The effects of Pilates exercise in comparison to other forms of exercise on pain and disability in individuals within chronic non-specific low back pain: A systematic review with meta-analysis

The effects of Pilates exercise in comparison to other forms of exercise on pain and disability in individuals within chronic non-specific low back pain: A systematic review with meta-analysis

Objective: To compare the effects of Pilates exercise (PE) with other forms of exercise on pain and disability in individuals with chronic non-specific low back pain (CNSLBP) and to inform clinical practice and future research.

Study design: Systematic review with meta-analysis conducted and reported in line the Preferred Reporting Items for Systematic review and Meta-analysis (PRISMA).

Literature search: Six electronic databases were searched from inception to April 2021.

Study selection criteria: Randomised controlled trials (RCTs) comparing the effect of PE with other forms of exercise for adults with CNSLBP on pain and disability

Data synthesis: Two reviewers assessed the risk of bias of the trials, guided by the Cochrane RoB2 tool. Available data were extracted for meta-analysis with subgroup analysis. PE was compared to general exercise (GE), direction-specific exercise (DSE) and spinal stabilisation exercise (SSE). Certainty of evidence was interpreted following the Grading of Recommendations Assessment, Development and Evaluation approach.

Results: Eleven RCTs were included. A low certainty of evidence supported PE was more effective than GE in pain reduction (Effect size (ES) 0.44). Moreover, very low levels of certainty were revealed for effectiveness of PE compared with DSE for pain reduction (ES 0.65) and equivalence of PE and SSE for pain and disability.

Conclusions: This review found no strong evidence for using one type of exercise intervention over another when managing patients with CNSLBP. Existing evidence does not allow this review to draw definitive recommendations. In the absence of a superior exercise form clinicians should work collaboratively with the patient, using the individual's goals and preferences to guide exercise selection. Further appropriately designed research is warranted to explore this topic further.

Keywords: Low back pain, Exercise therapy, Rehabilitation, Systematic review

Introduction

Low back pain (LBP) is a prevalent cause of disability worldwide, a challenge for healthcare systems and a significant social problem (Vos et al., 2020). Chronic non-specific low back pain (CNSLBP) is characterised as LBP without a definite pathological cause lasting more than 12 weeks and is estimated to account for more than 80% of all chronic LBP (Maher, Underwood, & Buchbinder, 2017). CNSLBP generates approximately 80% of the direct cost of LBP (Eliks, Zgorzalewicz-Stachowiak, & Zeńczak-Praga, 2019).

Various interventions have been suggested to manage CNSLBP. Previous reviews have demonstrated exercise training is more effective than non-exercise treatments in reducing pain in CNSLBP (Owen et al., 2020; Searle, Spink, Ho & Chuter, 2015; Yamato et al., 2015). There is a consistent recommendation from various international guidelines (UK, USA and Canada) that the management of CNSLBP should include some forms of exercise therapy (O'Connell, Cook, Wand, & Ward, 2016). Previous systematic reviews concluded Pilates exercise (PE) (Lim, Poh, Low, & Wong, 2011), spinal stabilisation exercise (SSE) (Rackwitz et al., 2006), and general exercise (GE) (with mixed exercise components) (Gordon & Bloxham, 2016) were more effective in reducing pain than non-exercise comparators in CNSLBP.

PE was developed by Joseph Pilates in the early 1900s (Hoffman & Gabel, 2015). Six principles underpin traditional PE. They include (1) centering - activation of the 'core' abdominal and back muscles, (2) concentration - focus and attention on proper performance of the exercise, (3) control - control of the movement and posture, (4) precision - attention to the quality of exercise, (5) breathing - specific breathing rhythm during exercise and (6) flow

- smoothness during and between exercise (Ehsani, Arab, Jaberzadeh, & Salavati, 2016). Moreover, PE places a strong emphasis on the alignment of body posture to achieve a neutral spine and the maintenance of spinal and pelvic stabilisation (Ehsani, Arab, Jaberzadeh, & Salavati, 2016; Owen et al., 2020). Based on these principles, PE has become increasingly popular in rehabilitation settings to support management of CNSLBP (Wells, Kolt, & Bialocerkowski, 2012).

Previous systematic reviews were conducted to investigate the effects of PE over other forms of interventions and exercises (Lin et al., 2016; Miyamoto, Costa, & Cabral, 2013; Patti et al., 2015; Wells et al., 2013, Yamato et al., 2015). While there has been consistent evidence showing exercises are better than minimal interventions, there has been no conclusive evidence for the comparative effectiveness between PE and other forms of exercise in managing CNSLBP (Hayden et al., 2021). More randomized controlled trials (RCTs) comparing PE and other exercises have been published since 2016 (Lin et al., 2016). An update of the evidence base regarding the comparative effectiveness of these exercise interventions is therefore needed. The evaluation of the comparative effectiveness of exercise interventions for CNSLBP can potentially be valuable to inform treatment options in clinical practice. The objectives of this systematic review are to compare the effectiveness of PE with other forms of exercise for CNSLBP in both pain and disability and synthesise current evidence to inform treatment options in clinical practice and future research.

Methods

Eligibility criteria

Published RCTs comparing the effects of PE with other forms of exercise were eligible for inclusion. Non-English and unpublished studies were excluded. Studies including individuals with LBP as a secondary problem from other comorbidities or specific causes (such as scoliosis, systemic inflammatory disease, and trauma) were excluded. Variation of PE was accepted, including PE on a mat or on an apparatus (such as Cadillac and Reformer). Co-interventions were accepted only if they were added into both the experiment group (PE) and comparison group (other forms of exercise).

Information sources

An electronic search was completed in the following databases: MEDLINE Ovid, PEDro, CENTRAL, EMBASE, CINAHL, SPORTDiscus. The reference lists of the included studies were also reviewed. The search was completed in databases from their inception to 20 April 2021.

Search strategy

Sensitivity-maximising strategy for LBP and RCTs recommended by Cochrane was used for main databases (MEDLINE and EMBASE). Search terms “Pilate*” and “Pilates” were used, aiming to search for interventions explicitly named as “Pilate”. The search strategy is summarised using the STARLITE framework (**APPENDIX A)**.

Study Selection

Eligible studies were screened using the selection criteria (framed by PICO search tool) through the abstract and full text. Studies were included only if (1) their participants (18 years of age or older) were symptomatic with non-specific LBP lasting for at least 12 weeks, (2) an exercise named explicitly as ‘Pilates’ was used in the trial, (3) PE and interventions with exercise components were compared in the trial and (4) either pain or disability was measured as an outcome. Study selection was completed independently by two authors (CW and BR) and then compared. Inconsistency was discussed to reach a consensus. Covidence software was used in the process of study selection (Covidence, Australia). Covidence is an online-based software-as-a-service review platform recommended by the Cochrane.

Data collection and items

Data on participant, inclusion and exclusion criteria, description of interventions and reported outcomes were extracted using Covidence software. Responding authors of the trials were contacted if any information required for data analysis was missing. Data collection was performed by one review author (CW). Self-reported outcomes measuring the construct of pain intensity and change of disability directly were considered comparable and extracted (**TABLE 1**).

Risk of bias assessment

Two authors (CW and BR) independently conducted the risk of bias (RoB) assessment using the Cochrane RoB 2 tool (Sterne et al., 2019). Individual judgment was compared, and inconsistency was discussed to reach a consensus. The RoB assessment was guided by the algorithm and handbook which accompanies the Cochrane RoB2 tool (Sterne et al., 2019). Five domains were carefully examined, including randomization process, deviations from the intended intervention (intention-to-treat), missing outcome data, measurement of the outcome and selection of the reported results. Included trials were judged and given ‘low risk’, ‘some concerns’ or ‘high risk’ depending on their methodological quality. More details in **APPENDIX B**.

Effect measures

Since various scales were used in outcomes, standardized mean differences (SMD) with 95% confidence intervals (95% CI) was considered a more appropriate representation of the estimated effects. The effect size calculated with SMD was interpreted as small (0.2), medium (0.5) or large (0.8) effects (Kinney, Eakman, & Graham, 2020). Trials conducted with the same sample were pooled once only to avoid double counting. A positive value of the effect sizes (as shown in SMD) indicated that PE was more effective than the type of exercises being compared in reducing pain or disability.

Synthesis methods

Data synthesis was completed by one author (CW). The mean differences (MD) and standard deviations of the outcomes from trials were extracted. If not available, the MD was calculated by subtracting the baseline values from the post-intervention values whereas the standard deviations were estimated based on the standard error of the mean change (Higgins et al., 2019). Available data were computed in a meta-analysis using RevMan5 with a random-effects model.

Data presented from the trials in a format other than the mean and standard deviation were converted to an estimated value required for the meta-analysis. In cases where median and interquartile range (IQR) were reported it was assumed that the median was an estimate of the mean value whereas the width of IQR was 1.35 times the standard deviation (Higgins et al., 2019). It was noted that the robustness of this conversion method was uncertain and there might be potential errors.

Heterogeneity across studies was examined using the Chi-square test and I^2 statistics. A probability value of less than 0.05 was indicative of significant heterogeneity. The findings of I^2 were interpreted as follows: low heterogeneity (I^2 = 0%-30%), moderate heterogeneity (I^2 = 30-60%), substantial heterogeneity (I^2 = 50%-90%) and high heterogeneity (I^2 = 75%-100%).

Subgroup analysis was planned in case of possible heterogeneity among the included trials based on the characteristics of the exercise interventions in comparison to PE. Sensitivity analysis of the pooled results was performed if the estimate of effects from individual trials deviated significantly from the rest of the estimates.

Reporting bias assessment

Reporting biases from missing results in a synthesis was assessed by the visual representation of funnel plots. The effect sizes (in SMD) for each outcome were plots against the standard error. Publication bias was indicated if an asymmetrical funnel plot was present.

Certainty of evidence

The certainty of the evidence for each outcome was judged based on the GRADE. There were four key domains to determine the level of certainty of evidence. They included the risk of bias, inconsistency, imprecision, and indirectness (Rubinstein et al., 2012). More details can be found in **TABLE 4**.

Results

Study selection

Results of the selection process of eligible studies is reported (**FIGURE 1**). Eleven studies were included for this review. One study based on one sample was published as two separate reports (Brooks, Kennedy, & Marshall, 2012; Marshall, Kennedy, Brooks, & Lonsdale, 2013). Data from these reports are referred to as a single study in RoB assessment and data extraction (Marshall, Kennedy, Brooks, & Lonsdale, 2013). Among the included studies, three were reviewed (Anand, Caroline, Arun, & Gomathi, 2014; Marshall, Kennedy, Brooks, & Lonsdale, 2013; Wajswelner, Metcalf, & Bennell, 2012) by previous systematic reviews on relevant topics (Lin et al., 2016; Miyamoto, Costa, & Cabral, 2013; Patti et al., 2015; Wells et al., 2014; Yamato et al., 2015). There were seven trials which had not been included in pervious pair-wise meta-analysis.

