Purpose
Kinesiology taping is a common intervention used to treat individuals with shoulder pain. While there have been several studies published to date evaluating the effectiveness of this intervention, a systematic review with meta-analysis synthesizing the collective effectiveness of kinesiology taping is not available. Therefore, the purpose of this study was to perform a systematic review with meta-analysis investigating the efficacy and treatment fidelity of kinesiology taping (KT) in combination with conservative interventions for shoulder pain.

Methods
Databases (PubMed, EMBASE, SportDiscus, CINAHL) of studies published in English meeting criteria were searched to October 2019. Methodologic quality was assessed utilizing the Modified Downs and Black checklist. Treatment fidelity was evaluated using a modified fidelity tool. Random effects meta-analyses were performed when an outcome (disability, pain, range of motion (ROM)) was reported by two or more studies. Overall effect size (pooled random effects) was estimated for studies with acceptable clinical homogeneity.

Results
When KT was used with conservative treatments, meta-analysis revealed large effect sizes for improvements in disability (standard mean difference (SMD) = -1.35; 95% CI, -2.09 to -0.60) and ROM (SMD = 0.96; 95% CI, 0.60-1.33) with no significant effects for pain. The average Modified Downs & Black score for bias was 11.5 ± 3.9. Of 10 retained studies, only two had good treatment fidelity.

Conclusions
Adding KT to interventions performed in clinical settings appears to demonstrate efficacy regarding disability and ROM when compared to conservative interventions alone. However, despite reasonably good methodologic quality, fidelity was lacking in a majority of studies. Because of its impact on the implementation of evidence-based practice, lower fidelity should be considered when interpreting results.
INTRODUCTION

Shoulder pain affects up to 67% of the adult population throughout their lifespan.\textsuperscript{1,2} The etiology of shoulder pain is multifactorial and inclusive of various impairments, including but not limited to restricted range of motion (ROM), pain, decreased strength, and impaired motor coordination.\textsuperscript{3–5} A variety of interventions are utilized in a multimodal approach for treating shoulder pain and dysfunction with the primary focus on restoring pain-free shoulder mobility and mitigating ensuing activity restrictions.\textsuperscript{5–7} In recent years, kinesiology taping (KT) procedures have become increasingly popular. The proposed benefits of KT include providing a tactile cue via the skin to assist or limit movements, modulate pain, and increase microcirculatory flow which is thought to assist in reducing inflammation.\textsuperscript{8–12} Kinesiology tape refers to various types of elastic adhesive materials applied to joints in order to achieve various therapeutic effects.\textsuperscript{13,14} In particular, KT is widely utilized to treat impairments of the shoulder due to ease of accessibility and the ability to adapt to various body morphologies.

Current evidence investigating the effects of KT on shoulder movement and overall function among both asymptomatic and symptomatic individuals has been contradictory. Some studies have concluded that KT assists in improving shoulder joint mechanics and functional movement, while others suggest KT has no impact on joint position sense and in some cases can even negatively impair joint position sense.\textsuperscript{9,14–18} Furthermore, KT may impair muscular performance and not be well-tolerated by some individuals.\textsuperscript{12} Results from other studies have demonstrated that KT has no significant difference when compared to placebo treatment, manual therapy, or different types of tape.\textsuperscript{12,13} This contradictory evidence makes it difficult to determine the clinical utility of KT for those individuals with shoulder disorders seeking conservative treatment options. Additionally, the treatment fidelity associated with studies investigating the efficacy of KT has not been investigated but is a requisite for the implementation of evidence-based clinical practice. Treatment fidelity is utilized in order to ensure that the particular intervention being studied is carried as described in the original protocol of an investigation. Without treatment fidelity it is uncertain as to whether those providing the intervention did so in the same manner each time and also makes it difficult to recreate similar investigations in the future.

