SYSTEMATIC REVIEW

Effectiveness of conservative therapy on pain, disability and quality of life for low back pain in pregnancy: A systematic review of randomized controlled trials

Laísa B. Maia\textsuperscript{a}, Letícia G. Amarante\textsuperscript{b}, Débora F.M. Vitorino\textsuperscript{b}, Rodrigo O. Mascarenhas\textsuperscript{b}, Ana Cristina R. Lacerda\textsuperscript{a}, Bianca M. Lourenço\textsuperscript{a}, Vinicius C. Oliveira\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a} Postgraduate Program in Rehabilitation and Functional Performance, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Diamantina, Brazil
\textsuperscript{b} Department of Physical Therapy, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Diamantina, Brazil

Received 8 May 2020; received in revised form 28 May 2021; accepted 30 June 2021
Available online 22 July 2021

Abstract

Background: Efficacy of conservative therapy for low back pain in pregnancy (PLBP) is unclear.
Objective: To investigate the efficacy of conservative therapy on pain, disability, and quality of life in PLBP.
Methods: The protocol of this systematic review was prospectively registered at PROSPERO (CRD42020164640). Search strategy was conducted on six databases up to August 24 2020 without date or language restrictions. Minimal intervention (i.e., placebo, sham, waiting list or no intervention) was the comparator of interest. Selection of randomized controlled trials, data extraction and methodological quality assessment of included trials were conducted independently by two reviewers. The PEDro scale (0–10) was used to assess methodological quality. Effect sizes for specific therapies were pooled when possible, using random-effects models. The quality of the evidence was assessed using the Grading of Recommendations Assessment (GRADE) approach.
Results: Ten included trials provide uncertain evidence (low to very low quality) about the effects of auriculotherapy, education, exercise, exercise plus education, oil treatment, and osteopathy in pain, disability, and quality of life at short- and long-term. At short-term, mean differences (MDs) and 95% confidence intervals (CI) on a 0–10 points pain intensity scale were: for oil treatment, 2.8 points (2.6, 3.1) (\(n\) = one trial, 114 participants); for auriculotherapy, 1.6 points (1.2, 2.0) (\(n\) = one trial, 112 participants); for exercise, 2.2 points (-1.8, 6.2) (\(n\) = three trials, 297 participants).
Conclusion: There is an urgent need for larger high-quality trials investigating effects of conservative therapy in pain, disability, and quality of life in this population.

© 2021 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier España, S.L.U. All rights reserved.

KEYWORDS

Gestation; Low back pain; Meta-analysis; Non-invasive

https://doi.org/10.1016/j.bjpt.2021.06.007
Introduction

Low back pain (LBP) is defined as pain in the posterior region of the body, between the lower margin of the twelfth rib and the gluteal fold, lasting, at least, one day with or without leg pain.\(^1\) LBP in pregnancy (PLBP) and pelvic girdle pain (PGP) are considered public health issues, with prevalence ranging from 4% to 76%.\(^2\) It causes disability to pregnant women, costs, and affects their quality of life.\(^3,4\)

Guidelines for management of PLBP are still reporting unclear evidence\(^5,6\) and current evidence supporting therapies for general population with LBP\(^7\) might not be appropriate as it could be harmful to pregnant women and the fetus.\(^8-10\) For instance, pharmacological therapies such as non-steroidal anti-inflammatory and opioids that have been recommended by some guidelines to the general population with LBP\(^10,11\) are contraindicated in pregnancy due to potential long-term effects on the fetus.

Previous reviews\(^2,12-20\) investigated effects of conservative therapy in PLBP (e.g., exercise) and these are considered the first treatment option to improve symptoms, disability, and quality of life of this population.\(^2,13,19,20,22\) However, evidence is limited by: comparing therapies with other active interventions,\(^2,13,14,20,21\) lack of investigation of the quality of the current evidence,\(^14\) trials combining PLBP with pelvic pain\(^12\) and reviews without systematic design.\(^2,13\) Therefore, efficacy of conservative therapy in PLBP is still unclear. The aim of this systematic review of randomized controlled trials was to investigate the effectiveness of conservative therapy on pain, disability, and quality of life in PLBP and to summarize the quality of the current evidence using the Grading of Recommendations Assessment (GRADE) approach.\(^24\)

