Therapeutic ultrasound for knee osteoarthritis: A systematic review and meta-analysis with grade quality assessment

Lucas Ogura Dantas\textsuperscript{a,b}, Mikala C. Osani\textsuperscript{c}, Raveendhara R. Bannuru\textsuperscript{c,d,*}

\textsuperscript{a} Division of Rheumatology, Tufts Medical Center, Boston, MA, USA
\textsuperscript{b} Physical Therapy Department, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil
\textsuperscript{c} Center for Treatment Comparison and Integrative Analysis, Division of Rheumatology, Tufts Medical Center, Boston, MA, USA
\textsuperscript{d} Center for Complementary and Integrative Medicine, Division of Rheumatology, Tufts Medical Center, Boston, MA, USA

Received 4 March 2020; received in revised form 1 July 2021; accepted 30 July 2021
Available online 9 September 2021

Abstract

Background: Therapeutic ultrasound (US) is a widely used intervention in physical therapy to manage pain and to aid in the healing of soft tissue.

Objective: This systematic review aimed to determine the effects of therapeutic US on knee osteoarthritis (KOA) symptoms.

Methods: PubMed, MEDLINE, EMBASE, Google Scholar, and the Cochrane databases were searched from inception to April 2019. Randomized controlled trials (RCTs) involving adults with symptomatic KOA that compared therapeutic US with a sham or other control were included. The methodological quality of the trials was assessed at the study level using the Cochrane Risk of Bias tool. The quality of evidence at the outcome level- and overall- was assessed using GRADE methodology. Meta-analyses were conducted using random effects models, and heterogeneity was assessed using the \(I^2\) statistic.

Results: Four studies (\(N = 234\) participants) were eligible for inclusion in our primary analyses assessing therapeutic US versus sham. The methodological quality of the included RCTs ranged from moderate to very low. Treatment with therapeutic US resulted in small, statistically significant benefits for pain (approximate 9.6\% improvement on a 0–100 visual analog scale [95\% confidence interval: 2, 17.4\%]) and self-reported measures of function (approximate 12.8\% improvement on a 0–100 visual analog scale [0.4, 25.2\%]). The overall quality of the evidence was very low. No adverse events were reported in any of the included studies.

Conclusions: The use of therapeutic US may provide additional benefits to physical therapy regimens in terms of symptom relief in individuals with KOA. However, it is not possible to make any meaningful recommendations for clinical practice due to the small number of applicable RCTs and the low methodological quality of the RCTs deemed eligible for this study.

KEYWORDS
Analgesia;
Chronic pain;
Clinical applications;
Electrophysical agent;
Physical therapy

* Corresponding author at: Center for Complementary and Integrative Medicine, Division of Rheumatology, Tufts Medical Center, 800 Washington St., #406, Boston, MA 02111, USA.
E-mail: rbannuru@tuftsmedicalex.org (R.R. Bannuru).

https://doi.org/10.1016/j.bjpt.2021.07.003
1413-3555/© 2021 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier España, S.L.U. All rights reserved.
Introduction

Knee osteoarthritis (KOA) affects approximately 250 million people worldwide and has a high societal and economic burden. Because there is no known cure or spontaneous remission of the disease, therapeutic regimens must often address chronic symptom management. However, the long-term use of conventional pharmacologic treatments, particularly opioids, has raised concerns due to the potential for serious adverse side-effects and misuse. Current clinical practice guidelines recommend a combination of non-pharmacological and pharmacological treatments to manage KOA symptoms and have increasingly emphasized reducing the use of pharmacologic analgesics. Despite the evidence in support of non-pharmacologic treatment options, a recent study suggested that non-pharmacologic therapies are still underused in the practical management of knee and hip osteoarthritis (OA). One of the key barriers to uptake of non-pharmacologic treatment modalities, including therapeutic ultrasound (US), is a perceived lack of evidence in support of these treatments on the part of providers.

Therapeutic US is used as a complementary treatment in physical therapy regimens focused on managing pain and aiding in the healing of soft tissue injuries. The treatment exerts therapeutic effects through thermal (continuous US) and non-thermal (pulsed US) modalities via a variety of application parameters (i.e., intensity, wavelength, duty cycle, and frequency). Continuous US achieves the thermal effect and is purported to produce analgesia through temperature elevation, which increases capillary permeability and tissue metabolism, thereby enhancing fibrous tissue extensibility and pain thresholds. Non-thermal effects are achieved by modulating cell membrane permeability, increasing protein synthesis, and activating immune response near the injury site, which may stimulate regeneration of damaged tissue.