Currently, systematic reviews have attempted to synthesize the literature surrounding KT. However, these systematic reviews either focus on musculoskeletal conditions as a whole or a specific shoulder pathology such as neurological conditions and rotator cuff tendinopathy.\textsuperscript{18–20} A recent systematic review sought to investigate the impact of KT in combination with a therapeutic intervention but focused solely on subacromial pain syndrome (SAPS) and retained only 4 studies which did not allow for any meaningful conclusion to be drawn.\textsuperscript{21}

Given that the combination of applying KT in addition to traditional physical therapy intervention closely mirrors how KT is utilized in a majority of clinic settings, it would appear most relevant to investigate this relationship if our analysis is to offer clinical utility.\textsuperscript{22–25} Thus, the purpose of this study was to perform a systematic review with meta-analysis investigating the efficacy and treatment fidelity of KT in combination with conservative interventions for shoulder pain. Multi-modal interventions were considered for inclusion as a means of supporting external validity and clinical impact.

METHODS

GUIDELINES

This systematic review used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines during the search and reporting phase of the research process. The PRISMA 2009 statement includes a 27-item checklist developed to improve reporting of systematic reviews and meta-analyses.\textsuperscript{26}

LITERATURE SEARCH

An online literature search was conducted utilizing PubMed, EMBASE, SportDiscus and CINAHL from the dates of their origin until October 2019. The search strategy was created and performed by a biomedical librarian. An example of the search strategy used for the PubMed database can be seen in Supplement 1 and similar strategies were utilized for the three remaining databases. This study was registered using the International Prospective Register of Systematic Reviews (PROSPERO) in July 2017 with the corresponding reference number CRD42017074147.

STUDY SELECTION

Appraisal of all titles was performed independently by two authors (CC) and (HC) after the initial online literature search. The abstracts of the retained titles were then reviewed to determine if inclusion criteria were met. Full-text versions of those studies whose abstracts met the inclusion criteria were accessed and read to determine their eligibility for the review. The same two authors undertook the study selection process throughout with discrepancies being decided by an independent author if necessary (no discrepancies were present). The inclusion criteria consisted of: (a) studies where KT was applied to the shoulder complex only; (b) studies including subjects diagnosed with shoulder pathology; (c) studies reporting objective measures of pain and function; and (d) studies comparing KT in conjunction with conservative treatment interventions. In addition, all study designs (including case studies) were included without restrictions on publication date or age of subjects, or duration of symptoms. Studies were excluded if (a) they were not published in the English language or (b) they had a mixed patient population (individuals with other orthopedic and neurological conditions).

DATA EXTRACTION

The data and results from each study selected for review was extracted using a standardized Population, Treatment, Comparison, Outcome and Setting (PICO) format.\textsuperscript{27} This format includes the characteristics of the population studied, treatments performed, comparative treatments, pri-
METHODOLOGICAL QUALITY TOOL

The Modified Downs and Black checklist was used to evaluate methodological quality within the individual studies. The Downs and Black checklist in its original format contains 26 items and has been shown to be a valid and reliable measure of methodological quality for randomized and nonrandomized studies. Given that various study designs were present among those articles retained for review, the authors chose to utilize a modified version of the Downs and Black checklist. The Modified Downs and Black checklist has been widely utilized to assess methodological quality. Six sections are present as part of the modified checklist and include patient selection bias (items 1–4), comparison (item 5), outcomes (items 6–8), reporting findings/statistical analysis (items 9–11), confounding (12 & 15), and power (items 14 & 15). Each item within the Downs and Black tool corresponds to a question and answered either yes (Y), no (N), or unable to determine (U). The maximum score for the checklist is 16 with all individual items rated as either yes (= 1) or no/unable to determine (= 0) except for item 12 that may be rated as yes (= 2), partially (= 1), or no/unable to determine (= 0). Two authors (PS) and (CC) independently scored the retained articles for methodological quality. Discrepancies were handled through discussion to reach a consensus.