Methods

Study reporting and protocol registration

This systematic review of randomized controlled trials followed the Cochrane Handbook for Systematic Reviews of Interventions\(^25\) and the Updated Method Guideline for Cochrane Back and Neck Group.\(^26\) Furthermore, this systematic review was reported based on the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) checklist.\(^27\) The protocol was prospectively registered at PROSPERO (CRD42020164640) and at the Open Science Framework ([https://osf.io/4yspu/].\(^28\)

Search strategy

Search strategy was conducted on MEDLINE, Embase, AMED, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews, and Physiotherapy Evidence Database (PEDro) from the date of inception to 24 August 2020 without date or language restrictions. The search terms were related to ‘randomized controlled trial’ and ‘low back pain’. Descriptors related to conservative therapy were not used to increase sensitivity of our search strategy, avoiding exclusions of potential conservative therapy that we were unaware of. In addition, we hand searched previous systematic reviews identified in the field for potentially relevant full text articles. Detailed search strategy is in the Supplemental Online Material.

Inclusion criteria

Study design and participants

Randomized controlled trials (i.e., parallel-group; crossover, and cluster randomized controlled trials) investigating participants with PLBP were included. Trials investigating PGP or PLBP mixed with PGP were excluded.

Interventions

To be included, trials had to investigate effects of any stand-alone conservative therapy (i.e., defined as any kind of non-invasive intervention including pharmacological therapies) or combined therapies.

Comparisons

Comparator of interest was minimal intervention (i.e., placebo, sham, waiting list, or no intervention) to investigate the isolated effect of therapies of interest in PLBP. Trials using usual/standard care as comparator group only were included when participants did not receive any type of active intervention (i.e., only exams and/or routine prenatal consultation); for the purpose of this review, this type of comparator was labelled as “no intervention”.

Outcomes

Primary outcomes of interest stated a priori were pain intensity (e.g., Numerical Rating Scale – NRS\(^20\), Visual Analog Scale – VAS\(^20\)); disability (e.g., Oswestry Disability Index – ODI\(^29\), Quebec Back Pain Disability Scale - QBSPDS\(^30\)); and quality of life (e.g., Short Form Health Survey-36).\(^31\) Pain intensity, disability, and quality of life could have been assessed by any validated instrument.

Selection of trials

After the searches, references were exported to an EndNote file (Software EndNote X8 Thomson Reuters, Philadelphia, PA, USA), duplicates were removed, and two independent reviewers (LBM and LGA) screened titles and abstracts and assessed potential full texts according to our eligibility criteria.

Assessment of methodological quality of included trials

Methodological quality was assessed independently by two reviewers (LBM and ROM) using the Physiotherapy Evidence Database (PEDro) scale (0–10), where 2–9 items were used specifically for the assessment of methodological quality, and 10 and 11 items were used for the statistical reporting.\(^32\) Higher scores indicate low risk of bias.\(^33-36\) The PEDro scale is widely used to evaluate risk of bias,\(^30\) and is valid and reliable.\(^33\) When possible, we used scores already available in the PEDro database ([https://www.pedro.org.au/].\(^33-35\) Between-reviewer discrepancies were resolved by a third author (VCO).
Data extraction

Two independent reviewers (BML and LBM) extracted characteristics and outcome data from the included trials. A third author (VCO) resolved between-reviewer discrepancies. Data extracted included: bibliometric data, source of participants; sample size; age; description of conservative therapy and comparator; outcomes; instrument measures; and follow-ups. Outcome data extracted included: sample sizes; means; and standard deviations (SDs) for all intervention groups.

Some outcome data were not reported in the included trials and authors were contacted twice with one week interval for further information. When authors did not answer, means and/or SDs were imputed using confidence intervals, p-value, and data presented in graphs, following the Cochrane Handbook for Systematic Reviews of Interventions. When trials investigated more than one similar conservative therapy compared with minimal intervention, we combined outcome data also following the Handbook. Placebo or sham were our first options for comparison when more than one comparator group was investigated in the same trial. For cross-over trials, we only considered results from the first randomization period to avoid carry-over effects.