High-quality evidence is needed to determine the efficacy of therapeutic US. Previous systematic reviews regarding US effectiveness on KOA are outdated and the latest reviews presented methodological limitations, such as the inclusion of mixed interventions, that hindered the evidence synthesis and the inclusion of pulsed US (low-intensity) only. We aimed to perform a more focused and comprehensive evidence synthesis targeting the isolated effects of therapeutic US to more clearly define its contributing role as an adjuvant treatment in rehabilitative regimens for KOA. Therefore, we conducted a systematic review and meta-analysis assessing the effects of therapeutic US compared with sham ultrasound on pain, function, and adverse events in individuals with KOA.

Methods

Data sources and searches

This study was conducted in accordance with PRISMA guidelines. We searched PubMed, MEDLINE, EMBASE, Google Scholar, and the Cochrane databases from inception through May 2019. We manually searched the reference lists of the most recent systematic reviews and meta-analyses and reviewed the supplements of conference proceedings published until May 2019. Our search terms included randomized controlled trials (RCTs) involving adults with symptomatic knee and/or hip OA that compared therapeutic US against a sham or other control. We incorporated hip OA into our search criteria to ensure that trials involving mixed populations of patients with knee and hip OA that provided separate data for knee OA participants were not omitted from our study. We did not register our study protocol in the PROSPERO database because a significant portion of the data acquisition and quality assessment had been conducted prior to the initiation of the present work, in the context of another project that was completed by two of the investigators (MO, RB). The full search strategy is available in the Supplemental Online Material.

Study selection

A web-based screening platform (http://rheumatology.tuftsmedicalcenter.org/CTCIA) was used to conduct abstract and full text screening of the references gathered from the literature search. Based on the PICO framework, we sought the following essential elements within each reference throughout the screening process: Population of participants with knee OA, Intervention of therapeutic US, Comparator of sham US (or inert comparison), and Outcome(s) of pain, function, and/or adverse events. Mixed interventions (e.g., US combined with exercise, hot packs, etc.) and mixed populations (e.g. patients with knee and hip OA, without separate results for knee OA only) were excluded. Though mixed interventions may be common in practice, the allowance of inconsistent concomitant therapies among included studies may result in clinical heterogeneity that may in turn increase the heterogeneity of the effect estimates produced by a meta-analysis. Studies that did not report outcomes of interest were also excluded. Two reviewers (LD, MO) independently assessed the title and abstract of each reference to determine potential eligibility during the abstract screening stage. Articles included during this stage were deemed eligible for full text screening. During the full text screening stage, full manuscripts for each abstract were obtained and examined thoroughly by the same independent reviewers (LD, MO). After abstract and full text screening concluded, respectively, discrepancies in inclusion were adjudicated by a third investigator (RB).

Data extraction and quality assessment

Quality was assessed at the study level using the Cochrane Risk of Bias tool. Data on outcomes of interest were extracted by a researcher (MO) using RevMan software. Outcomes of interest included the following: subjective measures of pain and functional status, measured by any validated scale; objective measures of function, specifically,
the 6 m walk test, timed get up and go test, and/or timed walking distance; quality of life, measured by any validated scale; structural changes, including joint space narrowing and cartilage volume change; discontinuation due to adverse events; incidence of treatment-related adverse events, and serious adverse events. Incidence of one measure of subjective pain or function was reported, the WOMAC scales were prioritized for our analyses; and the component summary scores of the Short Form-36 Item Medical Outcome Survey (SF-36) were the prioritized reporting method for quality of life.