TREATMENT FIDELITY ASSESSMENT

Treatment fidelity assessments were performed to determine if the studies included in this review followed appropriate procedures to ensure that valid comparisons of replicable interventions were being executed. We chose to assess treatment fidelity by utilizing a modified version of a tool developed by Borelli et al. The original scale was appropriate for psychological interventions, which contained domains that are outside the scope of clinical interventions. A modified version of the tool was utilized to identify the 11 items that represented the essential items pertaining to a study investigating clinical interventions and not phycological interventions. Although this modified scale has not been validated, it has been utilized by others in similar reviews to examine treatment fidelity. Items are scored as 0 (not present), 1 (minimally described), or 2 (more than minimally described). Scores were recorded as both the summary (number present, maximum possible score 11) and total score (maximum possible 22) and were determined by a single author (PS), a researcher with a clinical background in musculoskeletal orthopaedics and experience in fidelity scoring. We identified item summaries > 6 and total scores ≥ 12 as studies that exhibited good treatment fidelity. These scores are based on prior investigations in this area and less likely to misrepresent the actual effect of the treatment provided in any given study.

DATA SYNTHESIS AND ANALYSIS

The mean and standard deviation in the outcome variables of pain, disability, and ROM preintervention and postintervention, as well as sample sizes for the experimental and control groups were extracted. The assessments administered at the end of the intervention or closest to one-month post intervention were used in the meta-analysis. Effect sizes based on standard mean differences (SMDs) were calculated for each outcome variable. When not provided, the standard deviation of the change score was computed from the standard deviations of the pre-post scores, using a correlation of 0.50. Studies were grouped by the comparison category (e.g., kinesiology taping plus exercise vs. exercise only). Random effects meta-analyses were performed for each subgroup of studies to examine an outcome variable when an outcome was reported by two or more studies. A pooled random-effects estimate of the overall effect size was estimated for all studies with acceptable clinical homogeneity. For pain and disability scales (the numeric pain rating scale (NPRS) and visual analog scale (VAS) for pain, Disability of Arm and Shoulder Questionnaire (DASH), and Shoulder Pain and Disability Index (SPADI)), higher scores indicate greater pain or disability. By contrast, higher scores reflect higher quality of shoulder function in the Constant score. Therefore, the direction of Constant score was reversed when pooled with other disability scales. For an increase in ROM, a positive effect size favors the experimental group over the control group. The effect size of the interventions was categorized into one of the three levels based on SMD (<0.40 = small, 0.40 to 0.70 = moderate, >0.70 = large).

The presence of heterogeneity was evaluated using the Q statistic after the meta-analysis. The Q statistic follows a chi-squared distribution with n-1 degrees of freedom. A small p-value (p<0.05) for the Q statistic indicates that heterogeneity is present, and the meta-analysis model has some unaccounted-for bias. All meta-analyses were conducted in Stata (version 12.1) statistical software (StatCorp, College Station, TX) and the metan command. All data analysis was performed by (XL), an academician with a specialty in biostatistics and health informatics.

RESULTS

SEARCH STRATEGY

Database searches yielded 618 independent study titles, resulting in 357 after duplicates were removed. After a title and abstract search, 345 studies were removed because they did not meet the a priori inclusion/exclusion criteria. After review of the full-text studies, 10 were ultimately retained for analysis (Figure 1). Study characteristics for each of the 10 included studies are included in the PICOS table (Appendix A).

METHODOLOGICAL QUALITY

Methodological quality assessment revealed a range of values (Table 1). Of a 15 total item maximum and 16 total possible points, the average score was 11.5 ± 3.9 (median 12,
Figure 1: PRISMA flow diagram outlining study selection process

range 2-16). Only one of the 10 studies retained for analysis contained elements for each item resulting in a score of 16.42 There was one study that demonstrated a low quality and received a total score of 2.41 The remaining studies received a total score of 10 or greater.22,23,25,43–47 Although there are no validated cut points for what constitutes a good, fair, or poor score for methodological quality as identified by the modified Downs and Black Checklist, most of the retained studies received a score greater than half of the total possible score. Item 7 (Were the main outcome measures used accurate (valid and reliable)? Accuracy not reported but method clearly described) was scored as yes in only two of the retained studies.42,46

TREATMENT FIDELITY

Treatment fidelity assessments produced a wide value range (Table 2). Out of a maximum of 11 items and 22 total points, the average items identified were 7 ± 1.9 (median 7.5, range 3–9) and the total score average was 9.4 ± 3.0 (median 10.5, range 3–13). No individual study had all of the items represented; items 3 (If more than one intervention is described, are they all described equally well?) and 6 (Characteristics to be sought and avoided by treatment provider are addressed a priori, make some mention of credentials) were identified in all studies and item 15 (Were non-specific treatment effects evaluated) was not identified in any of the studies. While no validated cut-off exists for this tool, using our criteria, two studies25,46 had item scores ≥ 6 and total scores ≥ 12; these studies were independently identified as good treatment fidelity.