Data analysis and synthesis

Meta-analysis was conducted using random-effects model. Mean differences (MDs) for individual studies or pooling results and 95% confidence intervals (CIs) were presented for each specific conservative therapy in forest-plots for short- and long-term effects. All analyses were conducted using Comprehensive Meta-analysis software, version 2.2.04 (Biostat, Englewood, NJ). Short-term effect was considered follow-up up to 12 weeks, and long-term effect was considered follow-up more than 12 weeks (time points were considered during the prenatal period). When more than one time point was available within the same follow-up period, the one closer to the end of the intervention was considered. When the same outcome was measured with different scales, we converted them to a similar scale before pooling (i.e., a 0–100 points scale). We considered the presence of heterogeneity when $I^2 > 50\%$ or when there was no overlapping among the confidence intervals of individuals trials included in the meta-analysis. The MCID values for this specific population were not found and we considered: small effect for estimates $<10\%$ of the scale; medium effect for estimates $\geq 10\%$ up to $20\%$ of the scale; and large effect for estimates $>20\%$ of the scale.

Two independent reviewers (LB and ROM) assessed the quality of the evidence using the GRADE approach, and between-reviewer discrepancies were resolved by a third author (VCO). We considered the following GRADE domains: imprecision; risk of bias; inconsistency, and publication bias. Quality of the evidence ranges from high to very low quality (i.e., high, moderate, low, and very low) and starts as high quality but can be downgraded by one or two levels in each of the domains. Lower quality levels indicate uncertainty (future high-quality trials are likely to change estimates).

For imprecision, quality of evidence was downgraded by one level (i.e., serious) if the sample analyzed was $<400$ participants and downgraded by two levels (i.e., very serious) if the sample analyzed was $<200$ participants. For risk of bias, we downgraded the quality of evidence by one level when $>25\%$ of the participants were from trials with a high risk of bias (i.e., PEDro score $<6$ out of 10). For inconsistency, we downgraded by one level (i.e., serious) when $I^2$ statistics was $\geq 50\%$ and we downgraded by two levels (i.e., very serious) if analyzed trials showed poor overlapping of confidence intervals. Effect estimates provided by single randomized trials (with $<400$ participants) were considered inconsistent and imprecise (i.e., sparse data) and provided low quality evidence. This could be further downgraded to very low quality evidence if sample size $<200$ participants or if there were also limitation in design. We planned to evaluate publication bias using the visual inspection of funnel plots and the Egger’s test adopting an $\alpha = 0.1$; however, it was not possible because of the small number of included trials (i.e., $<10$ trials analyzed).

We also planned sensitivity and subgroup analyses to investigate whether poor methodological quality, removing trials with high risk of bias from the meta-analysis, $<6$ points on PEDro scale (0–10), and characteristics of participants or therapy dosage impacted on the estimates.

Results

The flow of studies through the review is summarized in Fig. 1. After identification of 2266 records, 61 potential full texts were assessed, and 11 original trials were included in the qualitative analysis (total of 1854 participants). From the 11 included trials, one (total of 106 participants) was excluded from the quantitative analysis because data were not reported and imputations were not possible. When possible, means and/or SDs were imputed using confidence intervals, $p$-value and data presented in graphs.

Characteristics of included trials

The mean ± SD age of the participants in the included trials was 26.5 ± 4.7 years old. Trials included in the qualitative analysis were conducted in Europe, Asia, and North America. They investigated different forms of conservative therapy: exercise ($n = 440$ participants), exercise combined with education ($n = 377$ participants); education ($n = 134$ participants); oil treatment ($n = 120$ participants); osteopathy ($n = 400$ participants); tapping ($n = 106$ participants) and auriculotherapy ($n = 159$ participants). Pain was investigated in 10 trials (90.9%, $n = 10$) and disability was investigated in seven trials (63.6%, $n = 7$). Quality of life was investigated in three trials (27.3%, $n = 3$) out of the 11 included trials.

Instruments used to evaluate the outcomes of interest were: VAS, NRS, and the Iranian version of the Quebec Back Pain Disability Scale (KEBK) for pain; ODI and Disability Rating Index (DRI) for disability; and the Roland Morris Disability Questionnaire (RMDQ) for disability; and Short Form Health Survey-36 (SF-36) and quality of life. Short-term effects were investigated in 10
trials (90.9%, n = 10), and long-term effects were investigated in two trials (18.2%, n = 2). Detailed characteristics of included trials are in Table 1.