Study characteristics and results

The characteristics reported results and outcomes of individual study are summarized in **TABLE 1** and **TABLE 3**.

Participants

It was noted that the baseline duration of LBP symptoms was only mentioned in four studies, ranging from less than a year to more than 14 years (Bhadauria & Gurudut, 2017; Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mazloum et al., 2018; Wajswelner, Metcalf, & Bennell, 2012). Three notable inconsistencies among the inclusion and exclusion criteria in the studies were identified. Firstly, only two studies explicitly reported the inclusion of LBP participants with or without leg pain (Akodu, Akinbo, & Okonkwo, 2016; Wajswelner, Metcalf, & Bennell, 2012). Three studies excluded LBP individuals with radiculopathy or radiating leg pain (Bhadauria & Gurudut, 2017; Dsa, Rengaramanujam, & Kudchadkar, 2014; Marshall, Kennedy, Brooks, & Lonsdale, 2013). Secondly, only four studies explicitly excluded individuals who previously received physiotherapy or exercise interventions for their LBP (Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mazloum et al., 2018; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012). Thirdly, only three studies mentioned the exclusion of participants who presented with psychological or psychiatric disorders (Anand, Caroline, Arun, & Gomathi, 2014; Bhadauria & Gurudut, 2017; Mazloum et al., 2018).

*Intervention*s

The duration of the PE program ranged from two to eight weeks with an hour in length. Only four studies explicitly mentioned that the Pilates interventions were individualised (Anand, Caroline, Arun, & Gomathi, 2014; Hasanpour-Dehkordi, Dehghani, & Solati, 2017; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012). Most of the included studies did not report the intervention protocols with sufficient information. Essential information such as intensity and compliance of the PE programs were poorly described. Only two studies provided full details of the interventions, including a list of exercises, repetitions, and descriptions (Akodu, Akinbo, & Okonkwo, 2016; Wajswelner, Metcalf, & Bennell, 2012). While some studies introduced the theoretical concept of PE, discrepancies of the concept underpinning PE across studies were observed (Bhadauria & Gurudut, 2017; Dsa, Rengaramanujam, & Kudchadkar, 2014; Kofotolis et al., 2016; Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mostagi et al., 2015).

Comparators

Three groups of exercise were used as comparator interventions in the trials, including (1) General exercise (GE) which included mixed forms of multidirectional and nonspecific exercises, such as stationary bike exercise, floor exercise, bodyweight exercises and lower limb stretching (Anand, Caroline, Arun, & Gomathi, 2014; Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012), (2) Direction specific exercise (DSE) which included exercise protocols with a clear directional bias, such as 'extension-based exercise' or 'McKenzie exercise' (Hasanpour-Dehkordi, Dehghani, & Solati, 2017; Mazloum et al., 2018), and (3) Spinal stabilisation exercise (SSE) which generally included Swiss ball and floor exercises with an emphasis on abdominal bracing/hollowing, and termed 'core stabilisation exercise' or 'lumbar stabilisation exercise' or 'dynamic/trunk strengthening exercise' (Akodu, Akinbo, & Okonkwo, 2016; Bhadauria & Gurudut, 2017; Dsa, Rengaramanujam, & Kudchadkar, 2014; Kofotolis et al., 2016).

It was observed that the operational definitions and differences between PE and SSE were vaguely presented across these studies. Only one study provided sufficient details to demonstrate the clear difference between the interventions of interest (Akodu, Akinbo, & Okonkwo, 2016). For a study to be classified into the subgroup of SSE, it had to be a specific exercise targeting the training to the trunk muscles but not described as PE and did not have any Pilates-related principles involved in the exercise, for example, describing focus on postural alignment control or specific breathing patterns.

Outcomes

Both pain and disability were measured by ten studies, with data analysed from 369 participants and 418 participants in total respectively. One study did not measure pain (Kofotolis et al., 2016) and one study did not measure disability (Hasanpour-Dehkordi, Dehghani, & Solati, 2017). Measurement time points of the outcomes varied across studies.

The trials with SSE, DSE and GE as comparators had their outcome measures at 2-4, 4-6, 6-8 weeks respectively. Data on reported outcomes and associated measurement time points from trials was summarized in **TABLE 3**.

Risk of bias assessment

The results of the RoB assessment for individual studies are shown (**TABLE 2**). Overall, one study was at low risk (Marshall, Kennedy, Brooks, & Lonsdale, 2013) and two studies were with some concerns (Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012). The rest of the included studies were at high risk. The distribution of the RoB assessment by domains was presented (**FIGURE 2**). More details in **APPENDIX B**.

Effects of interventions on pain

Overall, the pooled result favoured PE over other forms of exercise in pain reduction (n = 317, ES 0.55, 95%CI 0.14 to 0.97). However, it was noted that there was moderate heterogeneity ($I^2 = 66\%$). Therefore, the results were further analysed by using subgroup analysis to highlight a more clinically meaningful comparison and to prevent a wash-out effect resulting from heterogeneity among trials. The results of subgroup analysis (**FIGURE 3**) are presented in three categories: (1) PE vs GE, (2) PE vs DSE, and (3) PE vs SSE.

Pilates exercise vs General exercise. Four studies reported data on pain measurements comparing PE with GE (Anand, Caroline, Arun, & Gomathi, 2014; Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012). One study (n=30) showing PE had a better improvement in pain when compared to GE, was excluded due to insufficient information on reported data. (Anand, Caroline, Arun, & Gomathi, 2014) (See **APPENDIX C**). The pooled result from the remaining three studies was highly homogeneous (I2 = 0%) (Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012). It showed that PE achieved a greater effect in pain reduction than GE (n = 173, ES 0.44, 95%CI 0.14 to 0.74). It was noted that one study reported the median and interquartile range, suggestive of the potential skewness of the primary data in that study, and that data collection timepoints in trials ranged from 6-8 weeks (Mostagi et al., 2015). Overall, the evidence has a low to moderate risk of bias of favouring PE in pain reduction over GE in individuals with CNSLBP but should be considered with caution considering the points highlighted above.

Pilates exercise vs Direction-specific exercise. Two studies compared PE with DSE in pain reduction, showing consistent evidence favouring PE over the DSE (n = 55, ES 0.65, 95%CI 0.10 to 1.19) (Hasanpour-Dehkordi, Dehghani, & Solati, 2017; Mazloum et al., 2018). The result of this subgroup was highly homogenous ($I^2 = 0\%$) but at a high risk of bias.

Pilates exercise vs Spinal stabilisation exercise. Three studies reported data on pain reduction of PE compared with SSE (Akodu, Akinbo, & Okonkwo, 2016; Bhadauria & Gurudut, 2017; Dsa, Rengaramanujam, & Kudchadkar, 2014). It is noted that the result from one study was questionable (see **APPENDIX D**) and deviated significantly from the result of the remaining two studies and thus was excluded in the analysis (**FIGURE 3)** (Dsa, Rengaramanujam, & Kudchadkar, 2014). The recomputed pooled result after exclusion (**FIGURE 4**) was consistent with low heterogeneity ($I^2 = 0\%$), showing a similar effect between PE and SSE in pain reduction (n = 56, ES -0.15, 95%CI -0.69 to 0.4). However, the pooled results in this subgroup were based on trials with a moderate to high risk of bias. In summary, the comparative effectiveness between SSE and PE in pain reduction for CNSLBP is unclear.

Effects of interventions on disability

Overall, the pooled result indicated that there was no significant difference between PE and other forms of exercise in improving disability (n = 333, ES 0.21, 95%CI -0.01 to 0.42) with low heterogeneity ($I^2 = 29\%$) indicated. The results were further analysed by using subgroup analysis. The results are also presented (**FIGURE 5**) in three categories: (1) PE vs GE, (2) PE vs DSE and (3) PE vs SSE.

Pilates exercise vs General exercise. Four studies reported data on disability measurements comparing PE with GE (Anand, Caroline, Arun, & Gomathi, 2014; Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012), with the exclusion of one study due to insufficient information (see **APPENDIX C**) (Anand, Caroline, Arun, & Gomathi, 2014). However, this study concluded that PE was superior to GE in improving disability. The remaining three studies were inconsistent for disability improvement in this subgroup with moderate heterogeneity ($I^2 = 40\%$). The pooled result showed that there was no difference between PE and GE in disability improvement (n = 173, ES 0.32, 95%CI -0.09 to 0.76). This result was based on trials with a low to moderate risk of bias.

Pilates exercise vs Direction-specific exercise. There was only one study which reported data on disability improvement, suggesting that PE was equally effective in improving disability when compared to DSE (n = 31, ES 0.51, 95%CI -0.21 to 1.23) (Mazloum et al., 2018). This study was judged to be at high risk of bias.

Pilates exercise vs Spinal stabilisation exercise. Four studies reported data on disability improvement of PE compared with SSE (Akodu, Akinbo, & Okonkwo, 2016; Bhadauria &

Gurudut, 2017; Dsa, Rengaramanujam, & Kudchadkar, 2014; Kofotolis et al., 2016). Data from one study were excluded in this subgroup analysis for this outcome due to questionable data. (See details in **APPENDIX D**) (Dsa, Rengaramanujam, & Kudchadkar, 2014). However, the authors reported that PE achieved a better improvement in disability than SSE in the trial. The data from the remaining three studies with high homogeneity ($I^2 = 0\%$) were pooled. The result indicated that there was no significant difference between SSE and PE on improvement in disability (n = 129, ES -0.07, 95%CI -0.42 to 0.28), supported by studies with a moderate to high risk of bias.

Reporting biases

Publication bias for each outcome was checked and the funnel plots were presented in **FIGURE 6** and **FIGURE 7**. Both funnel plots were symmetrical, offering a visual representation of the absence of significant publication bias. However, it was noted that the small number of included trials may limit the power of such estimate and thus they should be interpreted with caution.

Summary of findings - GRADE level of evidence

Overall, the findings from the comparison between PE and GE for both pain and disability were supported by evidence with a low level of certainty. The findings from the two comparisons of PE versus DSE and PE versus SSE for pain and disability were at a very low level of certainty, mainly downgraded by high risk of bias (more details and grading principles available in **TABLE 4**).

Discussion

The objectives of this review were to compare the effects of PE on pain and disability with other forms of exercise in CNSLBP and (2) to synthesise and update current evidence with seven new RCTs in the relevant topic to inform clinical practice.

Significance of findings

This review revealed PE was more effective than GE (supported by low certainty of evidence with small effect sizes) and DSE (supported by very low certainty of evidence with medium effect sizes) in reducing pain in CNSLBP. PE was also found to be equally effective in reducing pain when comparing to SSE. There was no significant difference between the effect on disability among different types of exercises.