META-ANALYSIS

Of ten studies retained in this systematic review,22,23,25,41–47 all but one41 provided sufficient data to be included and pooled in the meta-analysis. One study41 did not report standard deviation information for baseline and post-intervention assessments and was excluded from the meta-analysis. Figures 2-4 represent the meta-analyses for pain, disability, and ROM outcomes, respectively.

Pain meta-analysis. Seven studies25,42–47 investigating pain outcomes compared KT with exercise to other interventions. One study compared the addition of KT and exercise to non-steroidal anti-inflammatory drugs (NSAIDs)45 vs. a comparator treatment that did not include KT. One study42 utilized the NPRS while the remaining studies25,43–47 utilized the VAS. Five subgroups were analyzed depending on the intervention comparisons. The overall meta-analysis for comparing KT to no KT or sham KT for pain resulted in a small effect size in favor of the experimental group, but the difference was not statistically sig-
Significant (SMD = -0.16; 95% CI, -0.40 - 0.08). Significant heterogeneity was present among the included studies (heterogeneity statistic $Q = 42.92; df = 14; p < 0.001$), indicating that the variability in the SMD values were greater than that expected by sampling error alone (Figure 2).
Table 1: Modified Downs and Black Risk of Bias Assessment

<table>
<thead>
<tr>
<th>Authors</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devereaux et al 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>16</td>
</tr>
<tr>
<td>Djordjevic et al 2012</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>14</td>
</tr>
<tr>
<td>Gandhi et al 2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>Goksu et al 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>15</td>
</tr>
<tr>
<td>Kaya et al 2014</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>13</td>
</tr>
<tr>
<td>Onat et al 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>12</td>
</tr>
<tr>
<td>Pekyavas &amp; Baltaci 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>12</td>
</tr>
<tr>
<td>Simsek et al 2013</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>12</td>
</tr>
<tr>
<td>Subasi et al 2016</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>11</td>
</tr>
<tr>
<td>Tantawy &amp; Kamel 2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>10</td>
</tr>
</tbody>
</table>

1=y|es, 0=no/uns|ure, | item 12 scoring only; 2=y|es, 1=uns|ure, 0=no). Criteria 1-15, per the Modified Downes and Black Checklist.
The Efficacy and Treatment Fidelity of Kinesiology Taping in Conjunction With Conservative Treatment Interventions Among...

Table 2: Fidelity Scoring

<table>
<thead>
<tr>
<th>Author</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>Summary</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devereaux et al 2015</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Djordjevic et al 2012</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Gandhi et al 2016</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Goksu et al 2016</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Kaya et al 2014</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Onat et al 2016</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Pekyavas &amp; Baltaci 2016</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Subasi et al 2016</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Tantawy &amp; Kamel 2016</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>13</td>
</tr>
</tbody>
</table>

0 = not present; 1 = minimally described; 2 = more than minimally described
Key to Fidelity criteria:
1. Is there information about treatment dose in the intervention condition?
2. Is there information about treatment dose in the control/comparison condition?
3. If more than one intervention is described, are they all described equally well?
4. Method to ensure dose is equivalent between conditions
5. Method to ensure dose is equivalent within conditions
6. Characteristics to be sought and avoided by treatment provider are addressed a priori, make some mention of credentials
7. Is there mention of a theoretical model or clinical guidelines upon which the intervention is based?
8. Characteristics to be sought and avoided by treatment provider are addressed a priori, make some mention of credentials
9. Method to ensure that the dose of the intervention was being delivered as specified?
10. Is there a method to ensure that the content of the intervention was being delivered as specified?
11. Were non-specific treatment effects evaluated

Disability meta-analysis. Seven studies assessing disability were included that compared KT in addition to exercise (as well as manual therapy and high-intensity laser therapy in a single study) versus a comparator treatment that did not include KT. Two studies utilized the Constant score, four studies utilized the SPADI, and two studies utilized the DASH. The overall meta-analysis for disability and function resulted in a large and statistically significant effect size (SMD = -1.35; 95% CI, -2.09 to -0.60) in favor of the experimental group. The homogeneity statistic $Q= 303.18$ (df = 29, $p<0.001$) indicated significant heterogeneity among the included studies, and the variability in the SMD values were greater than that expected by sampling error alone (Figure 4).