Methodological quality and statistical reporting of included trials

The median score on the PEDro scale (0–10) for the included trials was 6, with scores ranging from 4 to 7 points. Four of the 11 included trials had high risk of bias (i.e., PEDro score <6 out of 10) (Table 2). None of the included trials blinded therapists (100%), and nine of the 11 trials did not blind participants and/or assessors (81.8%).

Effectiveness of conservative therapy on pain, disability, and quality of life for low back pain in pregnancy

Fig. 2 reports current evidence from the quantitative analysis for pain, disability, and quality of life at short- and long-term. The estimates are presented as MDs with their 95% CIs on a 0–10 points scale for pain and on a 0–100 points scale for disability and quality of life. The evidence is very
<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Source</th>
<th>Patient characteristic</th>
<th>Intervention</th>
<th>Outcome measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandpei et al. (2010)</td>
<td>Patients in gestational age 17 to 22 week, at least 12 weeks of LBP during pregnancy. Location: Iran (Asia).</td>
<td>$N = 120$ Age: N/A</td>
<td>Exp: Illustrated booklet on the exercises and ergonomic principles; 5 educational workshops, 20 min each, on the abdominal and back muscles, strengthening and stretching exercises; led by an expert midwife and a physiotherapist. Follow-up telephone calls to assure the sustainability of intervention ($N = 60$; age: N/A). Con: No intervention ($N = 60$; age: N/A).</td>
<td>Disability: ODI (0 - 100) Pain: VAS (0 - 100) Follow up: 12 weeks (short-term)</td>
</tr>
<tr>
<td>Eggen et al. (2012)</td>
<td>Patients between 18 and 40 years of age before gestation week 20. Location: Norway (Europe).</td>
<td>$N = 257$ Age: 30.3 ± 4.8</td>
<td>Exp: supervised exercises, including ergonomic advice, in groups and were advised to do home exercises. Each weekly group exercise session lasted 60 min, and the groups trained for 16 to 20 weeks (between gestation weeks 16 and 36). ($N = 129$; age: 30.6 ± 4.8). Con: No intervention ($N = 128$; age: 30.0 ± 4.8).</td>
<td>Disability: RMDQ (0 - 24) Pain: NRS (0 - 10) Quality of life: SF-8 Follow up: 12 weeks (short-term) and 20 weeks (long-term)</td>
</tr>
<tr>
<td>Holden et al. (2019)</td>
<td>Patients aged 18 to 39 years with uncomplicated pregnancies at 12 to 26 weeks Location: United States (North America).</td>
<td>$N = 20$ Age: 31.4 ± 4.7</td>
<td>Exp: Yoga procedure. 1 session per week ($N = 11$; age: 29.6 ± 5.1). Con: No intervention ($N = 9$; age: 33.4 ± 3.5).</td>
<td>Disability: RMDQ (0 - 24) Pain: VAS (0 - 10) Quality of life: SF-12 (0 - 100) Follow up: 12 weeks (short-term)</td>
</tr>
<tr>
<td>Kalinowski and Krawulska (2017)</td>
<td>Patients in the second and third trimesters of pregnancy with low back pain, who were patients of gynecological offices, and did not have any contraindications. Location: Poland (Europe)</td>
<td>$N = 106$ Age 29.5 ± 4.2</td>