The authors noted that there has been reviews with network meta-analyses published since the start of this review (Hayden et al., 2021; Owen et al., 2020). The results of this review agreed with those from the above reviews, suggesting PE may be chosen over some exercises interventions due to relative effectiveness. While the findings from the above recently publish reviews can be limited from its low certainty of evidence and its methodological bias regarding between-comparison heterogeneity from indirect comparison, this findings from this pair-wise meta-analysis might offer additional evidence and agreement on the relevant topic by direct comparison of exercises interventions.

Several systematic reviews were published to explore the comparative effectiveness of PE with other forms of intervention for CNSLBP (Lin et al., 2016; Miyamoto, Costa, & Cabral, 2013; Wells et al., 2013, Yamato et al., 2015). Only one review performed a meta-analysis to offer quantitative evidence on the comparison of the effects between PE and other forms of interventions (Yamato et al., 2015). However, the finding was limited to studies comparing PE to GE. Hence, the findings of this review offered direct comparison of PE including but not limited to GE, but also other forms of exercises with the consideration of recently published RCTs since 2015.

Comparison between Pilates exercise and General exercise

Among the studies showing the superiority of PE over GE in pain reduction, postural alignment or neutral spine principle was consistently mentioned in the description of PE groups (Marshall, Kennedy, Brooks, & Lonsdale, 2013; Mostagi et al., 2015; Wajswelner, Metcalf, & Bennell, 2012). While the exact reason is not clear, it is possible that the application of postural alignment or neutral spine principle in PE might have contributed to better symptom modification and restoration of motor control than GE in the trials.

The relationship between neutral spine deficit and CNSLBP was established in a previous study (Sheeran et al., 2012). It was suggested that the maintenance of a neutral spine could help reduce pain and improve disability in CNSLBP by avoiding additional loading and strain on the sensitized structures in the low back area (Hemming, Sheeran, Van Deursen, & Sparkes, 2019). Moreover, increased superficial abdominal muscle activity was also found to be associated with CNSLBP (Sheeran et al., 2012). Since the activation of deep trunk muscles such as transverse abdominis (TrA) and deep lumbar multifidus (LM) were suggested to be higher in a neutral spine position, PE with an emphasis on neutral spine might have helped to address the altered motor control presented in CNSLBP (Fujitani, Jiromaru, Kida, & Nomura, 2017; Wong et al., 2019). This was supported by a previous ultrasonographic study, showing higher automatic activation of TrA after motor control exercises than GE in participants with CNSLBP (Hemming, Sheeran, Van Deursen, & Sparkes, 2019). Also, the focus of isolated activation of deep trunk muscles (such as deep LM) in PE was shown to be effective in reducing the overactivation of superficial LM (Massé-Alarie, Beaulieu, Preuss, & Schneider, 2016). This was also supported by another study, pointing out the potential role of motor control training to normalize the overlapped mapping of primary motor cortex networks represented in people with CNSLBP (Brumagne et al., 2019).

Comparison between Pilates exercise and Direction-specific exercises/Spinal Stabilisation exercise

It was noted that there were discrepancies in the breathing patterns and trunk muscle activation technique in the PE used in the trials. Failure to implement these features could potentially explain the non-significant result obtained in the comparison between PE and other exercises. Firstly, precise breathing pattern was one of the core principles underpinning PE (Kim & Lee, 2017). Pilates breathing patterns were shown to significantly increase the activation of TrA and internal oblique muscles when compared to general breathing patterns in abdominal exercise with healthy subjects (Barbosa et al., 2015; Barbosa, Martins, Vitorino, & Barbosa, 2013). However, it was unclear from the included trials whether breathing patterns were implemented as they are recommended in PE. Potential non-adherence to the breathing patterns of PE might have undermined the effect of PE, which contributed to the non-significance results.

Secondly, there was inconsistency among the trials regarding the trunk muscle activation technique used in PE. Some trials used abdominal hollowing (also known as abdominal drawing-in manoeuvre) while others used abdominal bracing as an activation technique. It was shown that the hollowing technique could significantly increase the activation of TrA contraction independently, without increasing the activity of the superficial trunk muscles (such as rectus abdominis and external oblique) in healthy women. In contrast, exercising with the bracing technique was found to significantly increase the activation of superficial trunk muscles (Koh, Cho, & Kim, 2014). Thus, it was questionable whether the results from the included trials truly reflected the effect of PE by using proper activation techniques. This limitation might have made the exercises less distinct to compare, further leading to a non-significant pooled result between PE and another exercise in comparison.

However, it is also possible that the relatively subtle differences between properly implemented PE and SSE techniques are not sufficient to achieve a difference in outcome, or put differently, that they are similar enough in effect to achieve a similar outcome. This is plausible given the aim of both PE and SSE is to stabilise or control movement of the spinal region through activation of the spinal support muscles, and the differences between the other exercise approaches (GE and DSE) and PE are greater.

Implications for clinical practice

While the existing evidence and the findings of this review could only offer uncertain and limited evidence to the superiority of PE over other exercises, comprehensive assessment from a biopsychosocial perspective should also be emphasized to determine the use and justify the indication of a particular form of exercise. The knowledge and skills of the clinicians and the preference of patients regarding exercise intervention should be carefully considered. Clinicians might consider integrating the discussed PE principles into clinical practice to offer more specific training for postural alignment and deep trunk muscle activation to individuals with CNSLBP.

Implications for future research

It was previously suggested that individuals with non-specific LBP were not homogenous in clinical presentation and responsiveness to different treatments (Stolze, Allison, & Childs, 2012). Multiple classification systems were established to classify patients into different clinical subgroups and facilitate the diagnosis, prognosis, or treatment of non-specific LBP (Fairbank et al., 2011). Moreover, a biopsychosocial model was promoted based on the emerging evidence of the interaction between biological and psychosocial factors in LBP (Fersum et al., 2010). Thus, the involvement of psychosocial factors could have added another level of potential heterogeneity among the participants in the trials.

Inconsistencies were noted from the inclusion and exclusion criteria of the included trials, including the presence of leg pain and psychological disorders. It was likely that the discrepancy of the biological and psychological characteristics of participants at baseline might have influenced the accuracy of effect estimation of interventions in the trials. This idea was supported by a previous systemic review on a similar topic, suggesting that the prognostic heterogeneity among participants in LBP RCTs might dilute the positive treatment effect of the intervention (Fersum et al., 2010). Research into CNSLBP without subclassification was therefore once considered not likely to offer useful insight (Leboeuf-Yde & Manniche, 2001).

Future research should consider using existing classification systems or clinical prediction rules to identify homogeneous subgroups of patients for clinical trials. Future research should also consider psychosocial factors when classifying patients into subgroups to reflect the biopsychosocial nature of CNSLBP. This may increase the value of future research for clinical practice and provide clinicians with evidence regarding the selection of exercise interventions for subgroups of CNSLBP. However, it should be acknowledged no single set of classification systems or clinical prediction rules was considered the gold standard and each of them had its own methodological limitations (Fersum et al., 2010).

Another challenge involved in the investigation of the comparative effectiveness of exercise interventions for CNSLBP could be the fidelity of implementation of exercise interventions. The complexity of principles underpinning the exercise interventions, such as the application of the neutral spine principle and the adherence to specific breathing patterns and trunk muscle activation technique, and the adherence to these principles in the trials may be important to capture any difference more accurately in treatment effect among various forms of exercises. More RCTs comparing different exercise interventions with higher methodological quality and larger sample size are warranted.

Strength and Limitations

This review updated current evidence base of the comparative effectiveness between exercises interventions foe CNSLBP by offering direct comparison using pair-wise meta-analysis, supplementing the recently published reviews with similar research questions. This review critically highlighted some methodological limitations from the trials investigating the effectiveness of PE and explored the potential insufficiency of trial implementations. The synthesis and discussion of findings by drawing in current evidence offered implications for clinical practice and future research.

There were several limitations. Publication bias might arise since only trials published in English were included. The findings were limited by the low to very low certainty of evidence. The data extraction and data analysis were done by single author. Since subgroup analysis was used to pool the results to eliminate heterogeneity, the number of studies included in each subgroup was small. This might have limited the power of the results obtained. This review was completed as part of a master's dissertation, thus the review was not prospectively registered and protocol was unpublished. However, unpublished protocol can be found in supplementary files.

Conclusion

This review found no strong evidence for using one type of exercise intervention over another when managing patients with CNSLBP. Existing evidence does not allow this review to draw definitive recommendations. In the absence of a superior exercise form clinicians should work collaboratively with the patient, using the individual's goals and preferences to guide exercise selection. Further appropriately designed research is warranted to explore this topic further.

Reference

Akodu, A., Akinbo, S., Okonkwo, C. (2016). Comparative efficacy of core stabilisation exercise and Pilates exercise on patients with non-specific low back pain. *Romanian journal of physical therapy*, 22, 14-22.

Anand, U. A., Caroline, P. M., Arun, B., & Gomathi, G. L. (2014). A study to analyse the efficacy of modified pilates based exercises and therapeutic exercises in individuals with chronic non specific low back pain: a randomized controlled trial. *International journal of physiotherapy and research*, *2*(3), 525-29.

Barbosa, A. W. C., Guedes, C. A., Bonifácio, D. N., de Fátima Silva, A., Martins, F. L. M., & Barbosa, M. C. S. A. (2015). The Pilates breathing technique increases the electromyographic amplitude level of the deep abdominal muscles in untrained people. *Journal of bodywork and movement therapies*, *19*(1), 57-61.

Barbosa, A. W. C., Martins, F. L. M., de Melo Vitorino, D. F., & Barbosa, M. C. S. A. (2013). Immediate electromyographic changes of the biceps brachii and upper rectus abdominis muscles due to the Pilates centring technique. *Journal of bodywork and movement therapies*, *17*(3), 385-390.

Bhadauria, E. A., & Gurudut, P. (2017). Comparative effectiveness of lumbar stabilization, dynamic strengthening, and Pilates on chronic low back pain: randomized clinical trial. *Journal of exercise rehabilitation*, *13*(4), 477.

Brooks, C., Kennedy, S., & Marshall, P. W. (2012). Specific trunk and general exercise elicit similar changes in anticipatory postural adjustments in patients with chronic low back pain: a randomized controlled trial. *Spine*, *37*(25), E1543-E1550.

Brumagne, S., Diers, M., Danneels, L., Moseley, G. L., & Hodges, P. W. (2019). Neuroplasticity of sensorimotor control in low back pain. *Journal of orthopaedic & sports physical therapy*, *49*(6), 402-414.

Dsa, C. F., Rengaramanujam, K., & Kudchadkar, M. S. (2014). To assess the effect of modified pilates compared to conventional core stabilization exercises on pain and disability in chronic non-specific low back pain-randomized controlled trial. *Indian Journal of physiotherapy and occupational therapy*, *8*(3), 202.