DISCUSSION

The aim of this study was to perform a systematic review and meta-analysis of the current literature to determine the efficacy and treatment fidelity related to the application of KT on the shoulder complex in conjunction with conservative treatment interventions within a symptomatic population. Within this review, KT was paired with various interventions in order to best mirror how KT is currently utilized in the clinical setting and was compared to other standard practices seen in the treatment of patients with shoulder pathology. Essentially, clinical practice often dictates multi-modal approaches to care, thus adding KT to conservative care would offer improved clinical translation.

All but one study scored greater than or equal to 10/16 on the modified Downs and Black checklist. The majority of studies included in this systematic review described a...
patient population similar to that which is typically seen in the clinical setting with their characteristics clearly described, adding to the clinical utility of this review. All studies utilized a comparison group, which allowed comparison of kinesiology taping to commonly-used intervention strategies such as therapeutic exercise, manual therapy techniques, injections, and NSAIDs. These studies appeared to demonstrate clinical utility and ease of transference to practice; however, there appears to be lower methodological quality and questionable validity and reliability of outcomes measured. Only two studies clearly described the accuracy and reliability of the outcomes used; four studies clearly demonstrated appropriate adjustment for confounding variables; and only half of the studies had sufficient power to determine a clinically important effect. While the populations and treatments resemble those commonly seen in the clinic, the included studies may lack the internal validity needed to accurately assess the use of KT in place of interventions that are already supported by literature and currently utilized in practice.

Attention to the critical aspects of treatment fidelity underpins the validity of clinical research and is a prerequisite for the implementation of evidence-based clinical practice. Despite the importance of treatment fidelity, it is an often-neglected component of intervention. To our knowledge, this is the first study to evaluate treatment fidelity of clinical KT interventions. Only two studies were identified to have good treatment fidelity based on the criteria set forth. These results suggest the majority of studies included in this review could have been influenced by factors such as a lack of researcher training, lack of adherence to protocols, or interventions not performed as specified. Lack of treatment fidelity could also help to explain why results of studies investigating similar interventions may produce different results. This is important to note, especially when considering the use of KT in place of a more invasive or higher cost treatment, such as receiving injections or the prescription of NSAIDs. For example, Goksu et al. reported significant improvements after either injections or KT for shoulder impingement patients, although this study demonstrated a lack of treatment fidelity. In comparison, Subasi et al. had similar findings but demonstrated increased adherence to items associated with treatment fidelity; therefore, KT may be favored as a less invasive clinical treatment. Treatment fidelity should be considered when comparing outcomes of different studies utilizing similar interventions.

The meta-analysis revealed a large effect size for both disability and ROM favoring the experimental group and a
small, statistically insignificant effect size for pain in favor of the experimental group. Although the meta-analysis favors the experimental group for both disability and ROM, these findings must be considered in the presence of both the methodological quality and treatment fidelity of the included studies. Of those studies investigating the effect of KT and other conservative treatment on disability as well as ROM, only two studies\(^{25,46}\) had a treatment fidelity score considered to be "good." Furthermore, of the 10 studies retained for analysis, only four\(^{22,25,41,47}\) demonstrated a significant between-group improvement when compared to the paired intervention alone considering their chosen outcome measures. One of these studies\(^{41}\) had the lowest scores of both methodologic quality and treatment fidelity among those included.

The findings of these studies, even in the context of this review, cannot be interpreted alone without methodological quality and treatment fidelity being considered. When fundamental limitations in treatment fidelity exist such as those found among a majority of the retained studies; lack of reported clinician training, lack of adherence to protocols, and interventions not performed as specified, the validity of the findings may be questioned. Adhering to and appropriately reporting elements related to treatment fidelity are necessary to maintain the integrity of intervention-based research.