Quality was assessed at the outcome level using GRADE methodology. GRADE Evidence Profiles were constructed and Evidence Tables were produced by exporting the results of all analyses from RevMan into GRADEpro web-based software; evidence tables were produced for both full analysis sets and for the sensitivity analyses limited by study quality. GRADE assesses quality according to four principal domains: risk of bias, inconsistency, indirectness, and imprecision. The overall classification of the evidence is assessed as either “high,” “moderate,” “low,” or “very low.” We established and adhered to strict inclusion criteria to circumvent reassessment of the indirectness domain. For risk of bias, we established a priori criteria based on study-level ratings: a “very low” quality rating required that a study received ≥2 high risk of bias ratings or 1 specific high risk rating in the “Other” category in addition to ≥2 unclear ratings or ≥3 unclear ratings in dimensions other than the “Other” category using the Cochrane Risk of Bias tool; a “low” quality rating required that the study received at least 1 high risk rating; a “moderate” quality rating required that the study received 2 unclear risk of bias ratings, in categories related to Randomization and/or Blinding, and a “High” quality rating required that the study received 1 or fewer unclear risk of bias ratings. We used the following a priori I² cutoff values for inconsistency ratings: <50% = low/acceptable heterogeneity, no quality downgrading; ≥50% and ≤75% = moderate heterogeneity, “serious” quality downgrading; >75% = high heterogeneity, “very serious” quality downgrading. In grading the imprecision of continuous effects, effect size magnitude cutoffs proposed by Cohen were used as benchmarks in assessing the magnitude of the confidence intervals of observed effect estimates: a 95% confidence interval (CI) of an standardized mean difference (SMD) extending between >0.2 and <0.5 points in either direction, “serious” quality downgrading; 95% CI of an SMD exceeding ≥0.5 points in either direction, “very serious” quality downgrading. We closely followed GRADE imprecision guidelines to assess risk ratios, whereby a 1-point downgrading is applied if the 95% CI of the risk ratio crosses the null value. Additional downgrades for imprecision were applied for very small sample sizes in pooled analyses: a “serious” quality downgrade was used to classify sample sizes in one’s study arm of <50 individuals, and a “very serious” quality downgrade was used to classify total sample sizes ≤30 individuals. Because fewer than 10 published studies were gathered for outcomes of interest, we did not objectively assess publication bias using Egger’s test (Cochrane Handbook, 10.4.3.1).23,24 GRADE quality assessments were undertaken by two independent reviewers (MO, RB), who resolved conflicts through discussion and consensus.

Data synthesis and analysis

In analyzing continuous data, we calculated SMDs and 95% CIs. To account for clinical and methodological heterogeneity, we conducted meta-analyses using random effects models using the DerSimonian and Laird inverse variance method.25 To aid in interpretation of SMDs, we used the approach from Bliddal and Christensen, to convert SMDs to percentage improvement on a 0–100 visual analogue scale (VAS).26 Dichotomous outcomes were analyzed using the Mantel-Haenszel method, and results were reported as risk ratios with 95% CIs.27 We assessed between-trial variance using Tau squared, and quantified heterogeneity using the I² statistic.28,29 In the event of high levels of heterogeneity (I² ≥75%), and/or between-trial variability, we conducted analyses excluding any visibly apparent outliers (95% CIs not overlapping with a majority of included studies). All meta-analyses were conducted using RevMan software.20 A third investigator checked over all data extraction, analyses, and study quality ratings to ensure consistency and accuracy (RB).

To minimize the potential for bias, our primary efficacy analyses involved only RCTs utilizing adequate sham comparators, but we planned supplemental analyses including other controls a priori. All included RCTs contributed to analyses of safety, regardless of control.

Results

Study characteristics and quality

Of 265 potentially relevant references procured by our systematic search, only 5 RCTs30–34 (N = 264 participants) were eligible for inclusion in our analyses (Fig. 1). The most common reason for exclusion was the incorporation of at least one concomitant physical therapy intervention with therapeutic US. Included RCTs were published between 2009 and 2017, and the duration of the studies ranged from two to eight weeks. Characteristics of the included studies are described in Table 1. The mean age range of participants