<td>Exp: Elastic tape - 5 days of maintenance and the women were asked to remove the tape on their own. ($N = 53$; age: N/A). Con: Placebo elastic tape - 5 days of maintenance and the women were asked to remove the tape on their own. ($N = 53$; age: N/A)</td>
<td>Disability: RMDQ (0 – 24) Pain: VAS (0 - 10) Follow up: 1 week (short-term)</td>
</tr>
<tr>
<td>Author/Year</td>
<td>Source</td>
<td>Patient characteristic</td>
<td>Intervention</td>
<td>Outcome measurement</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------------------------------</td>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Kihlstrand et al. (1999)</td>
<td>Patients between weeks 15–18 of gestation. Location: Sweden (Europe).</td>
<td>N = 252 Age 28.5</td>
<td>Exp: Water-gymnastics group. 17–20 times (once a week), one hour (N = 124; age: 28). Con: No intervention (N = 128; age: 29). Exp1: Participants performed low-to-moderate intensity resistance exercise training twice per week for 12 weeks and, depending on availability, training was performed individually or in small groups (≤4 per group). Strength training sessions were supervised by an experienced exercise specialist. (N = 44; age: 28 ± 5).</td>
<td>Pain: VAS (0 - 10) Follow up: 16 weeks (long-term)</td>
</tr>
<tr>
<td>O’Connor et al. (2018)</td>
<td>Patients in the second trimester of pregnancy. Location: Georgia (Europe).</td>
<td>N = 134 Age: 28.7 ± 4.3</td>
<td>Exp1: Maternity nurses taught six bimonthly pregnancy education classes (20 per class, 60 min each) which covered several topics about pregnancy (N = 45; age: 29 ± 4).</td>
<td>Quality of life: SF-36 Follow up: 12 weeks (short-term)</td>
</tr>
<tr>
<td>Shirazi et al. (2017)</td>
<td>Patients at 12 to 33 weeks of gestation. Location: Iran (Asia).</td>
<td>N = 120 Age: 28 ± 4.2</td>
<td>Exp1: Rose oil (+carrier oil); 7 drops of oils topically for 100 cm² of the painful part of skin without massage, 2 times daily for 4 weeks + standard prenatal care (N = 40; age: 27.7 ± 4.9). Exp2: Rose oil (carrier almond); 7 drops of oils topically for 100 cm² of the painful part of skin without massage, 2 times daily for 4 weeks + standard prenatal care (N = 40; age: 27.9 ± 4.3).</td>
<td>Disability: RMDQ (0–24) Pain: VAS (0–10) Follow up: 4 weeks (short-term)</td>
</tr>
<tr>
<td>Suputtitada et al. (2002)</td>
<td>Patients at 26 to 36 weeks of gestation aged 20 to 35 years. Location: Thailand (Asia).</td>
<td>N = 74 Age: N/A</td>
<td>Exp: The Sitting pelvic tilt exercise: 2 times per day, one in the morning and one in the evening, 5 days per week for 8 weeks. (N = 32; age N/A).</td>
<td>Pain: VAS (0 - 10) Follow up: 8 weeks (short-term)</td>
</tr>
<tr>
<td>Wang et al. (2009)</td>
<td>Patients with a gestational age of 25–38 weeks, who had lower back and/or posterior pelvic pain. Location: United States (North America).</td>
<td>N = 159 Age: 32.4 ± 5</td>
<td>Exp: Acupuncture-auricular press needles at 3 points (N = 58; age: 33 ± 5). Con1: Sham acupuncture-auricular press needles at 3 nonspecific auricular acupuncture points. (N = 54; age: 32 ± 5).</td>
<td>Disability: DRI (0 - 100) Pain: VAS (0 - 100) Follow up: 2 weeks (short-term)</td>
</tr>
</tbody>
</table>