**LIMITATIONS**

There are limitations to this study, such as the inclusion of only those studies published in the English language. The authors also did not include a search of the grey literature as the content area being investigated with this type of search would not produce results beyond our current search given the search strategy initially implemented. Methodological quality was assessed utilizing the Modified Downs and Black Assessment Tool, which requires categorization of sub-elements within each study based on the assessor's determination, which could introduce bias. However, this was minimized by having two authors independently score each of the retained studies and discuss any discrepancies. The same limitation may also be considered regarding the assessment of treatment fidelity. Finally, the tool we utilized to evaluate treatment fidelity has not been validated in the modified form.

---

**Figure 3: Meta-analysis of kinesiology taping for disability, subgrouped by comparison type**

KT = kinesiology taping, EX = exercise, MT = manual therapy, HILT = high-intensity laser therapy, DASH = Disability of Arm and Shoulder Questionnaire (0 = no disability, 100 = most severe disability), Constant score (the higher the score, the higher the quality of the shoulder function), SPADI = Shoulder Pain and Disability Index (higher scores indicate greater pain and disability). NOTE: the direction of the Constant Score was reversed.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Outcome Variable</th>
<th>Exp. n</th>
<th>Exp. change</th>
<th>Exp. change SD n</th>
<th>Control n</th>
<th>Control change</th>
<th>Control change SD n</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KT + EX vs EX</td>
<td>Constant Score</td>
<td>33</td>
<td>-9.53</td>
<td>2.08</td>
<td>38</td>
<td>-8.29</td>
<td>2.01</td>
<td>-0.61 (-1.08, -0.13)</td>
</tr>
<tr>
<td>Peknyar  &amp; Baltaci Total SPADI score 20</td>
<td>-32.70</td>
<td>29.24</td>
<td>15</td>
<td>-6.01</td>
<td>17.94</td>
<td>-1.06 (-1.78, -0.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taranzi &amp; Karmel  Total SPADI score 33</td>
<td>-8.60</td>
<td>4.42</td>
<td>30</td>
<td>-2.00</td>
<td>4.89</td>
<td>-1.46 (-2.02, -0.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT + EX vs Injection + EX</td>
<td>Constant Score</td>
<td>30</td>
<td>-14.69</td>
<td>13.54</td>
<td>31</td>
<td>-21.33</td>
<td>13.94</td>
<td>0.48 (0.03, 0.99)</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT + EX vs MT + EX</td>
<td>Constant Score</td>
<td>16</td>
<td>-96.43</td>
<td>15.21</td>
<td>15</td>
<td>-6.01</td>
<td>17.94</td>
<td>-5.45 (-7.02, -3.88)</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT + EX vs sham KT</td>
<td>Constant Score</td>
<td>19</td>
<td>-20.42</td>
<td>14.45</td>
<td>19</td>
<td>-8.37</td>
<td>13.11</td>
<td>-0.87 (-1.54, -0.21)</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.35 (-2.09, -0.60)</td>
</tr>
</tbody>
</table>

*NOTE: Weights are from random effects analysis*
CONCLUSION

Although the addition of KT to conservative interventions may demonstrate some efficacy with regard to disability and ROM when compared to conservative interventions alone, these findings must be considered in light of the quality and treatment fidelity associated with these studies. A majority of the identified studies demonstrate fair methodologic quality; however, there were limitations surrounding treatment fidelity. Treatment fidelity is paramount to the implementation of evidence-based practice; thus, limited fidelity may in fact be of greater clinical value than quantitative changes when interpreting the result. Clinicians should consider these findings when determining if KT would be an appropriate adjunct in the treatment of those individuals presenting with shoulder pathology. Future studies on KT that adhere to the attributes of treatment fidelity may serve to reduce variability of results and provide a more practical outcome that can be replicated in the clinic.

ACKNOWLEDGEMENTS

Leila Ledbetter

CONFLICTS OF INTERESTS

The authors report no conflicts of interest.
REFERENCES


SUPPLEMENTARY MATERIALS

Supplement 1

Appendix A