---

**Fig. 1** Study flow diagram.
<table>
<thead>
<tr>
<th>Author, year, country</th>
<th>N of patients</th>
<th>Age, mean (years)</th>
<th>% Female</th>
<th>Intervention(s) vs. Comparator</th>
<th>Treatment Frequency</th>
<th>Application location</th>
<th>Rescue analgesia</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Külcü, 2009, Turkey</td>
<td>45</td>
<td>63.5</td>
<td>78%</td>
<td>Therapeutic Ultrasound (1 MHz, 1.5 W/cm² continuously, 10 min/session) vs. No Treatment</td>
<td>5 sessions per week for 3 weeks</td>
<td>Superomedial and lateral parts of the target knee</td>
<td>The control group was allowed to receive acetaminophen as needed during the study</td>
<td>3 weeks</td>
</tr>
<tr>
<td>Özgönenel, 2009, Turkey</td>
<td>67</td>
<td>54.9</td>
<td>81%</td>
<td>Therapeutic Ultrasound (either 1 MHz or 1 W/cm² continuously, 5 min/session) vs. Sham Ultrasound</td>
<td>5 times per week for 2 weeks</td>
<td>Lateral and medial margins, avoiding the patella of the target knee</td>
<td>None</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Tascioglu, 2010, Turkey</td>
<td>82</td>
<td>60.7</td>
<td>66%</td>
<td>Therapeutic Ultrasound (1 MHz, 2 W/cm² continuously, 5 min/session) vs. Pulsed Therapeutic Ultrasound (1 MHz, 2 W/cm² and pulsed mode duty cycle of 1:4 for 5 min) vs. Sham Ultrasound</td>
<td>5 times per week for 2 weeks</td>
<td>Superomedial and lateral parts of the target knee</td>
<td>None</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Loyola-Sánchez, 2012, Canada</td>
<td>27</td>
<td>61.9</td>
<td>78%</td>
<td>Pulsed Therapeutic Ultrasound (1 MHz, 0.2–1 W/cm², 9.5 min/session) vs. Sham Ultrasound</td>
<td>3 sessions per week for 8 weeks</td>
<td>Tibiofemoral joint medial to the patellar tendon of the more painful knee</td>
<td>NR</td>
<td>8 weeks</td>
</tr>
<tr>
<td>Yegün, 2017, Turkey</td>
<td>62</td>
<td>ND</td>
<td>82%</td>
<td>Therapeutic Ultrasound (1 MHz, 1 W/cm² continuously, 8 min/knee) vs. Sham Ultrasound</td>
<td>One treatment session</td>
<td>Superomedial and lateral parts of both knees</td>
<td>NR</td>
<td>4 weeks</td>
</tr>
</tbody>
</table>

NR = not reported.
<table>
<thead>
<tr>
<th>No. of studies</th>
<th>Study design</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>Therapeutic Ultrasound</th>
<th>Sham</th>
<th>Relative (95% CI)</th>
<th>Absolute (95% CI)</th>
<th>Certainty</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>not serious</td>
<td>serious</td>
<td>none</td>
<td>N = 131</td>
<td>N = 103</td>
<td>SMD 0.33 lower (0.6 lower to 0.07 lower)</td>
<td>@@@ LOW</td>
<td>CRITICAL</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>not serious</td>
<td>serious</td>
<td>none</td>
<td>N = 76</td>
<td>N = 76</td>
<td>SMD 0.33 lower (0.65 lower to 0.01 lower)</td>
<td>@@@ LOW</td>
<td>CRITICAL</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>not serious</td>
<td>not serious</td>
<td>none</td>
<td>N = 42</td>
<td>N = 45</td>
<td>AD 0.10 higher (0.04 higher to 0.16 higher)</td>
<td>@@@@ MODERATE</td>
<td>IMPORTANT</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>very serious</td>
<td>serious</td>
<td>none</td>
<td>N = 89</td>
<td>N = 58</td>
<td>SMD 0.36 lower (0.98 lower to 0.27 higher)</td>
<td>@@@@ VERY LOW</td>
<td>IMPORTANT</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>not serious</td>
<td>serious</td>
<td>none</td>
<td>N = 30</td>
<td>N = 32</td>
<td>SMD 0.02 lower (0.51 lower to 0.48 higher)</td>
<td>@@@ LOW</td>
<td>IMPORTANT</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>very serious</td>
<td>serious</td>
<td>none</td>
<td>N = 11</td>
<td>N = 12</td>
<td>SMD 0.11 lower (0.93 lower to 0.71 higher)</td>
<td>@@@@ VERY LOW</td>
<td>IMPORTANT</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>not serious</td>
<td>not serious</td>
<td>none</td>
<td>0/118 (0.00)</td>
<td>0/88 (0.00)</td>
<td>not pooled</td>
<td>see comment</td>
<td>@@@@ MODERATE</td>
<td>IMPORTANT</td>
</tr>
<tr>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>not serious</td>
<td>not serious</td>
<td>not serious</td>
<td>none</td>
<td>0/118 (0.00)</td>
<td>0/88 (0.00)</td>
<td>not pooled</td>
<td>see comment</td>
<td>@@@@ MODERATE</td>
<td>IMPORTANT</td>
</tr>
</tbody>
</table>