Ehsani, F., Arab, A. M., Jaberzadeh, S., & Salavati, M. (2016). Ultrasound measurement of deep and superficial abdominal muscles thickness during standing postural tasks in participants with and without chronic low back pain. *Manual therapy*, *23*, 98-105.

Eliks, M., Zgorzalewicz-Stachowiak, M., & Zeńczak-Praga, K. (2019). Application of Pilates-based exercises in the treatment of chronic non-specific low back pain: state of the art. *Postgraduate medical journal*, *95*(1119), 41-45.

Fairbank, J., Gwilym, S. E., France, J. C., Daffner, S. D., Dettori, J., Hermsmeyer, J., & Andersson, G. (2011). The role of classification of chronic low back pain. *Spine*, *36*, S19-S42.

Fersum, K. V., Dankaerts, W., O'sullivan, P. B., Maes, J., Skouen, J. S., Bjordal, J. M., & Kvåle, A. (2010). Integration of subclassification strategies in randomised controlled clinical trials evaluating manual therapy treatment and exercise therapy for non-specific chronic low back pain: a systematic review. *British journal of sports medicine*, *44*(14), 1054-1062.

Fujitani, R., Jiromaru, T., Kida, N., & Nomura, T. (2017). Effect of standing postural deviations on trunk and hip muscle activity. *Journal of physical therapy science*, *29*(7), 1212-1215.

Gordon, R., & Bloxham, S. (2016). A systematic review of the effects of exercise and physical activity on non-specific chronic low back pain. *Healthcare 4(*2), 22.

Hasanpour-Dehkordi, A., Dehghani, A., & Solati, K. (2017). A comparison of the effects of Pilates and McKenzie training on pain and general health in men with chronic low back pain: a randomized trial. *Indian journal of palliative care*, *23*(1), 36.

Hayden, J. A., Ellis, J., Ogilvie, R., Stewart, S. A., Bagg, M. K., Stanojevic, S., ... & Saragiotto, B. T. (2021). Some types of exercise are more effective than others in people with chronic low back pain: A network meta-analysis. *Journal of physiotherapy*, *67*(4), 252-262.

Hemming, R., Sheeran, L., van Deursen, R., & Sparkes, V. (2019). Investigating differences in trunk muscle activity in non-specific chronic low back pain subgroups and no-low back pain controls during functional tasks: a case-control study. *BMC musculoskeletal disorders*, *20*(1), 1-10.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hoffman, J., & Gabel, C. P. (2015). The origins of Western mind–body exercise methods. *Physical therapy reviews*, *20*(5-6), 315-324.

Kim, S. T., & Lee, J. H. (2017). The effects of Pilates breathing trainings on trunk muscle activation in healthy female subjects: a prospective study. *Journal of physical therapy science*, *29*(2), 194-197.

Kinney, A. R., Eakman, A. M., & Graham, J. E. (2020). Novel effect size interpretation guidelines and an evaluation of statistical power in rehabilitation research. *Archives of physical medicine and rehabilitation*, *101*(12), 2219-2226.

Kofotolis, N., Kellis, E., Vlachopoulos, S. P., Gouitas, I., & Theodorakis, Y. (2016). Effects of Pilates and trunk strengthening exercises on health-related quality of life in women with chronic low back pain. *Journal of back and musculoskeletal rehabilitation*, *29*(4), 649-659.

Koh, H. W., Cho, S. H., & Kim, C. Y. (2014). Comparison of the effects of hollowing and bracing exercises on cross-sectional areas of abdominal muscles in middle-aged women. *Journal of physical therapy science*, *26*(2), 295-299.

Leboeuf-Yde, C., & Manniche, C. (2001). Low back pain: time to get off the treadmill. *Journal of manipulative & physiological therapeutics*, *24*(1), 63-66.

Lim, E. C. W., Poh, R. L. C., Low, A. Y., & Wong, W. P. (2011). Effects of Pilates-based exercises on pain and disability in individuals with persistent nonspecific low back pain: a systematic review with meta-analysis. *Journal of orthopaedic & sports physical therapy, 41*(2), 70-80.

Lin, H. T., Hung, W. C., Hung, J. L., Wu, P. S., Liaw, L. J., & Chang, J. H. (2016). Effects of Pilates on patients with chronic non-specific low back pain: a systematic review. *Journal of physical therapy science*, *28*(10), 2961-2969.

Marshall, P. W., Kennedy, S., Brooks, C., & Lonsdale, C. (2013). Pilates exercise or stationary cycling for chronic nonspecific low back pain: does it matter? a randomized controlled trial with 6-month follow-up. *Spine*, *38*(15), E952-E959.

Massé-Alarie, H., Beaulieu, L. D., Preuss, R., & Schneider, C. (2016). Influence of paravertebral muscles training on brain plasticity and postural control in chronic low back pain. *Scandinavian journal of pain*, *12*(1), 74-83.

Mazloum, V., Sahebozamani, M., Barati, A., Nakhaee, N., & Rabiei, P. (2018). The effects of selective Pilates versus extension-based exercises on rehabilitation of low back pain. *Journal of bodywork and movement therapies*, *22*(4), 999-1003.

Miyamoto, G. C., Costa, L. O., & Cabral, C. (2013). Efficacy of the Pilates method for pain and disability in patients with chronic nonspecific low back pain: a systematic review with meta-analysis. *Brazilian journal of physical therapy*, *17*, 517-532.

Mostagi, F. Q. R. C., Dias, J. M., Pereira, L. M., Obara, K., Mazuquin, B. F., Silva, M. F., ... & Cardoso, J. R. (2015). Pilates versus general exercise effectiveness on pain and functionality in non-specific chronic low back pain subjects. *Journal of bodywork and movement therapies*, *19*(4), 636-645.

O'Connell, N. E., Cook, C. E., Wand, B. M., & Ward, S. P. (2016). Clinical guidelines for low back pain: a critical review of consensus and inconsistencies across three major guidelines. *Best practice & research Clinical rheumatology*, *30*(6), 968-980.

Owen, P. J., Miller, C. T., Mundell, N. L., Verswijveren, S. J., Tagliaferri, S. D., Brisby, H., ... & Belavy, D. L. (2020). Which specific modes of exercise training are most effective for treating low back pain? Network meta-analysis. *British journal of sports medicine*, *54*(21), 1279-1287.

Patti, A., Bianco, A., Paoli, A., Messina, G., Montalto, M. A., Bellafiore, M., ... & Palma, A. (2015). Effects of Pilates exercise programs in people with chronic low back pain: a systematic review. *Medicine*, *94*(4).

Rackwitz, B., de Bie, R., Limm, H., von Garnier, K., Ewert, T., & Stucki, G. (2006). Segmental stabilizing exercises and low back pain. What is the evidence? A systematic review of randomized controlled trials. *Clinical rehabilitation*, *20*(7), 553-567.

Rubinstein, S. M., Terwee, C. B., Assendelft, W. J., de Boer, M. R., & van Tulder, M. W. (2012). Spinal manipulative therapy for acute low-back pain. *Cochrane database of systematic reviews*, (9).

Searle, A., Spink, M., Ho, A., & Chuter, V. (2015). Exercise interventions for the treatment of chronic low back pain: a systematic review and meta-analysis of randomised controlled trials. *Clinical rehabilitation*, *29*(12), 1155-1167.

Sheeran, L., Sparkes, V., Caterson, B., Busse-Morris, M., & van Deursen, R. (2012). Spinal position sense and trunk muscle activity during sitting and standing in nonspecific chronic low back pain: classification analysis. *Spine*, *37*(8), E486-E495.

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*, *366*.

Stolze, L. R., Allison, S. C., & Childs, J. D. (2012). Derivation of a preliminary clinical prediction rule for identifying a subgroup of patients with low back pain likely to benefit from Pilates-based exercise. *Journal of orthopaedic & sports physical therapy*, *42*(5), 425-436.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., ... & Bhutta, Z. A. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet, 396*(10258), 1204-1222.

Wajswelner, H., Metcalf, B., & Bennell, K. (2012). Clinical Pilates versus general exercise for chronic low back pain: randomized trial. *Medicine & science in sports & exercise*, *44*(7), 1197-205.

Wells, C., Kolt, G. S., & Bialocerkowski, A. (2012). Defining Pilates exercise: a systematic review. *Complementary therapies in medicine*, *20*(4), 253-262.

Wells, C., Kolt, G. S., Marshall, P., Hill, B., & Bialocerkowski, A. (2013). Effectiveness of Pilates exercise in treating people with chronic low back pain: a systematic review of systematic reviews. *BMC medical research methodology*, *13*(1), 1-12.

Wells, C., Kolt, G. S., Marshall, P., Hill, B., & Bialocerkowski, A. (2014). The effectiveness of Pilates exercise in people with chronic low back pain: a systematic review. *Plos one*, *9*(7), e100402.

Wong, A. Y., Chan, T. P., Chau, A. W., Cheung, H. T., Kwan, K. C., Lam, A. K., ... & De Carvalho, D. (2019). Do different sitting postures affect spinal biomechanics of asymptomatic individuals?. *Gait & posture*, *67*, 230-235.

Yamato, T. P., Maher, C. G., Saragiotto, B. T., Hancock, M. J., Ostelo, R. W., Cabral, C. M., ... & Costa, L. O. (2015). Pilates for low back pain. *Cochrane database of systematic reviews*, (7).

Titles of figure

FIGURE 1. Flowchart of study selection process

FIGURE 2. Distribution of the risk of bias assessment by domains

FIGURE 3. Forest plot showing subgroup standard mean differences in pain between Pilates exercise (experimental) and other forms of exercise (control)

FIGURE 4. Forest plot showing subgroup standard mean differences in pain between Pilates exercise (experimental) and spinal stabilisation exercise (control) with the exclusion of data from Dsa (2014)

FIGURE 5. Forest plot showing subgroup standard mean differences in disability between Pilates exercise (experimental) and other forms of exercise (control)

FIGURE 6. Funnel plot for the outcome of pain.

FIGURE 6. Funnel plot for the outcome of disability.

