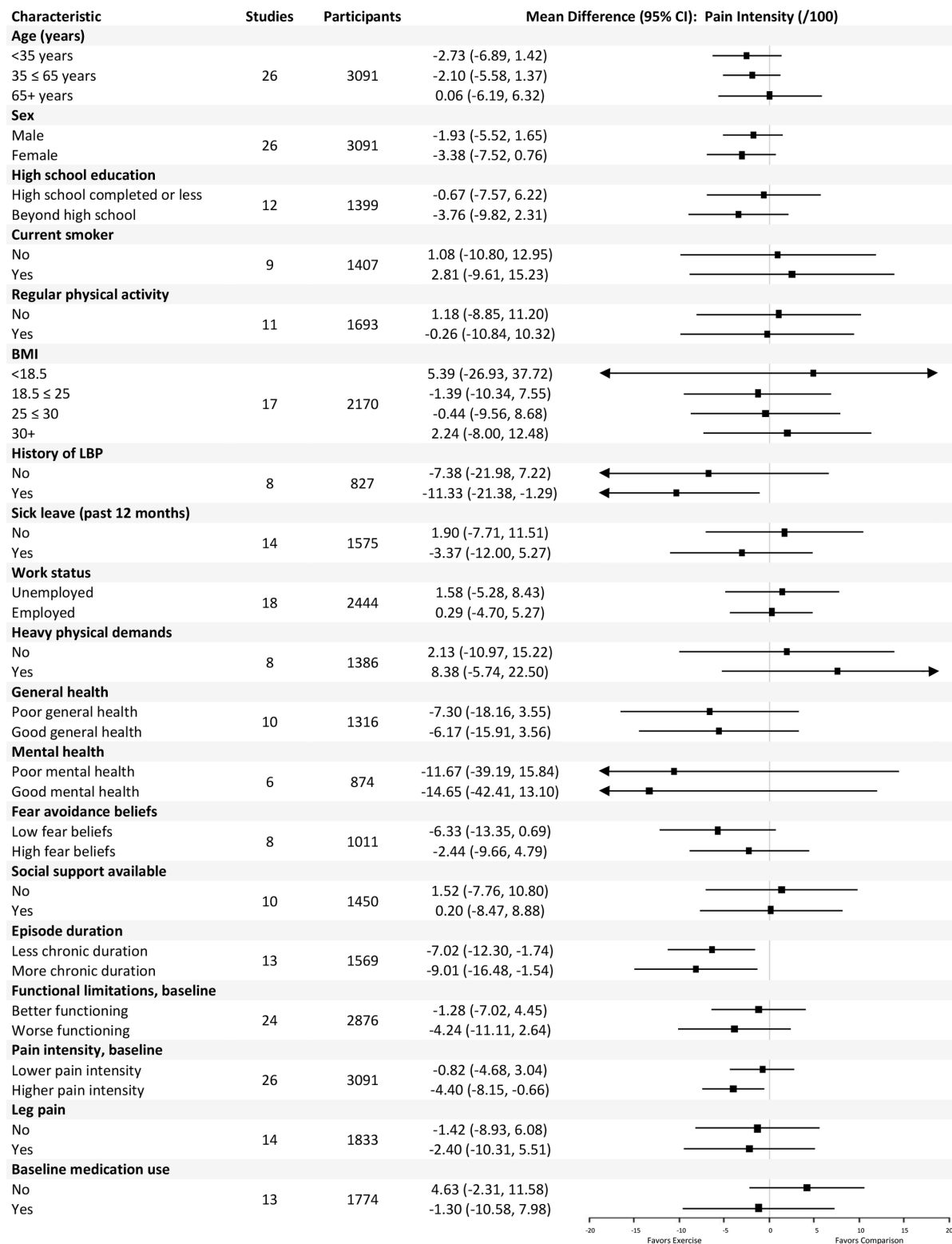
Characteristic	Studies	Participants	Mean Difference (95% CI): Pain Intensity (/100)
Age (years)			
<35 years			-2.73 (-6.89, 1.42)
35 ≤ 65 years	26	3091	-2.10 (-5.58, 1.37)
65+ years			0.06 (-6.19, 6.32)
Sex			
Male	26	3091	-1.93 (-5.52, 1.65)
Female			-3.38 (-7.52, 0.76)
High school education			
High school completed or less	12	1399	-0.67 (-7.57, 6.22)
Beyond high school			-3.76 (-9.82, 2.31)
Current smoker			
No	9	1407	1.08 (-10.80, 12.95)
Yes			2.81 (-9.61, 15.23)
Regular physical activity			
No	11	1693	1.18 (-8.85, 11.20)
Yes			-0.26 (-10.84, 10.32)
BMI			
<18.5			5.39 (-26.93, 37.72)
18.5 ≤ 25	17	2170	-1.39 (-10.34, 7.55)
25 ≤ 30			-0.44 (-9.56, 8.68)
30+			2.24 (-8.00, 12.48)
History of LBP			
No	8	827	-7.38 (-21.98, 7.22)
Yes			-11.33 (-21.38, -1.29)
Sick leave (past 12 months)			
No	14	1575	1.90 (-7.71, 11.51)
Yes			-3.37 (-12.00, 5.27)
Work status			
Unemployed	18	2444	1.58 (-5.28, 8.43)
Employed			0.29 (-4.70, 5.27)
Heavy physical demands			
No	8	1386	2.13 (-10.97, 15.22)
Yes			8.38 (-5.74, 22.50)
General health			
Poor general health	10	1316	-7.30 (-18.16, 3.55)
Good general health			-6.17 (-15.91, 3.56)
Mental health			
Poor mental health	6	874	-11.67 (-39.19, 15.84)
Good mental health			-14.65 (-42.41, 13.10)
Fear avoidance beliefs			
Low fear beliefs	8	1011	-6.33 (-13.35, 0.69)
High fear beliefs			-2.44 (-9.66, 4.79)
Social support available			
No	10	1450	1.52 (-7.76, 10.80)
Yes			0.20 (-8.47, 8.88)
Episode duration			
Less chronic duration	13	1569	-7.02 (-12.30, -1.74)
More chronic duration			-9.01 (-16.48, -1.54)
Functional limitations, baseline			
Better functioning	24	2876	-1.28 (-7.02, 4.45)
Worse functioning			-4.24 (-11.11, 2.64)
Pain intensity, baseline			
Lower pain intensity	26	3091	-0.82 (-4.68, 3.04)
Higher pain intensity			-4.40 (-8.15, -0.66)
Leg pain			
No	14	1833	-1.42 (-8.93, 6.08)
Yes			-2.40 (-10.31, 5.51)
Baseline medication use			
No	13	1774	4.63 (-2.31, 11.58)
Yes			-1.30 (-10.58, 7.98)