**OVERALL QUALITY OF EVIDENCE:** @@@ LOW.


**Explanation:**
- 50% of trials received "High" risk of bias ratings (≥ 1 out of 6 dimensions in the Cochrane Risk of Bias tool)
- 95% CI of an SMD extends between > 0.2 and ≤ 0.5 points in either direction (Cohen 1988*).
- 95% CI of an SMD extends > 0.5 points in either direction (Cohen 1988*).
- Total sample size < 30.
- Due to zero events in both study groups, a relative risk could not be estimated.
was 54 to 64 years old; and one study did not report age. Three RCTs assessed the efficacy of continuous therapeutic US only, and one assessed pulsed therapeutic US only; and one RCT was a three-arm trial involving both continuous and pulsed US. Four of the five studies involved a sham comparator and were included in our primary efficacy analyses; one study compared therapeutic US to no treatment. The use of pharmacologic rescue medication (in the event of insufficient pain relief from the study treatments) was not allowed or not reported in the majority of the studies. One study allowed the use of acetaminophen in the control group.

The methodological quality of the included RCTs ranged from moderate to very low (Supplemental Online Material). The one very low quality RCT was also the only trial that did not involve a sham comparator. All of the studies randomized fewer than 100 participants in total. Quality assessed at the outcome level using GRADE methodology is presented in Table 2. The overall quality of the evidence was very low by GRADE Criteria.

**Dosage of therapeutic ultrasound**

The parameters of therapeutic US, as well as the frequency of its application, varied considerably among the included studies. For the continuous form, the minimum application period was one session and the maximum was 15 sessions. The frequency used was 1 MHz, with the intensity ranging from 1 to 2 W/cm². The therapy application time ranged from 5 to 10 min per session. For the pulsed form, the minimum application period was 10 times within two weeks, and the maximum was 24 sessions within 3 weeks. The frequency used was 1 MHz, with the intensity ranging from 0.2 to 2 W/cm². The therapy application time ranged from 5 to 9.5 min per session.

**Efficacy of therapeutic ultrasound**

All four sham-controlled RCTs reported pain outcomes within 2 to 8 weeks of follow-up (N = 234). Therapeutic US demonstrated small, but statistically significant benefits on pain (Table 3). Supplemental analyses including the trial with a “no treatment” comparator resulted in slightly larger magnitude of the effect. Outcome-level quality for pain assessed by GRADE was “low” due to imprecision of the estimate and Risk of Bias concerns.

**Safety of therapeutic ultrasound**

Four RCTs reported safety outcomes, including discontinuation due to adverse events and treatment-related adverse events. No treatment-related adverse events or discontinuations due to adverse events were reported by participants receiving therapeutic US or sham/no treatment throughout the duration of any of the studies. Incidence of serious adverse events was not noted or reported.

**Discussion**

The results of this meta-analysis demonstrated small, but statistically significant benefits of the use of therapeutic US versus sham in relieving pain and improving self-reported physical function in individuals with KOA. No adverse events were observed over the course of the follow-up of any included study. However, due to the low level of methodological quality and the limitation in the sample sizes of the

---

**Table 3** Efficacy of therapeutic ultrasound versus sham in knee osteoarthritis.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N of RCTs</th>
<th>N of patients</th>
<th>Effect estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain (2–8 weeks of follow-up)</td>
<td>4</td>
<td>234</td>
<td>SMD -0.33 (-0.60, -0.07)</td>
</tr>
<tr>
<td>Pain - % Improvement (0–100 VAS)</td>
<td>4</td>
<td>234</td>
<td>9.6% (2%, 17.4%)</td>
</tr>
<tr>
<td>Pain [Supplemental Analysis] (2–8 weeks of follow-up)</td>
<td>5</td>
<td>264</td>
<td>SMD -0.45 (-0.77, -0.13)</td>
</tr>
<tr>
<td>Function (2–8 weeks of follow-up)</td>
<td>3</td>
<td>152</td>
<td>SMD -0.33 (-0.65, -0.01)</td>
</tr>
<tr>
<td>Function - % Improvement (0–100 VAS)</td>
<td>3</td>
<td>152</td>
<td>12.8% (0.4%, 25.2%)</td>
</tr>
<tr>
<td>Function [Supplemental Analysis] (2–8 weeks of follow-up)</td>
<td>4</td>
<td>182</td>
<td>SMD -0.45 (-0.80, -0.09)</td>
</tr>
<tr>
<td>Function - % Improvement (0–100 VAS)</td>
<td>4</td>
<td>182</td>
<td>17.4% (3.5%, 31%)</td>
</tr>
<tr>
<td>6-minute walk test (4–8 weeks of follow-up)</td>
<td>2</td>
<td>87</td>
<td>SMD 0.10 (0.04, 0.16)</td>
</tr>
</tbody>
</table>