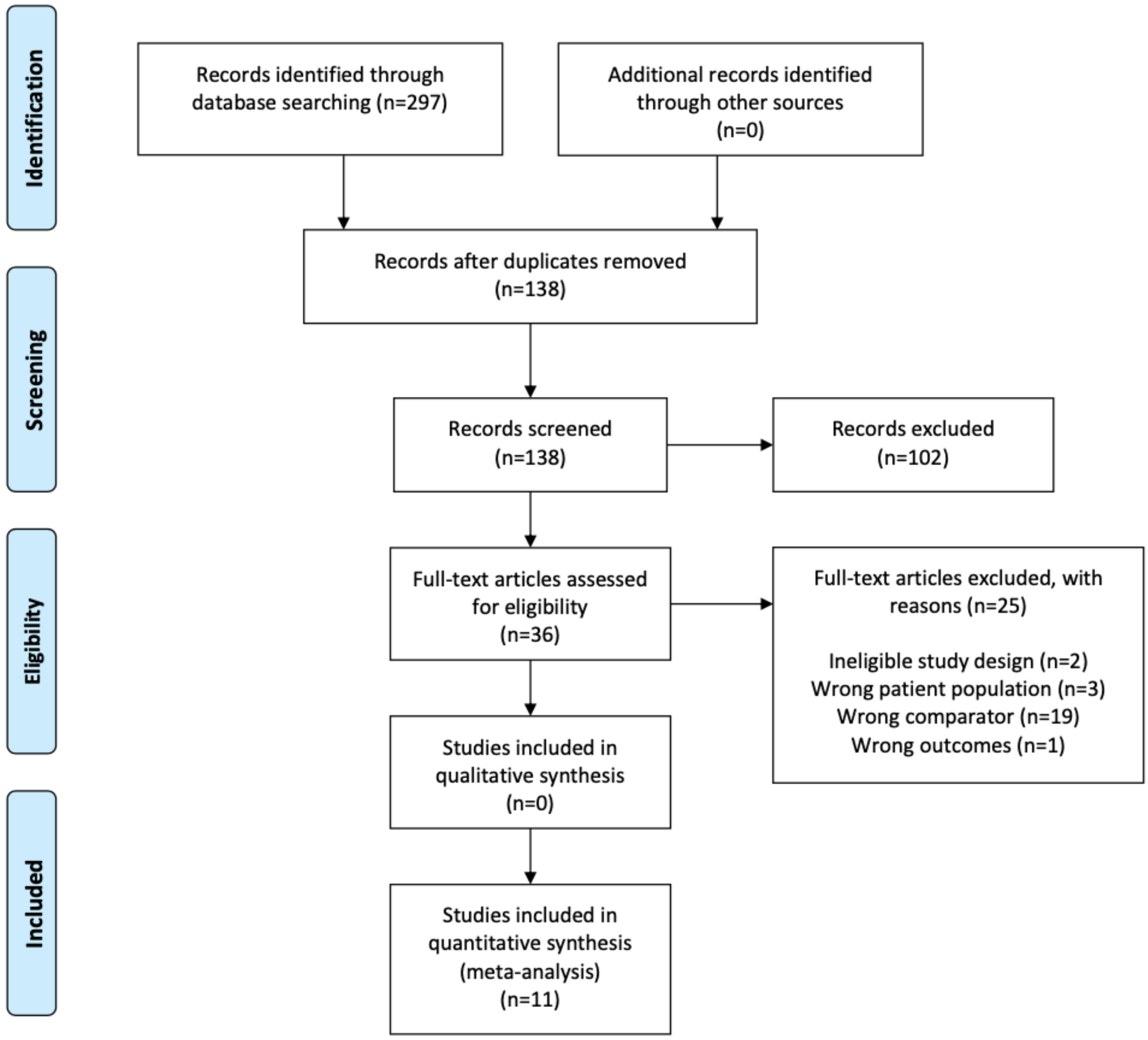

FIGURE 1. Flowchart of study selection process

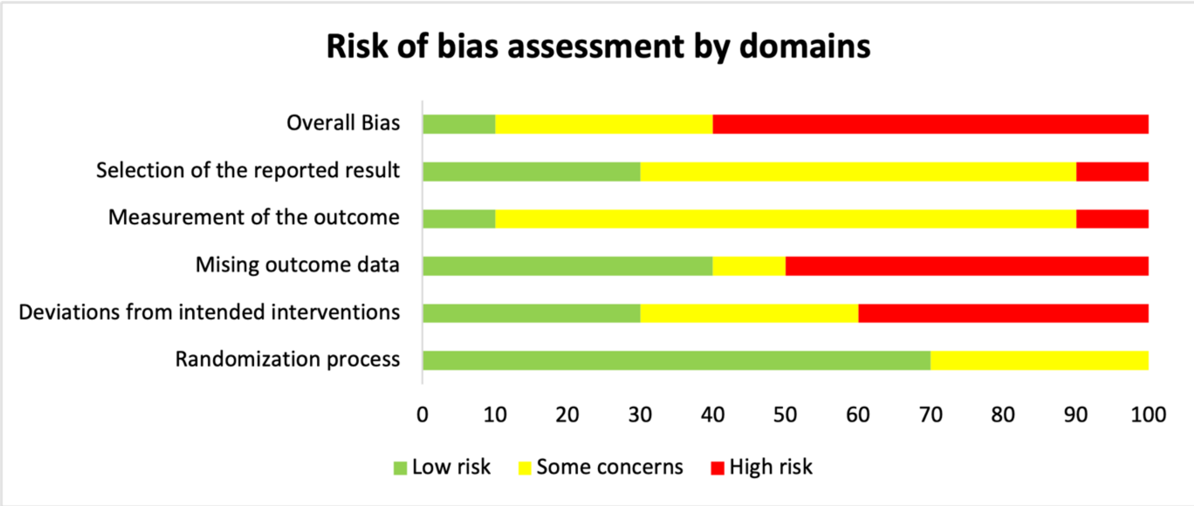

FIGURE 2. Distribution of the risk of bias assessment by domains

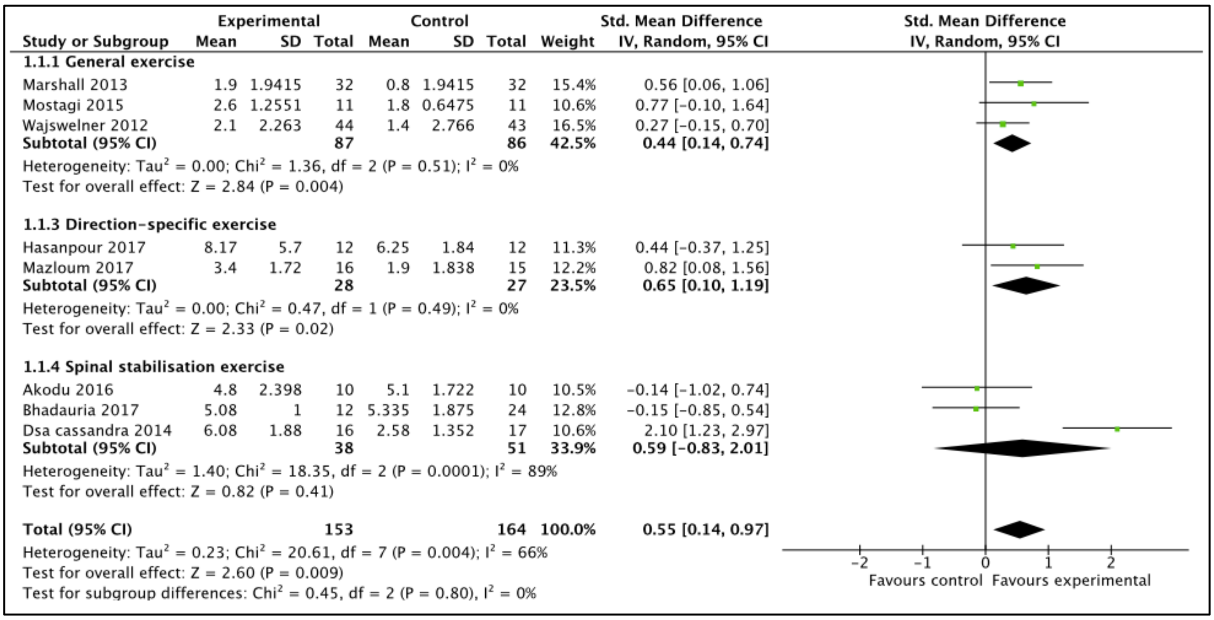

FIGURE 3. Forest plot showing subgroup standard mean differences in pain between Pilates exercise (experimental) and other forms of exercise (control).

Abbreviations: SD: standard deviation; CI: confidence intervals; I^2: inconsistency test.

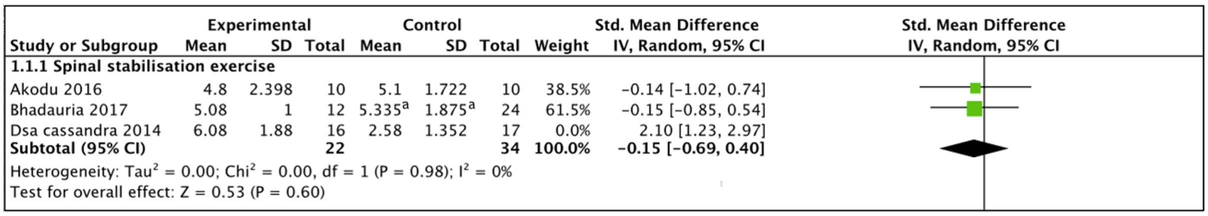

FIGURE 4. Forest plot showing subgroup standard mean differences in pain between Pilates exercise (experimental) and spinal stabilisation exercise (control) with the exclusion of data from Dsa (2014).

[a] Data of Bhadauria (2017) presented here were combined data from the trial due to the high similarity of the lumbar stabilisation and dynamic strengthening groups in the trial. The combination of the data did not result in any major change in the pooled result in the subgroup analysis

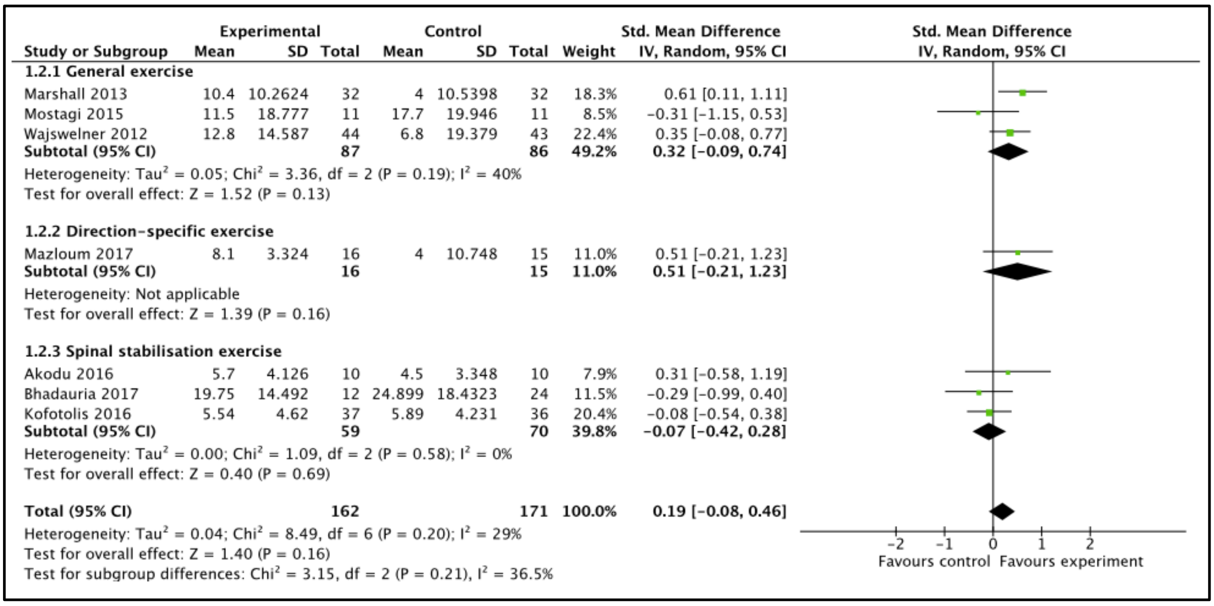

FIGURE 5. Forest plot showing subgroup standard mean differences in disability between Pilates exercise (experimental) and other forms of exercise (control).