Favors Exercise ← | → Favors Comparison
-20 -15 -10 -5 0 5 10 15 20

Figure 3 Treatment effect mean differences in pain outcome at follow-up closest to 3 months for subgroups of potential treatment effect modifiers; for presentation reasons, continuous variables were categorised at clinically relevant cut-points when possible or based on the observed median. We used uncentred values of the potential treatment effect modifier to calculate the mean treatment effect across categories, appropriately adjusted for covariates centred around their trial-specific means. BMI, body mass index; LBP, low back pain.

investigate subgroup effects. Second, we were able to attain consistent presentation of data; direct derivation of information independent of reporting and standardisation of analyses across studies allowed more usable data for meta-analyses. Third, we were able to conduct additional analyses to explore heterogeneity (more extensive use of available data to explore trial-level and participant-level factors in meta-analyses, and assessment of the variation in summary effects within participant subgroups to allow better understanding of the effects of exercise treatment).
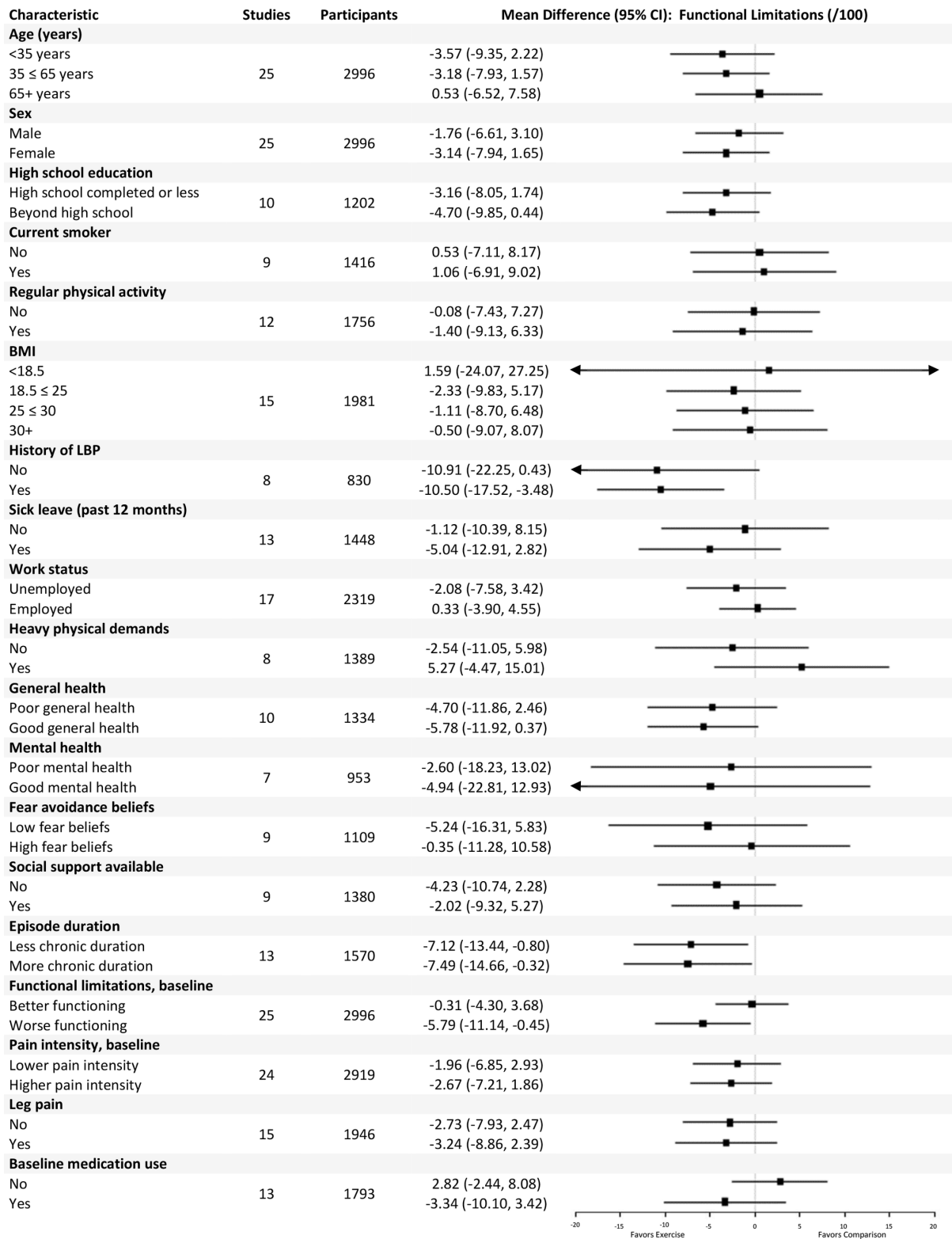
Characteristic	Studies	Participants	Mean Difference (95% CI): Functional Limitations (/100)
Age (years)			
<35 years			-3.57 (-9.35, 2.22)
35 ≤ 65 years	25	2996	-3.18 (-7.93, 1.57)
65+ years			0.53 (-6.52, 7.58)
Sex			
Male	25	2996	-1.76 (-6.61, 3.10)
Female			-3.14 (-7.94, 1.65)
High school education			
High school completed or less	10	1202	-3.16 (-8.05, 1.74)