Data are mean ± standard deviation. Con = comparison; Exp = experimental group; N = sample size; N/A = not available; LBP = low back pain; VAS = Visual Analog Scale; RMDQ = Roland Morris Disability Questionnaire; ODI = Oswestry Disability Index; DRI = Disability Rating Index; SF = Short Form Health Survey; QBPDS = Quebec Back Pain Disability Scale. KEBK = Iranian version of the Quebec Back Pain Disability Scale.
uncertain about the effect of oil treatment and auriculotherapy on pain and disability at short-term when compared with minimal intervention. The quality of the evidence for oil treatment was downgraded by serious risk of bias and inconsistency, and very serious imprecision, whereas for auriculotherapy, it was downgraded by very serious imprecision and serious inconsistency. MDs were: 2.8 points (95%CI: 2.6, 3.1) on a 0–10 points pain intensity scale (n = one trial, 114 participants) and 18.6 points (95%CI: 17.1, 20.1) on a 0–100 points disability scale (n = one trial, 114 participants) for oil treatment; and 1.6 points (95%CI: 1.2, 2.0) on pain (n = one trial, 112 participants) and 9.5 points (95%CI: 6.5, 12.6) on disability (n = one trial, 112 participants) for auriculotherapy.

In addition, evidence is very uncertain about the effect of other investigated therapies and may result in little to no difference in pain, disability, and quality of life when compared with minimal intervention at short- and long-term (low to very-low quality evidence). For example: MDs for exercise were 2.2 points (−1.8, 6.2) on pain (n = three trials, 297 participants) at short-term and 0.6 points (−0.3, 1.5) on pain (n = one trial, 244 participants) at long-term; 0.8 (−3.1, 4.7) on disability (n = one trial, 20 participants) at short-term; 2.2 (−4.1, 8.5) on quality of life (n = two trials, 109 participants) at short-term. The reasons for downgrading the quality of the evidence of all the estimates were imprecision (24 levels), inconsistency (19 levels) and risk of bias (five levels). Detailed MDs and 95%CI s are reported in Fig. 2. Table 3 summarizes the investigated therapies, their estimates, and their current quality of the evidence.

**Discussion**

This systematic review showed that the available evidence supporting conservative therapy for pain, disability, and quality of life in PLBP is scarce, mainly at long-term. The quality of the current evidence ranged from low to very low, so estimates are uncertain and likely to change with future high-quality trials with larger samples. There is an urgent need to clarify evidence for therapies such as oil treatment, auriculotherapy, and exercise.

Our systematic review shows that the effect of exercise is very uncertain on pain at short-term, even though previous systematic reviews found positive effects for exercise in PLBP. Exercise is recommended in pregnancy to decrease complications (urinary incontinence, gestational diabetes mellitus, hypertension, and pre-eclampsia). Recommended exercises for pregnant women include aerobic and resistance training (strong recommendation) and Kegel exercises (weak recommendation), at least 150 min of moderate-intensity or over, a minimum of three days per week. A potential explanation for the discrepancy between our findings for exercise and the previous evidence...
is the methodology of our systematic review that only considered the usual/standard care as comparator group if participants did not receive any type of active intervention (i.e., only exams and/or routine prenatal consultation), and excluded trials investigating PGP.

Previous systematic reviews investigated the effects of some conservative therapies (e.g., exercise and acupuncture) but it included PLBP and PGP (i.e., clinical heterogeneity) and comparison was not appropriate to investigate efficacy.10 Comparisons with other conservative therapy provide information about whether a specific conservative intervention is superior to another conservative intervention but do not provide evidence of efficacy because we do not have a comparison with minimal or no intervention. Comparison with minimal intervention (i.e., placebo, sham, and waiting list) or no intervention must be the primary option to investigate efficacy. We believe that the current review improved the methodological standard, and it is vital in this field where competing therapies, many of little value, are commonly used in clinical practice. Furthermore, the lack of trials investigating effects of noninvasive pharmacological therapies during pregnancy does not surprise due to the evidence showing their harmful effects in the gestational period mainly for the fetus, making these therapies not the first choice of treatment.8–10