Abbreviations: SD: standard deviation; CI: confidence intervals; I^2: inconsistency test.

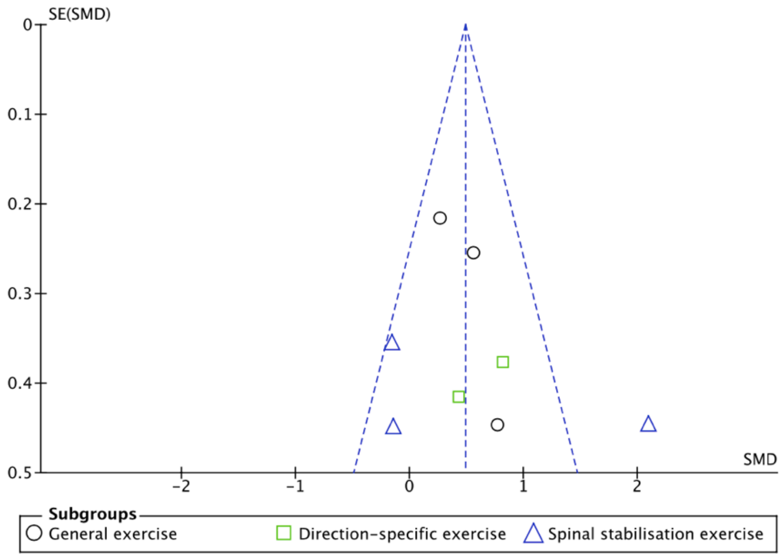

FIGURE 6. Funnel plot for the outcome of pain.
Abbreviations: SE: standard error; SMD: standard mean difference.

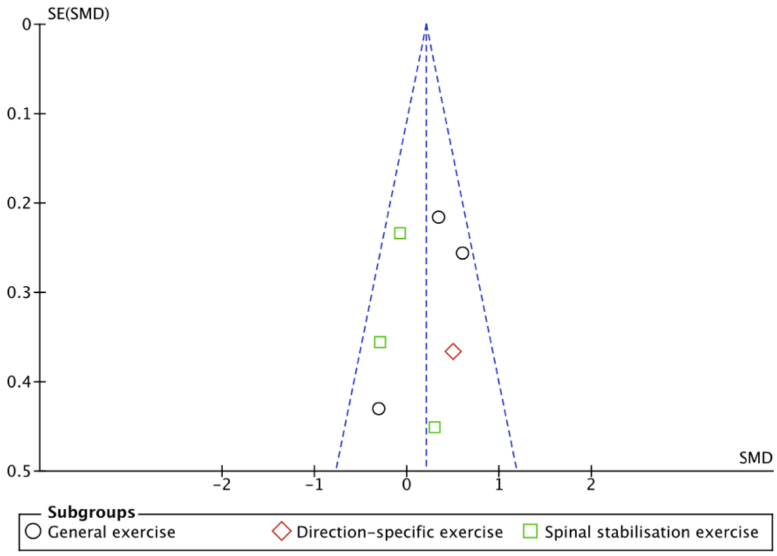

FIGURE 7. Funnel plot for the outcome of disability
Abbreviations: SE: standard error; SMD: standard mean difference.

Titles of tables

TABLE 1. Methodological characteristics of the included studies

First author (Year)/ Country	Participant	Inclusion criteria	Exclusion criteria	Description of interventions	Outcome measurements	Results	Included in previous reviews
Subgroup 1: Pilates versus General exercises							
Anand (2014) India	IG: *n* = 15 Loss to follow-up = NR Mean age (SD): NR Sex: NR Duration of symptoms: NR	Subjects with LBP not more than 5 in Visual analog scale, aged 18-60 yrs, pain >3 months, doing normal ADL activity, working population, normal BMI, no previous research physiotherapy involvement for the past 2 months, no psychological or yellow flag	Subjects with Intervertebral disc prolapsed, radiating pain, stenosis, severe spondylitis, and spondylolisthesis, cardiovascular problems, tumours, Infection or fracture, osteoporosis, radicular syndrome, inflammatory disorder, structural deformity not optimal for exercises and psychologically unstable patients	Experimental intervention: Standard back care + General flexibility exercises (15 mins) + Modified Pilates-based exercises on mat (45mins) for 8/52 Individualised: Yes	Pain (VAS) Disability (ODI) At 8/52 from baseline	Pain and disability were both reported to have a statistically significant improvement in Modified Pilates-based exercises group than Therapeutic exercises group. However, results were reported using plain text and appeared to be unclear. No table was presented	Yes - Patti et al. (2015)
	CG: *n* = 15 Loss to follow-up = NR Mean age (SD): NR Sex: NR Duration of symptoms: NR			Comparator intervention: Standard back care + General flexibility exercises (15 mins) + Therapeutic exercises (45 mins) including floor exercises, stationary bicycle and swiss ball coordination exercises for 8/52 Individualised: Yes			
Brooks (2012) † Australia	IG: n = 32 Loss to follow-up = 3 Mean age (SD): 36.2 (8.2) yrs Male:12 (37.5%) Duration of	Men and women aged 18-50 yrs, ongoing recurrent LBP (>12/52) located between the costal margins and inferior gluteal folds	Severe postural abnormality, pain radiating below the knee, known lumbar disc hernia or fracture, history of back surgery, diagnosed inflammatory joint	Experimental intervention: Specific Pilates exercises in different body positions with 8 components on mat and apparatus (50-60 mins), x3/week for 8/52 Individualised: NR	Pain (VAS) Disability (ODI) At 8/52 from baseline	There was a statistically significant reductions on current pain (P < 0.05) and disability (P = 0.018) in the Specific Pilates exercise group when compared to the Indoor stationary cycle training program group at 8/52	No

	symptoms: 9.5 (8.0) yrs		disease, known severe osteoporosis, known metabolic or neuromuscular disease, or recent (<3 mo) participation in an exercise program or any form of therapeutic treatment (i.e., manipulation, mobilization, massage)				
	CG: n = 32 Loss to follow-up = 9 Mean age (SD): 36.2 (6.2) yrs Male: 12 (37.5%) Duration of symptoms: 11.1 (7.9) yrs			Comparator intervention: Indoor stationary cycle training program with 8 components (50-60 mins), x3/week for 8/52 Individualised: No, group class with 10:1 ratio			
Marshall (2013) † Australia	Same as in Brooks et al. (2012)	Same as in Brooks et al. (2012)	Same as in Brooks et al. (2012)	Same as in Brooks et al. (2012)	Pain (VAS) Disability (ODI) At 8/52 from baseline, 6-month follow-up	No long-term difference was observed at 6-moth follow-up for both pain and disability	Yes - Wells et al., (2014), Patti et al. (2015), Lin et al. (2016)
Mostagi (2015) Brazil	IG: n = 11 Loss to follow-up = 1 Mean age (SD): 36.1 (9) yrs Male: 2 (18.2%) Duration of symptoms: NR	Sedentary and had not undergone physical therapy for at least 6 months, presented an exclusive medical diagnosis of non-specific chronic low back pain over a period >12/52 and aged 18-55 yrs	Diagnosis of protrusion of the intervertebral disc, scoliosis, spondylolisthesis, previous spine surgery, radicular symptoms, inflammatory disease, rheumatic	Experimental intervention: Direction specific Pilates method exercises on apparatus (60 mins), x2/week for 8/52 Individualised: Yes	Pain (VAS) Disability (QBPQ) At 8/52 from baseline, 3-month follow-up	No statistically significant difference between groups on both pain and disability at 8/52 and 3-month follow-up	No

	CG: n = 11 Loss to follow-up = 4 Mean age (SD): 34.7 (8.1) yrs Male: 2 (18.2%) Duration of symptoms: NR		disease, cancer or pregnancy	Comparator intervention: Standardised generic exercises including stationary bicycling, trunk and lower limb stretching, spine mobilisation and trunk muscle strengthening (60 mins), x2/week for 8/52 Individualised: Yes			
Wajswelner (2012) Australia	IG: n = 44 Loss to follow-up = 3 Mean age (SD): 49.3 (14.1) yrs Female: 25 (57%) Duration of symptoms: 13.6 (14.2) yrs	Age 18-70 yrs, symptoms of pain or stuffiness in the lower back with or without lower limb symptoms on most days of the week for more than 3 months, average pain score in the past week at telephone screening >4 on NRS and good understanding of written and spoken English	Spinal surgery; fever, infection, night sweats or rigors; unexplained weight loss or loss of appetite; history of cancer or malignancy; cauda equina lesion, loss of bladder or bowel control, or saddle paraesthesia; pregnancy or the possibility of pregnancy in the next 6 months; spinal fractures or diagnosed osteoporosis; spinal inflammatory disease such as ankylosing spondylitis, rheumatoid arthritis; comorbidities that	Experimental intervention: 6-12 direction specific Pilates exercises on apparatus (60 mins) with 1-4 home exercises (floor/chair/wall exercises), x2/week for 6/52 Individualised: Yes	Pain (NRS) Disability (QBPQ) At 6/52 from baseline, 3-month and 6-month follow-up	No statistically significant difference between groups on both pain and disability at 6/52, 3-month and 6-month follow-up	Yes - Miyamoto et al. (2013), Wells et al., (2014), Patti et al. (2015), Lin et al. (2016)
	CG: n = 43 Loss to follow-up = 1 Mean age (SD): 48.9 (16.4) yrs Female: 23 (53%) Duration of symptoms: 14.2 (12.7) yrs			Comparator intervention: Standardised generic exercises including stationary bike, lower limb stretching, upper bodyweight, TheraBand, swiss ball and floor exercises (60 mins), x2/week for 8/52 Individualised: NR			

			would prevent exercise; previous participation in a clinical Pilates program or other regular therapeutic back exercise program in the last 6 months; inability to comply with trial requirements; or compensable back pain				
Subgroup 2: Pilates versus Direction-specific exercises							
Hasanpour (2017) Iran	IG: n = 12 Loss to follow-up = NR Mean age (SD): NR Male: all Duration of symptoms: NR	Aged 40–55 years in with chronic back pain (history of more than 3 months of low back pain) and no specific disease or other surgery	Low back arch or so-called army back, serious spinal pathology such as tumours, fractures, inflammatory diseases, previous spinal surgery, nerve root compromise in the lumbar region, spondylolysis or spondylolisthesis, spinal stenosis, neurological disorders, systemic diseases, cardiovascular diseases, and receiving other therapies simultaneously	Experimental intervention: Pilates training program (no details provided) (60 mins), x3/week for 6/52 Individualised: Yes	Pain (MPQ) At 6/52 from baseline	No statistically significant difference between groups on pain at 6/52	No
	CG: n = 12 Loss to follow-up = NR Mean age (SD): NR Men: all Duration of symptoms: NR			Comparator intervention: McKenzie exercises (4 extension-type exercises + 2 flexion-type exercises) (60mins), daily for 20 days Individualised: Yes			