Beyond high school			-4.70 (-9.85, 0.44)
Current smoker			
No	9	1416	0.53 (-7.11, 8.17)
Yes			1.06 (-6.91, 9.02)
Regular physical activity			
No	12	1756	-0.08 (-7.43, 7.27)
Yes			-1.40 (-9.13, 6.33)
BMI			
<18.5			1.59 (-24.07, 27.25)
18.5 ≤ 25	15	1981	-2.33 (-9.83, 5.17)
25 ≤ 30			-1.11 (-8.70, 6.48)
30+			-0.50 (-9.07, 8.07)
History of LBP			
No	8	830	-10.91 (-22.25, 0.43)
Yes			-10.50 (-17.52, -3.48)
Sick leave (past 12 months)			
No	13	1448	-1.12 (-10.39, 8.15)
Yes			-5.04 (-12.91, 2.82)
Work status			
Unemployed	17	2319	-2.08 (-7.58, 3.42)
Employed			0.33 (-3.90, 4.55)
Heavy physical demands			
No	8	1389	-2.54 (-11.05, 5.98)
Yes			5.27 (-4.47, 15.01)
General health			
Poor general health	10	1334	-4.70 (-11.86, 2.46)
Good general health			-5.78 (-11.92, 0.37)
Mental health			
Poor mental health	7	953	-2.60 (-18.23, 13.02)
Good mental health			-4.94 (-22.81, 12.93)
Fear avoidance beliefs			
Low fear beliefs	9	1109	-5.24 (-16.31, 5.83)
High fear beliefs			-0.35 (-11.28, 10.58)
Social support available			
No	9	1380	-4.23 (-10.74, 2.28)
Yes			-2.02 (-9.32, 5.27)
Episode duration			
Less chronic duration	13	1570	-7.12 (-13.44, -0.80)
More chronic duration			-7.49 (-14.66, -0.32)
Functional limitations, baseline			
Better functioning	25	2996	-0.31 (-4.30, 3.68)
Worse functioning			-5.79 (-11.14, -0.45)
Pain intensity, baseline			
Lower pain intensity	24	2919	-1.96 (-6.85, 2.93)
Higher pain intensity			-2.67 (-7.21, 1.86)
Leg pain			
No	15	1946	-2.73 (-7.93, 2.47)
Yes			-3.24 (-8.86, 2.39)
Baseline medication use			
No	13	1793	2.82 (-2.44, 8.08)
Yes			-3.34 (-10.10, 3.42)