CI, confidence interval; MD, mean difference; N, number; RCT, randomized controlled trial; SMD, standardized mean difference; VAS, visual analogue scale;

Bold letters indicate statistically significant results or p values lower than 0.05.

---

The therapy application intensity ranging from 0.2 to 2 W/cm². The therapy application time ranged from 5 to 9.5 min per session.
included studies, there is still not adequate evidence to recommend therapeutic US as an effective treatment for patients with KOA.

**Interpretation of the main findings**

Pain reduction induced by therapeutic US may be due to both thermal and non-thermal mechanisms, which can be adjusted to warm superficial soft tissues or to accelerate tissue healing at the cellular level. There is rich literature suggesting that OA has an inflammatory component and that painful symptoms in OA may be the result of activation of inflammatory pathways that cause an increased response of peripheral joint nociceptors. It is possible that therapeutic US could mitigate inflammatory events by potentializing the repair phase, thus enhancing the overall healing process. In addition, a recent systematic review demonstrated positive benefits of low-intensity pulsed US on some properties of cartilage formation, especially in the increase of type II collagen, which is the principal fiber structure of articular cartilage, and in the reduction of a specific type of metalloproteinase expression, an inflammatory mediator that predisposes to OA. However, a major limitation was that the majority of the analyzed studies were done with animals and therefore, the findings cannot be fully translated to humans.

Regarding physical function, our results showed a beneficial effect for self-reported measures, but not for performance-based measures. Though valid and reliable patient-reported outcome measures are widely used in KOA, patients tend to over- or under-estimate their abilities depending on their perceptions, which may be influenced by symptoms or other demographic factors. In this meta-analysis, all of the included studies used the WOMAC questionnaire as the self-reported physical function measure. The WOMAC questionnaire is more influenced by the amount of pain experienced by individuals with KOA when compared with performance-based measures. Thus, because our results suggest that therapeutic US was significantly more effective than sham for pain relief, participants may have overestimated their subjective physical function.

The two most recent reviews assessing the efficacy of therapeutic US in OA concluded that US may be effective in improving symptoms in individuals with KOA. However, it is important to highlight that the composition of their contributing analyses and resulting conclusions are not...
consistent with our own, primarily due to methodological differences. The first review included 9 studies comparing therapeutic US with sham, but in many of these studies, US was paired with other therapies such as muscle strengthening and stretching exercises, manual therapy, and intra-articular hyaluronic acid injection. Moreover, in four of the included studies, both groups (US and control) were using concomitant therapies, such as hot packs, warm-up exercises, and passive range of motion exercises. The second review, which evaluated only the effects of low-intensity pulsed US on KOA, also included two studies involving additional therapies in the intervention and control groups, including a home-based exercise program and diclofenac, respectively. The incorporation of additional treatments within studies and variation in types of additive therapies between studies could have added heterogeneity into the analyses of these studies, because neither the intervention nor the sham groups were compared consistently.

The recommendations of current clinical guidelines are in agreement with our results regarding US for KOA. The American College of Rheumatology, the National Institute for Health and Care Excellence, the Ottawa Panel, and the European League Against Rheumatism, either did not achieve expert panel consensus with respect to a recommendation for US or did not consider it as a potential therapy. In the current OA guidelines published by the Osteoarthritis Research Society International, an uncertain recommendation was issued for the use of the therapy as a complementary treatment option for people with KOA.