Mazloum (2017) Iran	IG: n = 20 Loss to follow-up = 4 Mean age (SD): 37.1 (9.5) yrs Sex: NR Duration of symptoms: 32.3 (18.2) months	Adults aged of 18-55 yrs, diagnosis of non-specific LBP, lasting signs and symptoms >3 months, exercise indication for the subject based on clinical evaluation, the satisfaction of the person to participate in the study	History of trauma to the spinal column, any misalignment or specific condition in the lumbar spine, spondylosis or spondylolisthesis, history of spinal surgery, neurological or psychological conditions, receiving physical therapy or other treatment interventions in the past six months	Experimental intervention: Selective Pilates exercises on mat, x3/week for 6/52 Individualised: NR	Pain (VAS) Disability (ODI) At 4/52 from baseline, 6/52 follow-up	There was a statistically significant difference on pain (P < 0.01) but not disability (P = 0.851) in the Specific Pilates exercise group when compared to the Extension-based exercises group at 4/52	No
	CG: n = 20 Loss to follow-up = 5 Mean age (SD): 42.7 (8.1) yrs Sex: NR Duration of symptoms: 30.8 (15.3) months			Comparator intervention: Extension-based exercises (first four weeks: extension-type exercise, last two weeks: flexion-type exercise), x3/week for 6/52 Individualised: NR			
Subgroup 3: Pilates versus Spinal stabilisation exercises							
Akodu (2016) Nigeria	IG: n = 14 Loss to follow-up = 13 in the whole sample Mean age (SD): 45.3 (11.31) yrs Sex: NR Duration of symptoms: NR	Subjects with a history of non-specific chronic low back pain with or without pain radiating to one or both lower limbs, and Patients with recurrent history of	Subjects confirmed to be pregnant, subjects with specific LBP, subjects with medical or surgical conditions that might hinder exercise performance	Experimental intervention: Infra-red radiation + Pilates exercise protocol with 8 exercises on mat, x2/week for 4/52 Individualised: NR	Pain (NRS) Disability (RMDQ) At 2/52, 4/52 from baseline	No statistically significant difference between Pilates exercises group and Core stabilisation exercises group on both pain and disability at 2/52 and 4/52	No

	CG: n = 14 Loss to follow-up = 13 in the whole sample Mean age (SD): 49.1 (11.85) yrs Sex: NR Duration of symptoms: NR	LBP of not less than 3 months		Comparator intervention: Infra-red radiation + Core stabilisation exercise protocol with 9 exercises (bracing throughout all exercises), x2/week for 4/52 Individualised: NR			
Bhadauria (2017) India	IG: n = 12 Loss to follow-up = 3 Mean age (SD): 35.33 (12.88) yrs Male: 91.6% Duration of symptoms: 1.53 (1.64) yrs	All male and female adults aged 20-60 yrs, subjects with nonspecific back pain >3 months, and subjects willing to participate in the study	Subjects with specific back pain (fracture, osteoporosis or degenerative changes, prolapse intervertebral disc, bone disorders, arthritis, tumour), subjects with neurological involvement (radiculopathy, myelopathy), subjects with previous spinal surgery, subjects with spinal infections, and subjects with severe psychiatric disorder	Experimental intervention: Hot moist pad + Interferential current therapy + 10 Pilates exercises on mat, (whole session 60 mins), x10 sessions in 3/52 Individualised: NR	Pain (VAS) Disability (Modified ODI) At 3/52 from baseline	Pain (P = 0.0068) and disability (P = 0.0001) were both reported to have a statistically significant difference among 3 groups as measured in mean difference at 3/52. Also concluded that lumbar stabilisation exercises was more superior to Pilates exercises and Dynamic strengthening exercises. However, no further details regarding comparison between groups was reported clearly with data	No
	CG1: n = 12 Loss to follow-up = 3 Mean age (SD): 35.33 (12.88) yrs Male: 50% Duration of symptoms: 0.58 (0.54) yrs			Comparator 1 intervention: Hot moist pad + Interferential current therapy + 16 Lumbar stabilisation exercises (whole session 60 mins), x10 sessions in 3/52 Individualised: NR			
	CG2: n = 12 Loss to follow-up = 2 Mean age (SD): 36.67			Comparator 2 intervention: Hot moist pad + Interferential current therapy + 14 Lumbar dynamic strengthening			

	(10.74) yrs Male: 58.33% Duration of symptoms: 0.31 (0.42) yrs			exercises (whole session 60 mins), x10 sessions in 3/52 Individualised: Yes			
Dsa (2014) India	IG: *n* = 17 Loss to follow-up = 1 Mean age (SD): NR Male: 6 (37.5%) Duration of symptoms: NR	Chronic non-specific low back pain for at least 12 weeks, aged 18-45 yrs, patient is otherwise medically fit to perform exercises (subjects with no systemic disease)	Back pain attributed to any other pathology, malignancies, major surgery within the past years (back surgery), radiating pain in the lower limbs (neural involvement)	Experimental intervention: Moist heat (10 mins) + 6 Pilates exercises on mat with instruction 'tucking in the stomach' for all exercises, for 2/52 Individualised: NR	Pain (VAS) Disability (RMDQ) At 2/52 from baseline	There was a more statistical significant improvement on both pain ($P < 0.01$) and disability ($P < 0.01$) in the Pilates exercises group than the Core stabilisation exercises group at 2/52. However, the data was presented with confusion	No
	CG: *n* = 21 Loss to follow-up = 4 Mean age (SD): NR Male: 6 (35.2%) Duration of symptoms: NR			Comparator intervention: Moist heat (10 mins) + 6 Core stabilisation exercises on mat with instruction 'tucking in the stomach' for 5 seconds in each exercise, for 2/52 Individualised: NR			
Kofotolis (2016) Greece	IG: n = 40 Loss to follow-up = 3 Mean age (SD): 41.22 (8.49) yrs Female: all Duration of symptoms: NR	Female, aged 25–65 yrs, a new episode of non-specific LBP lasting more than 12 weeks, and an inability to resume daily activities in the last three weeks	Acute LBP, spinal stenosis or surgery, inflammatory disease affecting the spine, fracture, spondylolysis or spondylolisthesis, genetic spinal structure abnormality, acute LBP, pregnancy, use	Experimental intervention: 16 Pilates exercises on mat (60 mins), x3/week for 8/52 Individualised: NR	Disability (ODI) At 4/52, 8/52 from baseline, 3-month a follow-up	The Pilates exercises group was reported to have a statistically significant greater improvement on disability ($P < 0.05$) when compared to Trunk strengthening exercises at 8/52 and 3-month follow-up.	No
	CG: n = 40 Loss to follow-			Comparator intervention: Trunk strengthening			

	up = 4 Mean age (SD): 39.11 (8.68) yrs Female: all Duration of symptoms: NR		of medication that affects heart rate and/or blood pressure and pelvic girdle pain	exercises for abdominals, back extensors and the whole-body (60 mins), x3/week for 8/52 Individualised: NR		However, no clear data was presented	

† Brooks (2012) and Marshall (2013) are two separate reports but shared the same set of samples.
Abbreviations: IG: intervertion group; CG: control group; NR: not reported; yrs: years; /52: weeks; x: times, SD: standard deviation; mins: minutes; mo: month; RMDQ: Roland Morris Disability Questionnaire; ODI: Oswestry Disability Index; QBPQ: Quebec Back Pain Questionnaire; NRS: Numeric Rating Scale; VAS: Visual Analogue Scale; MPQ: McGill Pain Questionnaire.

TABLE 2. Risk of bias assessment

Study (Year)	**Subgroup**	Randomisation process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	**Overall**
Anaad et al., (2014)	1	+	-	-	?	?	-
Brooks et al., (2013) or Marshall et al., (2013)	1	+	+	+	+	+	+
Mostagi et al., (2015)	1	+	+	+	?	+	?
Wajswelner et al., (2012)	1	+	+	+	?	+	?
Hasanpour et al., (2017)	2	?	-	-	-	?	-
Mazloum et al., (2017)	2	?	-	?	?	?	-
Akodu et al., (2016)	3	+	-	-	?	?	-
Bhadauria et al., (2017)	3	?	?	+	?	?	?
Dsa et al., (2014)	3	+	?	-	?	-	-
Kofotolis et al., (2016)	3	+	?	-	?	?	-

+	Low risk
?	Some concerns
-	High risk

TABLE 3. Reported outcomes collected by timepoints (pain and disability)

	Subgroup	General exercise				Direction-specific exercise		Spinal stabilisation exercise			
Outcome	Intervention / Study	Anand (2014)	Brooks (2012)	Mostagi (2015) †	Wajswelner (2012) †	Hasanpour (2017)	Mazloum (2017) †	Akodu (2016) †	Bhadauria (2017)	Dsa (2014)	Kofotolis (2016) †
Pain	Pilates exercise	-	8 wks change: 1.9 (1.9415)	0 wk: 3 (3.259)	0 wk: 4.9 (1.6)	6 wks change: 8.17 (5.70)	0 wk: 6.8 (1.4)	0 wk: 6.9 (2.02)	3 wks change: 5.08 (1.0)	2 wks change: 6.08 (1.88)	-
		-		8 wks: 0.4 (2.59)	6 wks: 2.8 (1.6)		4 wks: 3 (0.9)	4 wks: 2.1 (1.91)			-
	Control (other exercises)	-	8 wks change: 0.8 (1.9415)	0 wk: 2.3 (1.481)	0 wk: 4.6 (1.8)	6 wks change: 6.25 (1.84)	0 wk: 7.2 (1.3)	0 wk: 6.2 (1.14)	3 wks change: 6.00 (0.85) ‡	2 wks change: 2.58 (1.352)	-
		-		8 wks: 0.5 (1.55)	6 wks: 3.2 (2.1)		4 wks: 4.8 (1.1)	4 wks: 1.1 (1.29)			-
									3 wks change: 4.67 (1.07) ‡		
Disability	Pilates exercise	-	8 wks change: 10.4 (10.262)	0 wk: 27 (15.7)	0 wk: 28.1 (11.4)	-	0 wk: 30.8 (1.2)	0 wk: 11.1 (2.8)	3 wks change: 19.75 (9.23)	2 wks change: 39.32 (14.67)	0 wk: 11.32 (4.11)
		-		8 wks: 15.5 (10.3)	6 wks: 15.3 (9.1)	-	4 wks: 22.9 (3.6)	4 wks: 5.4 (3.03)			4 wks: 5.78 (2.11)
	Control (other exercises)	-	8 wks change: 4 (10.539)	0 wk: 29.4 (17.8)	0 wk: 23.9 (14)	-	0 wk: 27.2 (7.6)	0 wk: 11.4 (2.67)	3 wks change: 14.33 (7.01) §	2 wks change: 9.32 (8.781)	0 wk: 12.41 (3.69)
		-		8 wks: 11.7 (9)	6 wks: 17.1 (13.4)	-	4 wks: 23.1 (7.5)	4 wks: 3.6 (2.54)			4 wks: 6.52 (2.07)
						-			3 wks change: 32.83 (10.68) §		