Figure 4 Treatment effect mean difference in functional outcome at follow-up closest to 3 months for subgroups of potential treatment effect modifiers; for presentation reasons, continuous variables were categorised at clinically relevant cut-points when possible or based on the observed median. We used uncentred values of the potential treatment effect modifier to calculate the mean treatment effect across categories, appropriately adjusted for covariates centred around their trial-specific means. BMI, body mass index; LBP, low back pain.

A limitation of our IPD study is the small sample size of included trials. Small trials, common with low back pain treatment studies, are not individually powered to detect a meaningful treatment or moderating effect, may be of lower quality, or reflect publication bias. A benefit of meta-analysis is providing sufficient power through synthesis. However, inclusion of invalid trials in our study may have led to misleading results, particularly related to overall treatment effect. We addressed study
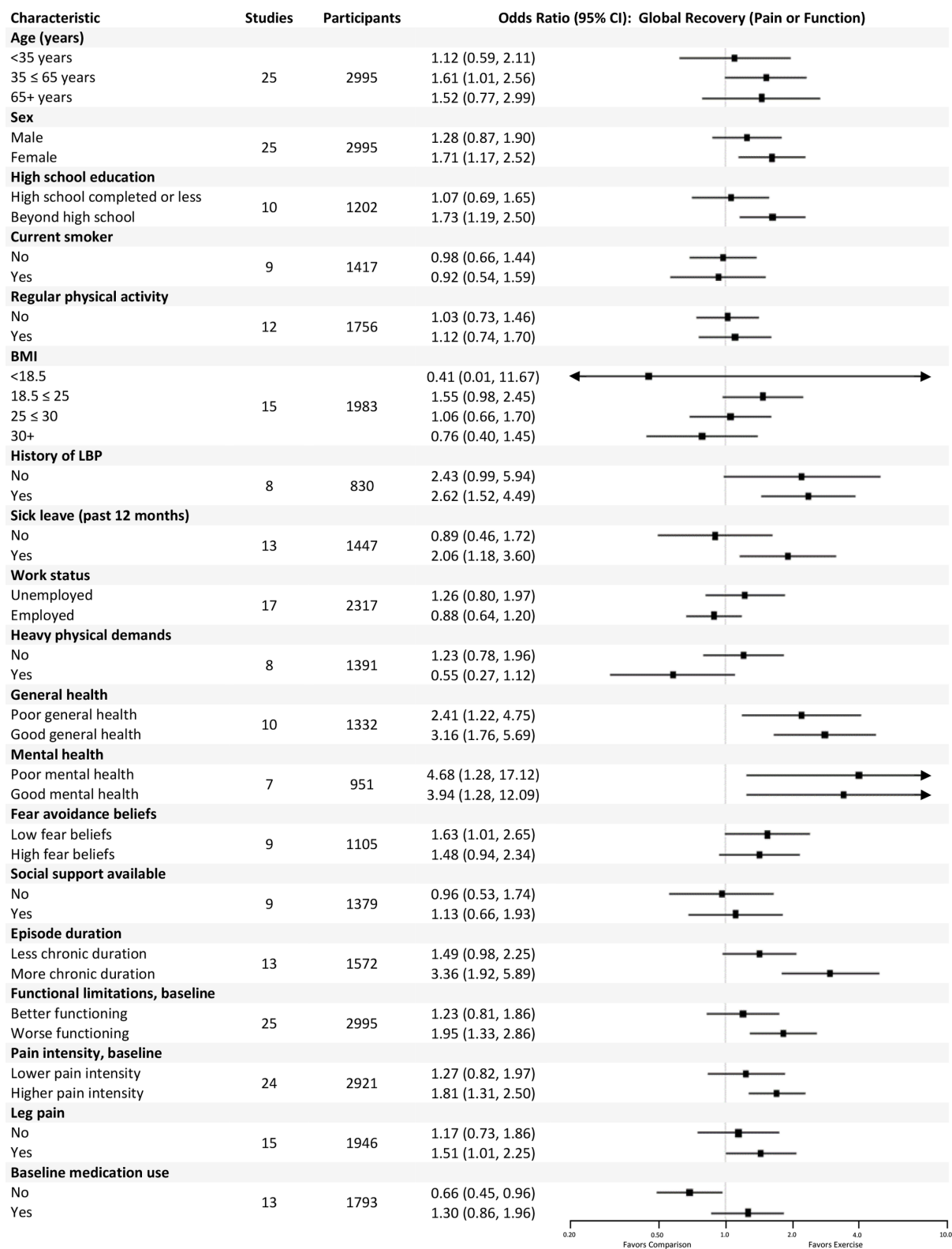
Characteristic	Studies	Participants	Odds Ratio (95% CI): Global Recovery (Pain or Function)
Age (years)			
<35 years			1.12 (0.59, 2.11)
35 ≤ 65 years	25	2995	1.61 (1.01, 2.56)
65+ years			1.52 (0.77, 2.99)
Sex			
Male			1.28 (0.87, 1.90)
Female	25	2995	1.71 (1.17, 2.52)
High school education			
High school completed or less			1.07 (0.69, 1.65)
Beyond high school	10	1202	1.73 (1.19, 2.50)
Current smoker			
No			0.98 (0.66, 1.44)
Yes	9	1417	0.92 (0.54, 1.59)
Regular physical activity			
No			1.03 (0.73, 1.46)
Yes	12	1756	1.12 (0.74, 1.70)
BMI			
<18.5			0.41 (0.01, 11.67)
18.5 ≤ 25	15	1983	1.55 (0.98, 2.45)
25 ≤ 30			1.06 (0.66, 1.70)
30+			0.76 (0.40, 1.45)
History of LBP			
No			2.43 (0.99, 5.94)
Yes	8	830	2.62 (1.52, 4.49)
Sick leave (past 12 months)			
No			0.89 (0.46, 1.72)
Yes	13	1447	2.06 (1.18, 3.60)
Work status			
Unemployed			1.26 (0.80, 1.97)
Employed	17	2317	0.88 (0.64, 1.20)
Heavy physical demands			
No			1.23 (0.78, 1.96)
Yes	8	1391	0.55 (0.27, 1.12)
General health			
Poor general health			2.41 (1.22, 4.75)
Good general health	10	1332	3.16 (1.76, 5.69)
Mental health			
Poor mental health			4.68 (1.28, 17.12)
Good mental health	7	951	3.94 (1.28, 12.09)
Fear avoidance beliefs			
Low fear beliefs			1.63 (1.01, 2.65)
High fear beliefs	9	1105	1.48 (0.94, 2.34)
Social support available			
No			0.96 (0.53, 1.74)
Yes	9	1379	1.13 (0.66, 1.93)
Episode duration			
Less chronic duration			1.49 (0.98, 2.25)
More chronic duration	13	1572	3.36 (1.92, 5.89)
Functional limitations, baseline			
Better functioning			1.23 (0.81, 1.86)
Worse functioning	25	2995	1.95 (1.33, 2.86)
Pain intensity, baseline			
Lower pain intensity			1.27 (0.82, 1.97)
Higher pain intensity	24	2921	1.81 (1.31, 2.50)
Leg pain			
No			1.17 (0.73, 1.86)
Yes	15	1946	1.51 (1.01, 2.25)
Baseline medication use			
No			0.66 (0.45, 0.96)
Yes	13	1793	1.30 (0.86, 1.96)