We attempted to improve upon previous methodologies by updating the search with stricter selection criteria and by restricting our analyses to only RCTs that administered therapeutic US as monotherapy compared to a control. We also updated the evidence base with a recently published trial that has not been included in previous syntheses.

Main limitations

This meta-analysis has some limitations. Both the small number of studies included and the small sample sizes hindered the evidence synthesis. Because all of the studies randomized less than 100 participants, our results may be subject to small study effects. Meta-epidemiologic research in OA trials has suggested that small study effects are further amplified in complementary medical interventions. Therefore, the magnitude and significance of the effects must be interpreted with caution until more robust research contributes to the body of evidence. The fact that data extraction was executed by a single reviewer based on resource and time restraints is another key limitation of our study, as this may have introduced bias and/or error into the estimates. However, given that GRADE assessments were conducted in tandem by two investigators, it is likely that any anomalies in bias assessments and/or effect size estimation would have been discovered during the thorough review of the RevMan file necessitated by each outcome-level assessment.

The variety found in the parameters of therapeutic US, as well as the frequency of its application may have introduced heterogeneity into our analysis, and ultimately these differences complicate the interpretation of our results. Interestingly, the studies with the lowest and highest application frequencies were the ones that demonstrated the least effect on pain and functional outcomes. However, due to the small number of RCTs assessing each dosing regimen, it was not reasonable to perform subgroup analyses based on dose, nor is it possible to make further assumptions regarding the potential dose effects of therapeutic US. This is a topic for further research, because the dose effect of therapeutic US is debated in the literature.

The follow-up periods of the studies also varied and were relatively short, ranging from 2 to 8 weeks. Because OA is a chronic condition, longer term studies, perhaps involving cyclical therapeutic sessions, would be required to determine the value of therapeutic US as part of an ongoing OA treatment regimen. Our study was also limited by the quality of the available data. Most of the studies were assessed to be of moderate to low quality. We attempted to mitigate the potential for performance bias and/or detection bias by limiting our primary analyses to RCTs that involved an adequate sham comparator. Finally, our search strategy may have been limited by the omission of content-specific databases such as PEDro, SPORTDiscus, or CINAHL.

Implications for clinical practice

Clinicians and physical therapists have expressed interest in determining the effectiveness of therapeutic US in managing KOA, based on its potential to reduce pain, improve function, and stimulate cartilage repair, and due to the low risk of adverse events associated with the intervention. However, as our systematic review showed, there is still a lack of large, high quality studies that evaluate the efficacy of the therapy in individuals with KOA. Therefore, even though US is currently used in clinical settings worldwide, the body of evidence is not yet strong enough to support US, administered in isolation, as an effective treatment for patients with KOA. The lack of high-quality evidence may also contribute to a lack of uptake of the intervention by providers in some clinical settings.

Scientific recommendations

Future investigators should aim to follow gold standard methods of designing and reporting trials, such as the Consolidated Standards of Reporting Trials (CONSORT) Statement for Randomized Trials of Nonpharmacologic Treatments. Moreover, researchers should consider using the Template for Intervention, Description, and Replication (TIDier) checklist and guide, which improve the completeness of reporting, and ultimately, the replicability of interventions.

Conclusion

The results of this meta-analysis suggest that therapeutic US is a safe non-pharmacological treatment option that may provide additional pain relief and functional improvement when used as an adjuvant therapy in individuals with KOA. However, the current body of evidence is not adequate in size or quality to make any meaningful recommendations for clinical practice. Future research should consist of larger, higher quality RCTs to determine the efficacy of therapeutic US in individuals with KOA.
Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgments

L.O. Dantas, M.C. Osani, R.R. Bannuru contributed to the conception and design of this study. M.C. Osani and R.R. Bannuru contributed to the analyses of the data. All authors contributed to the interpretation of the data. Article drafts were written by L.O. Dantas and M.C. Osani and critically revised by all authors. The final version of the article was approved by all authors. L.O. Dantas takes responsibility for the integrity of the work as a whole.

L.O. Dantas was financially supported by Sao Paulo Research Foundation (FAPESP, Process number #2018/23705-3). Dr. Bannuru was supported by the National Center for Complementary and Integrative Health of the National Institutes of Health (K23AT009374). The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the NCCIH.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjpt.2021.07.003.

References

29. Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC
Chae KJ, Choi MJ, Kim KY, Ajayi FF, Chang IS, Kim IS.