† Mean difference was calculated by subtracting the baseline values from the post-intervention values whereas the standard deviations were estimated based on the standard error of the mean change.
‡ Data were combined for meta-analysis. Combined value: 5.335 (1.875).
§ Data were combined for meta-analysis. Combined value: 24.899 (18.4323).
Data was shown in mean (standard deviation).
Data of Anand (2014) could not be extracted due to missing information.
Hasanpour (2017) did not include disability outcomes.
Kofotolis (2016) did not include pain outcomes.
Abbreviations: wk: week; wks: weeks.

TABLE 4. Summary of findings - GRADE level of evidence for studies †

Outcome	Participants	Design (studies)	Measurement	Risk of bias	Inconsistency	Indirectness	Imprecision	Certainty of evidence
Subgroup 1: Pilates exercise vs General exercise								
Pain	225	4 RCTs	NRS, VAS	Moderate	Not serious	Not serious	Serious	Low
Disability	225	4 RCTs	ODI, QBP	Moderate	Not serious	Not serious	Serious	Low
Subgroup 2: Pilates exercise vs Direction-specific exercise								
Pain	55	2 RCTs	MPQ, VAS	High	Not serious	Not serious	Serious	Very low
Disability	31	1 RCT	ODI	High	NA	Not serious	Serious	Very low
Subgroup 3: Pilates exercise vs Spinal stabilisation exercise								
Pain	89	3 RCTs	NRS, VAS	High	Very serious	Not serious	Serious	Very low
Disability	162	4 RCTs	ODI, RMDQ	High	serious	Not serious	Serious	Very low

GRADE: Grading of Recommendations, Assessment, Development and Evaluations; RCT: Randomised controlled trials; NA: Not applicable; RMDQ: Roland Morris Disability Questionnaire; ODI: Oswestry Disability Index; QBPQ: Quebec Back Pain Questionnaire; NRS: Numeric Rating Scale; VAS: Visual Analogue Scale; MPQ: McGill Pain Questionnaire.

† The grading followed the principles as below

(1) Risk of bias: the certainty of the evidence was rated down if more than 25% of the participants were from studies with a high risk of bias.

(2) Inconsistency: the certainty of the evidence was downgraded if there was a significant heterogeneity (I2 > 50%) presented or there was a large difference in the estimate of effects.

(3) Indirectness: the certainty of the evidence was downgraded if more than 50% of the participants were out of the target population of interest (individuals with CNSLBP).

(4) Imprecision: the certainty of the evidence was downgraded if the total number of participants was less than 400 for each continuous outcome. The imprecision of the evidence was also considered inconsistent.

The certainty of the evidence was determined by first considering the imprecision. If the evidence was imprecise, the certainty was judged to be 'low' regardless of the rest of the domains. Then, the certainty of the evidence was further downgraded to be 'very low' if there was a potential risk of bias and indirectness. The certainty of the evidence could be interpreted as follows:

- High level: Further research is very unlikely to change our confidence in the estimate of effect. There is sufficient data with narrow confidence intervals. There are no known or suspected reporting biases.
- Moderate level: Further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate; one of the domains is not met.
- Low level: Further research is very likely to have an important impact on confidence in the estimate of effect and is likely to change the estimate; two of the domains are not met.

- Very low level: Great uncertainty about the estimate; three of the domains are not met

Appendix A – STARLITE strategy

Sampling strategy	Selective: attempts to identify all relevant studies within specified limits
Type of study	Randomised control study
Approaches	Database search, citation search
Range of year	Inception – 30 April 2021
Limits	Human studies, English
Inclusion/ Exclusion	Inclusion Published RCTs comparing the effects of Pilates exercise (PE) with other forms of exercise were eligible for inclusion. Variation of PE was accepted, including PE on a mat or on an apparatus (such as Cadillac and Reformer). Co-interventions were accepted only if they were added into both the experiment group (PE) and comparison group (other forms of exercise). Eligible studies were screened using the selection criteria (framed by PICO search tool) through the abstract and full text. Studies were included only if • (1) their participants (18 years of age or older) were symptomatic with non-specific LBP lasting for at least 12 weeks • (2) an exercise named explicitly as 'Pilates exercise' was used in the trial • (3) PE and interventions with exercise components were compared in the trial and • (4) either pain or disability was measured as an outcome. Exclusion Non-English and unpublished studies were excluded. Studies including individuals with LBP as a secondary problem from other comorbidities or specific causes (such as scoliosis, systemic inflammatory disease, and trauma) were excluded.
Terms used	Note: Sensitivity-maxmising strategy for low back pain and RCTs recommended by Cochrane was used for main databases (MEDLINE and EMBASE). Search terms "Pilate*" and "Pilates" were used, aiming to search for interventions explicitly named as "Pilates". This aligned with the inclusion criteria of this review that interventions were only included if it was explicitly named as "Pilates" **Medline (Ovid) & EMBASE (Ovid)** 1. randomi?ed controlled trial.mp. 2. controlled clinical trial.mp. 3. randomi?ed.mp. 4. placebo.mp. 5. clinical trials as topic/ 6. randomly.mp. 7. trial*.mp. 8. 1 or 2 or 3 or 4 or 5 or 6 or 7 9. (animals not humans).mp. 10. 8 not 9 11. dorsalgia.mp. 12. exp back pain/ 13. backache.mp. 14. exp low back pain/ 15. (lumbar adj pain).mp.

	16. coccyx.mp. 17. coccydynia.mp. 18. sciatica.mp. 19. sciatic neuropathy/ 20. spondylosis.mp. 21. lumbago.mp. 22. back disorder*.mp. 23. 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 24. (pilates or pilate).mp. 25. 23 and 24 26. 10 and 25 **PEDro** Title and Abstract: back pain AND pilate* Method: Clinical trial **CENTRAL** Title Abstract Keyword: low back pain or dorsalgia or *spin* pain or back ache or lumbgo in AND Title Abstract Keyword: pilate* or pilates method in AND Publication Type: randomi?ed controlled trial* or controlled clinical trial* (32) **CINAHL (EBSCOhost) and SPORTDiscus (EBSCOhost)** S1. TI low back pain or lumbar pain or lumbar spine pain or non specific low back pain or chronic low back pain dorsalgia or *spin* pain or backache or lumbago S2. AB low back pain or lumbar pain or lumbar spine pain or non specific low back pain or chronic low back pain dorsalgia or *spin* pain or backache or lumbago S3. MW low back pain or lumbar pain or lumbar spine pain or non specific low back pain or chronic low back pain dorsalgia or *spin* pain or backache or lumbago S4.(S1 OR S2 OR S3) S5. TI pilate* S6. AB pilate* S7. MW pilate* S8. (S5 OR S6 OR S7) S9. (S8 AND S4) S10. PT randomi?ed controlled trial* or controlled clinical trial* S11. TI randomly or placebo or trial or randomi?ed S12. AB randomly or placebo or trial or randomi?ed S13. (S10 OR S11 OR S12) S14. (S13 AND S9)
Electronic Sources	MEDLINE Ovid, PEDro, CENTRAL, EMBASE, CINAHL, SPORTDiscus

Appendix B – Risk of bias assessment

Assessment criteria

For a study to be given 'low risk', the study needed to be judged to be at low risk in all five domains. Studies were judged to be at 'some concerns' or 'high risk' if they had at least one domain resulting in either 'some concerns' or 'high risk' respectively. The judgment was made based on the algorithm suggested by the RoB2 tool (Sterne et al., 2019).

The mentioned RoB2 tool and algorithm can be found on the following links:
https://methods.cochrane.org/risk-bias-2
https://www.riskofbias.info/

Appendix C – Exclusion of Anand at el (2014) due to questionable data

The data on pain and disability from Anand at el (2014) was exclused from the meta-analysis due to insufficient information on the data. The authors reported the data in a very brief plain text without any information regarding the data analysis. It was unclear whether the data reported was a mean change or a post-intervention measurement. It was also impossible to understand the numbers reported in the text and to input for meta-analysis. An attempt was made to contact the trial authors for extra information but there was no reply. Thus, the data reported from this trial was considered not suitable to include in the meta-analysis, which aimed at comparing mean changes in pain across studies.

Appendix D – Exclusion of Dsa et. al (2014) due to questionable data

The data on pain from Dsa et. al (2014) was considered problematic. In Dsa et al. (2014), data was only presented in a table with no details of data analysis. Although the author did not report the nature of the data (mean change or post-intervention measurement), the data appeared to be a change of the mean in pain score and thus was input into the meta-analysis. Contact was made to the responding author but no reply was received. Based on the above questionable data, the result in this subgroup was pooled excluding Dsa et al. (2014) considering as a statistical outlier for the analysis.

In Dsa et al. (2014), disability was measured by RMDQ which has a maximum score of 24. However, the data of disability presented in this trial did not match with the scale used. The average maximum score presented was 67 and the mean was 39.92. The data was questionable and deemed to be inappropriate to include in the meta-analysis.

Reference

Anand, U. A., Caroline, P. M., Arun, B., & Gomathi, G. L. (2014). A study to analyse the efficacy of modified pilates based exercises and therapeutic exercises in individuals with chronic non specific low back pain: a randomized controlled trial. *International journal of physiotherapy and research*, *2*(3), 525-29.

Dsa, C. F., Rengaramanujam, K., & Kudchadkar, M. S. (2014). To assess the effect of modified pilates compared to conventional core stabilization exercises on pain and disability in chronic non-specific low back pain-randomized controlled trial. *Indian Journal of physiotherapy and occupational therapy*, *8*(3), 202.

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*, *366*.