0.20   0.50   1.0   2.0   4.0   10.0
Favors Comparison   Favors Exercise

Figure 5 Treatment effect global recovery outcome at follow-up closest to 3 months for subgroups of potential treatment effect modifiers; for presentation reasons, continuous variables were categorised at clinically relevant cut-points when possible or based on the observed median. We used uncentred values of the potential treatment effect modifier to calculate the mean treatment effect across categories, appropriately adjusted for covariates centred around their trial-specific means. BMI, body mass index; LBP, low back pain.

internal validity by selecting trials judged to not be at high risk of bias using the Cochrane Back and Neck group recommended criteria, but may have missed other sources of bias. We do not think that systematic bias related to our primary treatment effect

modification results is likely; however, our results should still be interpreted cautiously. A challenge of smaller studies that should be considered by future researchers undertaking IPD meta-analyses relates to feasibility. Small studies add little information

relative to the time required to test, map and include their data. However, this should be balanced against enhanced generalisability of results with trials representing real heterogeneity in populations and exercise interventions.

A second limitation of this study was the inconsistent availability and measurement of some individual characteristics, limiting the ability to assess all potential treatment effect modifiers with the most valid, reliable continuous measures. Although the overall sample size was large, some potential treatment effect modifiers were available and measured consistently in only a small subset of studies. For example, 8 trials with 1386 participants provided usable data on heavy physical demands at work, and 13 trials with 1774 participants provided usable data on use of pain medication at baseline, analysed as a dichotomous measure, including analgesic, anti-inflammatory or opioids. Almost 40% of trials compared different exercise types, with no non-exercise comparison available.

We were unable to explore treatment effect modifiers for specific exercise therapies, non-exercise comparisons, or for no treatment separate from placebo or usual care comparisons due to insufficient homogeneous types across included trials. Furthermore, we were unable to investigate some potential treatment effect modifiers we had originally planned due to low availability across studies, including presence and number of comorbidities, alcohol use and socioeconomic status. While we only received data from approximately half of requested trials and the observed effect of exercise versus other comparisons was smaller for pain outcomes, trial-level characteristics did not significantly differ between those received and requested. A commonly stated benefit of IPD meta-analysis is a more consistent presentation of data and approach to analysis across included trials, allowing for more homogeneity. However, this is achievable only if the necessary participant characteristics and outcome variables are reported; in our study we did not find our IPD meta-analysis to have lower heterogeneity than previous aggregate meta-analyses. A limitation of our IPD analysis includes our assumption of linear interactions for continuous variables, which may have missed non-linear relationships. Finally, a challenge of the IPD approach was the considerable amount of time and effort that was involved in gathering, testing and compiling data from individual studies, which were published before 2013. However, we think that it is unlikely that newer trial data would be different in treatment effect modification results, which is the focus of this project.

Patients with low back pain are heterogeneous and the treatment is complex. We will need large data sets of reliably and consistently measured variables to better understand treatment effect modifiers and identify relevant treatment subgroups for exercise overall and for specific types of exercise (ie, yoga, aerobic exercise and so on). Specifically, the factors that we identified in our study to be potential treatment effect modifiers should be further investigated. Future trials should measure a comprehensive set of variables to define potential subgroups, evaluate treatment effect modification and include non-exercise comparisons. This is unlikely to be feasible with small individual studies, so it will need to be facilitated by increased international collaboration, prospective planning of multisite and multicountry trials, standardising measurement of prognostic factors, and potentially by sharing of data through accessible repositories. Future prospective coordination and collaboration for more consistent data collection will help researchers identify treatment effect modifiers. This will further advance a personalised management approach for persistent low back pain.

CONCLUSIONS

Our IPD meta-analysis combined data from 27 randomised trials, which allowed us to examine a large sample with consistent data. We assessed the effectiveness of exercise therapy to provide context to our study and explored the impact of potential treatment effect modifiers. In our sample, exercise therapy was minimally effective for persistent non-specific low back pain outcomes, and it appears that for individuals using medication for low back pain, and possibly for those with no heavy physical demands at work, they may benefit more from exercise than other treatments. This study provides potentially useful information to help design future studies of exercise interventions that are better matched to specific subgroups.

What is already known

► While there is no consensus on the best course of treatment, exercise therapy is on average moderately effective for persistent low back pain and is recommended by clinical guidelines.
► To choose the most appropriate care for patients, it is important to understand which individual characteristics (eg, work status, pain medication use) are associated with better or worse treatment outcomes.

What are the new findings

► The research team obtained data sets for 27 randomised controlled trials, each of which examined the impact of various forms of exercise on pain or function for people with persistent low back pain; trial data were merged into a large data set of 3514 individuals and were analysed.
► Exercise therapy was more effective than other treatments for people who did not have heavy physical demands at work and who used medication to treat low back pain.
► Future studies of exercise therapy should prospectively test the modifying effect of factors identified in this study, and other untested factors, alone and in combination.