Our systematic review followed current methodological guidance from Cochrane and the GRADE working group to update and synthesize the evidence on efficacy for all available conservative therapies for PLBP. By reporting comparable effect estimates and a rating of the quality of evidence for each therapy, this review provides patients and clinicians with reliable information to contribute to their decision-making processes. As potential limitations, although the efficacy of elastic taping41 was investigated by one included trial, it was not possible to be included in our quantitative analysis because the first phase results of this cross-over trial was not reported separately. We tried to contact the authors requesting the necessary data, but they did not reply. Moreover, it was not possible to conduct sensitivity or subgroup analyses to explore potential impact of high risk of bias, dosage of the conservative therapy, and characteristics of the population because of the small number of included trials. For instance, as discussed above, exercise would be important for PLBP and future research should clarify the use of exercise specifically for PLBP compared with minimal intervention. Another potential limitation was related to our comparator, we excluded trials not clearly reporting that the usual/standard care did not involve any type of active intervention. In this way, four trials did not report the information and we had no replies from authors. One
<table>
<thead>
<tr>
<th>Certainty assessment</th>
<th>Study design</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Imprecision</th>
<th>No. of participants</th>
<th>Mean Difference</th>
<th>Quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Experimental</td>
<td>Minimal intervention</td>
<td></td>
</tr>
<tr>
<td>A. Pain: Auriculotherapy vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Serious</td>
<td>Very serious</td>
<td>58</td>
<td>54</td>
<td>1.6 (1.2, 2.0)</td>
</tr>
<tr>
<td>B. Disability: Auriculotherapy vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Serious</td>
<td>Very serious</td>
<td>58</td>
<td>54</td>
<td>9.5 (6.5, 12.6)</td>
</tr>
<tr>
<td>A. Quality of life: Education vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Serious</td>
<td>Very serious</td>
<td>45</td>
<td>45</td>
<td>–0.0 (–11.2, 11.1)</td>
</tr>
<tr>
<td>A. Pain: Exercise vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>149</td>
<td>148</td>
<td>2.2 (–1.8, 6.2)</td>
</tr>
<tr>
<td>B. Pain: Exercise vs minimal intervention (follow up &gt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>124</td>
<td>120</td>
<td>0.6 (–0.3, 1.5)</td>
</tr>
<tr>
<td>C. Disability: Exercise vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>11</td>
<td>9</td>
<td>0.8 (–3.1, 4.7)</td>
</tr>
<tr>
<td>D. Quality of life: Exercise vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>55</td>
<td>54</td>
<td>2.2 (–4.1, 8.5)</td>
</tr>
<tr>
<td>A. Pain: Exercise plus education vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>174</td>
<td>179</td>
<td>1.2 (–0.8, 3.4)</td>
</tr>
<tr>
<td>B. Pain: Exercise plus education vs minimal intervention (follow up &gt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>103</td>
<td>107</td>
<td>0 (–0.6, 0.7)</td>
</tr>
<tr>
<td>C. Disability: Exercise plus education vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>117</td>
<td>124</td>
<td>9.2 (–9.7, 28.2)</td>
</tr>
<tr>
<td>D. Disability: Exercise plus education vs minimal intervention (follow up &gt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>103</td>
<td>107</td>
<td>0.2 (–0.9, 1.3)</td>
</tr>
<tr>
<td>E. Quality of life: Exercise plus education vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>117</td>
<td>124</td>
<td>–0.3 (–2.5, 1.9)</td>
</tr>
<tr>
<td>F. Quality of life: Exercise plus education vs minimal intervention (follow up &gt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>103</td>
<td>107</td>
<td>0.2 (–2.3, 2.7)</td>
</tr>
<tr>
<td>A. Pain: Oil treatment vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>75</td>
<td>39</td>
<td>2.8 (2.6, 3.1)</td>
</tr>
<tr>
<td>B. Disability: Oil treatment vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>75</td>
<td>39</td>
<td>18.6 (17.1, 20.1)</td>
</tr>
<tr>
<td>A. Pain: Osteopathy vs minimal intervention (follow up &lt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>136</td>
<td>131</td>
<td>0 (–0.3, 0.4)</td>
</tr>
<tr>
<td>B. Disability: Osteopathy vs minimal intervention (follow up &gt; 12 weeks)</td>
<td>RCT</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Serious</td>
<td>136</td>
<td>131</td>
<td>–0.5 (–0.7, 0.7)</td>
</tr>
</tbody>
</table>

CI, confidence interval; RCT, randomized controlled trial; minimal intervention, placebo, sham, waiting list or no intervention.

a Downgraded due to imprecision: less than 400 participants included in the meta-analysis (sample of less than 200 was considered very serious imprecision and downgraded in two levels).

b Downgraded due to inconsistency: I² statistic was higher than 50% or pooling was not possible (poor overlap among confidence intervals of analyzed trials was considered very serious and downgraded in two levels).

c Downgraded due to serious risk of bias: more than 25% of the analyzed participants were from trials with a high risk of bias (i.e., PEDro score <6 of 10).
excluded trial compared usual/standard care with exercise and found effects on reducing pain.53 One trial comparing reflexology with usual/standard care found effects on reducing pain.54 Two trials compared osteopathy with usual/standard care on disability and found medium to large effects on reducing disability in PLBP.55,56 We recommend future trials to clearly describe all investigated groups. The other limitation of the review was that we did not investigate potential adverse events of the investigated therapies, which should be investigated in future research.

Conclusion
Our systematic review shows that the evidence is very uncertain about the effect of conservative therapy (e.g., oil treatment, auriculotherapy, and exercise) in pain, disability, and quality of life at short- and long-term. Future larger high-quality trials are needed to clarify the evidence.

Conflicts of interest
None.

Acknowledgments
None. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials
Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.bjpt.2021.06.007.

References