Author affiliations
[1]Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada
[2]Research Institute for Primary Care and Health Sciences, Keele University, Keele, UK
[3]Department of Health Sciences, VU University, Amsterdam, The Netherlands
[4]Center for Rheumatology and Spine Diseases, Rigshospitalet, and University of Copenhagen, Copenhagen, Denmark
[5]Fondazione Don Carlo Gnocchi, Scientific Institute, Florence, Italy
[6]Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de Sao Paulo, Sao Paulo, Brazil
[7]Department of Rheumatology, A2 Rheumatology and Sports Medicine, Hillerød, Denmark
[8]Institute of Bone and Joint Research, The Kolling Institute, University of Sydney, Sydney Medical School, Sydney, New South Wales, Australia
[9]Primary Care Centre Versus Arthritis, School for Primary, Community and Social Care, Keele University, Keele, UK
[10]College of Chiropractic Medicine, Keiser University, West Palm Beach, Florida, USA
[11]Department of Sports Science and Clinical Biomechanics, University of Southern Denmark; Nordic Institute of Chiropractic and Clinical Biomechanics, Odense, Denmark
[12]Department of Training Medicine and Training Physiology, Personnel Command, Royal Netherlands Army, Utrecht, The Netherlands
[13]Research Department and Department of Rheumatology, Rehabilitation Centre Valens, Valens, Switzerland
[14]Centre for Rehabilitation Science, University of Manchester, Manchester Royal Infirmary, Manchester, UK
[15]Unidad de la Espalda Kovacs, Hospital Universitario HLA-Moncloa, Red Española de Investigadores en Dolencias de la Espalda, Madrid, Spain

[16]School of Health and Social Studies, JAMK University of Applied Sciences, Jyväskylä, Finland
[17]Bonavista Physical Therapy, Calgary, Alberta, Canada
[18]School of Rehabilitation Science, McMaster University, Hamilton, Ontario, Canada
[19]University Hospital, UFMG-Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
[20]Institute for Musculoskeletal Health, Sydney School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, New South Wales, Australia
[21]Department of Family and Community Medicine, University of California San Francisco, San Francisco, California, USA
[22]Clinical Laboratory of Experimental Neurorehabilitation, Santa Lucia Foundation, Rome, Italy
[23]Back Center Copenhagen, Copenhagen, Denmark
[24]Department of Neurobiology Care Sciences, and Society, Division of Physiotherapy, Karolinska Institutet, Huddinge, Sweden
[25]School of Health and Life Sciences, Teesside University, Middlesbrough, UK
[26]Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland
[27]Department of Rehabilitation Medicine, Research School of CAPHRI, Maastricht University, Maastricht, The Netherlands
[28]Radboud Institute for Health Sciences, IQ Healthcare, Radboud University Medical Center, Nijmegen, The Netherlands
[29]Department of Neuromedicine and Movement Science, and Department of Public Health and Nursing, NTNU, Trondheim, Norway
[30]Discipline of Physiotherapy, School of Allied Health, Human Services and Sport, College of Science, Health and Engineering, La Trobe University, Melbourne, Victoria, Australia
[31]Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Kowloon, Hong Kong

Twitter Leonardo O P Costa @lcos3060, Luciana A C Machado @machadolac, Rob Smeets @smeets1964 and Ella W Yeung @ella_yeung

ORCID iDs
Jill A Hayden http://orcid.org/0000-0001-7026-144X
Leonardo O P Costa http://orcid.org/0000-0003-3309-5619
Jan Hartvigsen https://orcid.org/0000-0002-5876-7410
George A Koumantakis https://orcid.org/0000-0001-7709-2802
Francisco M Kovacs https://orcid.org/0000-0002-2984-4800
Luciana G Macedo https://orcid.org/0000-0002-1840-2951
Luciana A C Machado https://orcid.org/0000-0001-6303-2753
Chris G Maher https://orcid.org/0000-0002-1628-7857
Wolf Mehling https://orcid.org/0000-0002-0932-9844
Giovanni Morone https://orcid.org/0000-0003-3602-4197
Eva Rasmussen-Barr https://orcid.org/0000-0003-3015-1232
Cormac G Ryan https://orcid.org/0000-0001-5864-4325
Tuulikki Sjögren https://orcid.org/0000-0001-6908-4618
Rob Smeets https://orcid.org/0000-0002-9503-366X
J Bart Staal http://orcid.org/0000-0002-0083-6380
Ella W Yeung http://orcid.org/0000-0002-6396-1146

REFERENCES

1 Vos T, Allen C, Arora M, *et al*. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The Lancet* 2016;388:1545–602.
2 Freburger JK, Holmes GM, Agans RP, *et al*. The rising prevalence of chronic low back pain. *Arch Intern Med* 2009;169:251–8.
3 Hartvigsen J, Hancock MJ, Kongsted A, *et al*. What low back pain is and why we need to pay attention. *The Lancet* 2018;391:2356–67.
4 Hoy D, Brooks P, Blyth F, *et al*. The epidemiology of low back pain. *Best Pract Res Clin Rheumatol* 2010;24:769–81.
5 Hayden JA, Dunn KM, van der Windt DA, *et al*. What is the prognosis of back pain? *Best Pract Res Clin Rheumatol* 2010;24:167–79.
6 Chou R, Huffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American pain Society/American College of physicians clinical practice guideline. *Ann Intern Med* 2007;147:492–504.
7 Foster NE, Hill JC, O'Sullivan P, *et al*. Stratified models of care. *Best Pract Res Clin Rheumatol* 2013;27:649–61.
8 Mistry D, Patel S, Hee SW, *et al*. Evaluating the quality of subgroup analyses in randomized controlled trials of therapist-delivered interventions for nonspecific low back pain. *Spine* 2014;39:618–29.
9 Saragiotto BT, Maher CG, Moseley AM, *et al*. A systematic review reveals that the credibility of subgroup claims in low back pain trials was low. *J Clin Epidemiol* 2016;79:3–9.
10 Hayden JA, Cartwright J, van Tulder MW, *et al*. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev* 2012;25.
11 Hayden JA, van Tulder MW, Malmivaara AV, *et al*. Meta-Analysis: exercise therapy for nonspecific low back pain. *Ann Intern Med* 2005;142:765–75.
12 Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med* 2005;142:776–85.
13 Berlin JA, Santanna J, Schmid CH, *et al*. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;21:371–87.
14 Lambert PC, Sutton AJ, Abrams KR, *et al*. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol* 2002;55:86–94.
15 Thompson S, Higgins J. Can meta-analysis help target interventions at individuals most likely to benefit? *The Lancet* 2005;365:341–6.
16 Furlan AD, Malmivaara A, Chou R, *et al*. 2015 updated method guideline for systematic reviews in the Cochrane back and neck group. *Spine* 2015;40:1660–73.
17 Hayden JA, Cartwright JL, Riley RD, *et al*. Exercise therapy for chronic low back pain: protocol for an individual participant data meta-analysis. *Syst Rev* 2012;1:64.
18 van Tulder M, Furlan A, Bombardier C, *et al*. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine* 2003;28:1290–9.
19 Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ* 2011;344:d7762.
20 Salaffi F, Stancati A, Silvestri CA, *et al*. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain* 2004;8:283–91.
21 Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. low back pain: outcome measures. *J Rheumatol* 2001;28:431–8.
22 Kovacs FM, Abraira V, Royuela A, *et al*. Minimal clinically important change for pain intensity and disability in patients with nonspecific low back pain. *Spine* 2007;32:2915–20.
23 Ostelo RWJG, Deyo RA, Stratford P, *et al*. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 2008;33:90–4.
24 Riley RD, Lambert PC, Abo-Zaid G. Meta-Analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340:c221.
25 Riley RD, Lambert PC, Staessen JA, *et al*. Meta-Analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med* 2008;27:1870–93.
26 Hingorani AD, Windt DAvd, Riley RD, *et al*. Prognosis research strategy (progress) 4: stratified medicine research. *BMJ* 2013;346:e5793.
27 Riley RD, Hayden JA, Steyerberg EW, *et al*. Prognosis research strategy (progress) 2: prognostic factor research. *PLoS Med* 2013;10:e1001380.
28 Fisher DJ, Carpenter JR, Morris TP, *et al*. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;356.

29 Hua H, Burke DL, Crowther MJ, *et al*. One-Stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. *Stat Med* 2017;36:772–89.

30 Ferreira ML, Herbert RD, Ferreira PH, *et al*. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit–harm trade-off study. *J Clin Epidemiol* 2013;66:1397–404.

31 Christiansen DH, de Vos Andersen N-B, Poulsen PH, *et al*. The smallest worthwhile effect of primary care physiotherapy did not differ across musculoskeletal pain sites. *J Clin Epidemiol* 2018;101:44–52.

32 Stata Statistical Software. *Release 14 version*. College Station, TX, USA: StataCorp, 2015.

33 Manager R. *RevMan). version 5.3 version*. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration, 2014.

34 IBM Corp. *SPSS statistics for windows. version 24.0 ED*. Armonk, NY: IBM Corp, 2016.

35 SAS Institute Inc. *SAS*. Cary, NC, USA: SAS Institute Inc, 2015.

36 Albaladejo C, Kovacs FM, Royuela A, *et al*. The efficacy of a short education program and a short physiotherapy program for treating low back pain in primary care: a cluster randomized trial. *Spine* 2010;35:483–96.

37 Cecchi F, Molino-Lova R, Chiti M, *et al*. Spinal manipulation compared with back school and with individually delivered physiotherapy for the treatment of chronic low back pain: a randomized trial with one-year follow-up. *Clin Rehabil* 2010;24:26–36.

38 Chan CW, Mok NW, Yeung EW. Aerobic exercise training in addition to conventional physiotherapy for chronic low back pain: a randomized controlled trial. *Arch Phys Med Rehabil* 2011;92:1681–5.

39 Costa LOP, Maher CG, Latimer J, *et al*. Motor control exercise for chronic low back pain: a randomized placebo-controlled trial. *Phys Ther* 2009;89:1275–86.

40 Dufour N, Thamsborg G, Oefeldt A, *et al*. Treatment of chronic low back pain: a randomized, clinical trial comparing group-based multidisciplinary biopsychosocial rehabilitation and intensive individual therapist-assisted back muscle strengthening exercises. *Spine* 2010;35:469–76.

41 Ferreira ML, Ferreira PH, Latimer J, *et al*. Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. *Pain* 2007;131:31–7.

42 Macedo LG, Latimer J, Maher CG, *et al*. Effect of motor control exercises versus graded activity in patients with chronic nonspecific low back pain: a randomized controlled trial. *Phys Ther* 2012;92:363–77.

43 Gudavalli MR, Cambron JA, McGregor M, *et al*. A randomized clinical trial and subgroup analysis to compare flexion–distraction with active exercise for chronic low back pain. *Eur Spine J* 2006;15:1070–82.

44 Hartvigsen J, Morsø L, Bendix T, *et al*. Supervised and non-supervised Nordic walking in the treatment of chronic low back pain: a single blind randomized clinical trial. *BMC Musculoskelet Disord* 2010;11:30.

45 Helmhout PH, Harts CC, Staal JB, *et al*. Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. *Eur Spine J* 2004;13:537–47.

46 Kool J, Bachmann S, Oesch P, *et al*. Function-Centered rehabilitation increases work days in patients with nonacute nonspecific low back pain: 1-year results from a randomized controlled trial. *Arch Phys Med Rehabil* 2007;88:1089–94.

47 Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus General exercise versus general exercise only: randomized controlled trial of patients with recurrent low back pain. *Phys Ther* 2005;85:209–25.

48 Kuukkanen TM, Mälkiä EA. An experimental controlled study on postural sway and therapeutic exercise in subjects with low back pain. *Clin Rehabil* 2000;14:192–202.

49 Long A, Donelson R, Fung T. Does it matter which exercise? A randomized control trial of exercise for low back pain. *Spine* 2004;29:2593–602.

50 Machado LAC, Azevedo DC, Capanema MB, *et al*. Client-centered therapy vs exercise therapy for chronic low back pain: a pilot randomized controlled trial in Brazil. *Pain Med* 2007;8:251–8.

51 Mehling WE, Hamel KA, Acree M, *et al*. Randomized, controlled trial of breath therapy for patients with chronic low-back pain. *Altern Ther Health Med* 2005;11:44–52.

52 Morone G, Iosa M, Paolucci T, *et al*. Efficacy of perceptive rehabilitation in the treatment of chronic nonspecific low back pain through a new tool: a randomized clinical study. *Clin Rehabil* 2012;26:339–50.

53 Morone G, Paolucci T, Alcuri MR, *et al*. Quality of life improved by multidisciplinary back school program in patients with chronic non-specific low back pain: a single blind randomized controlled trial. *Eur J Phys Rehabil Med* 2011;47:533–41.

54 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;34:221–8.

55 Ryan CG, Gray HG, Newton M, *et al*. Pain biology education and exercise classes compared to pain biology education alone for individuals with chronic low back pain: a pilot randomised controlled trial. *Man Ther* 2010;15:382–7.

56 Sjögren T, Nissinen KJ, Järvenpää SK, *et al*. Effects of a workplace physical exercise intervention on the intensity of low back symptoms in office workers: a cluster randomized controlled cross-over design. *J Back Musculoskelet Rehabil* 2006;19:13–24.

57 Smeets RJEM, Vlaeyen JWS, Hidding A, *et al*. Chronic low back pain: physical training, graded activity with problem solving training, or both? the one-year post-treatment results of a randomized controlled trial. *Pain* 2008;134:263–76.

58 Sorensen PH, Bendix T, Manniche C, *et al*. An educational approach based on a non-injury model compared with individual symptom-based physical training in chronic LBP. A pragmatic, randomised trial with a one-year follow-up. *BMC Musculoskelet Disord* 2010;11:212.

59 Staal JB, Hlobil H, Twisk JWR, *et al*. Graded activity for low back pain in occupational health care: a randomized, controlled trial. *Ann Intern Med* 2004;140:77–84.

60 Unsgaard-Tøndel M, Fladmark AM, Salvesen Øyvind, *et al*. Motor control exercises, sling exercises, and general exercises for patients with chronic low back pain: a randomized controlled trial with 1-year follow-up. *Phys Ther* 2010;90:1426–40.

61 Wajswelner H, Metcalf B, Bennell K. Clinical Pilates versus general exercise for chronic low back pain: randomized trial. *Med Sci Sports Exerc* 2012;44:1197–205.

62 Paolucci T, Morone G, Iosa M, *et al*. Psychological features and outcomes of the back school treatment in patients with chronic non-specific low back pain. A randomized controlled study. *Eur J Phys Rehabil Med* 2012;48:245–53.

63 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. *Spine* 2011;36:1999–2010.

64 Smeets RJEM, Vlaeyen JWS, Hidding A, *et al*. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. *BMC Musculoskelet Disord* 2006;7:5.

65 Staal JB, Hlobil H, Köke AJA, *et al*. Graded activity for workers with low back pain: who benefits most and how does it work? *Arthritis Rheum* 2008;59:642–9.

66 Vasseljen O, Unsgaard-Tøndel M, Westad C, *et al*. Effect of core stability exercises on feed-forward activation of deep abdominal muscles in chronic low back pain. *Spine* 2012;37:1101–8.

67 Cambron JA, Gudavalli MR, Hedeker D, *et al*. One-Year follow-up of a randomized clinical trial comparing flexion distraction with an exercise program for chronic low-back pain. *J Altern Complement Med* 2006;12:659–68.

68 Ferreira ML, Ferreira PH, Latimer J, *et al*. Relationship between spinal stiffness and outcome in patients with chronic low back pain. *Man Ther* 2009;14:61–7.

69 Macedo LG, Maher CG, Hancock MJ, *et al*. Predicting response to motor control exercises and graded activity for patients with low back pain: Preplanned secondary analysis of a randomized controlled trial. *Phys Ther* 2014;94:1543–54.

70 Helmhout PH, Harts CC, Staal JB, *et al*. Rationale and design of a multicenter randomized controlled trial on a 'minimal intervention' in Dutch army personnel with nonspecific low back pain [ISRCTN19334317]. *BMC Musculoskelet Disord* 2004;5:40.

71 Kool JP, Oesch PR, Bachmann S, *et al*. Increasing days at work using function-centered rehabilitation in nonacute nonspecific low back pain: a randomized controlled trial. *Arch Phys Med Rehabil* 2005;86:857–64.

72 Kuukkanen T, Mälkiä E, Kautiainen H, *et al*. Effectiveness of a home exercise programme in low back pain: a randomized five-year follow-up study. *Physiother Res Int* 2007;12:213–24.