Title: Clinical Practice Guidelines on Interventional Management of Low Back Pain: A Synthesis of Recommendations

Title Page

Abstract

Objective: To summarize the recommendations on the interventional management of subacute and chronic non-radicular low back pain (LBP) from the 21 quality-appraised CPGs identified in the previously published paper: “Quality of Clinical Practice Guidelines on Interventional Management of Low Back Pain: A Systematic Review”. By disseminating this information, we aim to facilitate the implementation of these recommendations into clinical practice.

Literature Survey: Electronic bibliographic databases, guideline databases and grey literature were searched from January 2016 to January 2020 to identify CPGs that met study criteria.

Methodology: 21 CPGs were quality-appraised and interventional management recommendations were extracted and organized into several treatment categories including epidural steroid injections (ESIs), radiofrequency procedures (RF), facet injections, sacroiliac injections (SI), and prolotherapy. Within each treatment category, the recommendations were organized based on 2 factors: quality of CPG and strength of recommendation.

Synthesis: Overall, there was no consistency in recommendations for or against any interventional procedure, even when accounting for the quality of the CPG. In all of the CPGs reviewed, the most common strength of recommendation was weakly-for. The second, third and

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmrj.12930

This article is protected by copyright. All rights reserved.
fourth most common strength of recommendations were inconclusive, weakly-against and strongly-against respectively and the least common was strongly-for. The treatment categories with the greatest number of recommendations were RF procedures (most common strength of recommendation-weakly for) and facet procedures. Among the high-quality CPGs, the most common strength of recommendation was inconclusive.

**Conclusions:** Most of the interventional management recommendations for management of non-radiculor LBP in the 21 CPGs appraised in this review were either weakly-for, weakly-against or inconclusive, with several recommendations within each treatment category contradicting each other. AGREE II quality appraisals of CPGs on interventional management of LBP were of unclear utility in guiding clinical implementation.

**Key Words:** Dissemination and Implementation Research, Health Care Delivery, Pain, Spine-Low back, Systematic Reviews
INTRODUCTION

Chronic low back pain (LBP) is a highly prevalent medical issue with the Global Burden of Disease Project estimating that the global prevalence of low back pain at 7.8% in 2017. Chronic LBP also has a high societal cost, with an associated 60.1 million disability-adjusted life-years globally in 2015, an increase of 54% since 1990. In 2010, LBP ranked first among medical conditions contributing to disability in the US. The direct financial cost of LBP on the healthcare system is also high. In 2016, the US spent an estimated $134.5 billion on healthcare for neck and back pain—an annual increase of 6.7% from 1996—a large portion of these costs going towards interventional treatment costs including spinal injections, and surgery.

Interventional treatment for LBP has been consistently increased seen over the last few decades. Between 1994 and 2001 there was a 271% increase in lumbar ESIs and a 231% increase in facet injections. Similarly, several other interventional procedures increased 228% from 2000 to 2011. This increase in interventional management of LBP has however not been uniform, with high variability in utilization between providers, the top 10% of interventionalists performing nine times more procedures per patient compared with the bottom 10% and 4.5 times more procedures when compared to the median provider.

High practice variability in interventional pain management is partly due to the vast and frequently conflicting amount of available evidence. Tools such as clinical practice guidelines (CPGs) condense available evidence and provide evidence-based recommendations that can aid interventionalists in practicing evidence-informed clinical care. The IOM defines CPGs as “statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefit and harms of alternative care options”. However, there are several CPGs addressing interventional management of LBP that provide inconsistent and sometimes conflicting recommendations, making adoption of their recommendations into clinical practice a challenging task. Therefore, a synthesis of recommendations made by CPGs, stratified by quality, could possibly be a valuable tool for clinicians considering interventional treatment options for LBP. Interventionalists’ adherence to CPG recommendations can also further optimize the quality of care they provide to patients with LBP while minimizing ineffective, unsafe, and other low-value treatments, reduce variations in practice, and reassure fellow clinicians and patients that they are following best current practice.

In our first paper, “Quality of Clinical Practice Guidelines on Interventional Management of Low Back Pain: A Systematic Review”, we appraised the quality of 21 LBP CPGs that included interventional management recommendations, using the AGREE II tool. There was a wide range of overall quality scores, though most were recommended for clinical use (albeit some with recommendation of modification). The quality of appraised CPGs showed no association with their characteristics, including year of publication, whether the CPG was updated or new, and region of publication.

The objectives of this analysis are to summarize the main recommendations of the 21 appraised LBP CPGs and to examine the relationship between the quality of the CPGs and their recommendations with the goal of facilitating interpretation of these conflicting CPGs for use in clinical practice.
METHODS
This systematic review was registered with PROSPERO (registration number: CRD42020185582) and followed the guidelines set forth by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (Figure 1).10, 11

Detailed methodology for the quality assessment of the CPGs has been previously published10 and are summarized here. In brief, we searched for all pertinent CPGs in several databases. Next, CPGs were included or excluded based on preset study criteria. Once a final set of CPGs were accrued, their quality was appraised using the AGREE II tool.

Two medical librarians conducted the search in the following bibliographic databases: Ovid MEDLINE® (In-Process & Other Non-Indexed Citations and Ovid MEDLINE® 1946 to Present); Ovid EMBASE (1974 to present); and CINAHL with Full Text (EBSCO). The following guideline databases were also searched by team members: The American Academy of Orthopedics Ortho Guidelines, CPG Infobase, ECRI, Guidelines International Network, NICE Evidence Search, and the Scottish Intercollegiate Guideline Network. The full Ovid MEDLINE search strategy can be referred to in part 1 paper (Figure 2).10 No article type, date, or language restrictions were included in the initial search.

Results from the search were imported into Endnote (a citation management tool), and de-duplicated. The de-duplicated results were screened by title and abstract by two independent reviewers (with discrepancies mediated by a third reviewer) against study criteria (see part 1 for details). The resulting full texts were reviewed by four reviewers in two groups of two with discrepancies mediated by a third reviewer.

From this process, 21 final CPGs resulted.6, 12-31 Their quality was assessed using the AGREE II tool. This assessed 23 items across 6 domains (scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence). Each item was judged on a 7-point Likert scale ranging from 1 (strongly disagree), to 7 (strongly agree). Each CPG was assessed by 4 appraisers. There were 6 total appraisers, with each evaluating 14 CPGs. Each appraiser evaluated all 23 AGREE II items, gave a summative overall quality rating (on a 7-point scale), and gave an overall CPG recommendation (recommend, not recommend, or recommend with modifications). The findings of the appraisal are discussed in detail in the part 1 paper.10

Extraction of Recommendations:
Recommendations pertaining to the interventional treatment of LBP were then extracted from the 21 appraised CPGs, by three independent reviewers, each reviewer extracting from 7 CPGs. Recommendations were then organized into following interventional treatment modality categories: epidural steroid injections (ESIs), radiofrequency procedures (RF), facet injections, sacroiliac injections (SI), medial branch blocks (MBB), spinal cord stimulation (SCS), intradiscal injections, discography, prolotherapy, diagnostic selective nerve root blocks, intradiscal electrothermal therapy (IDET), trigger point injections, and botulinum toxin injections. Other treatment categories were categorized as “miscellaneous”. Some recommendations pertained to multiple treatment categories. The synthesis of and tabular and narrative descriptions of the recommendations from the following treatment categories—ESI, MBB, RF and SI injections—being the most common interventional procedures, are presented in the main body of this manuscript. Recommendations in the following categories are presented as supplemental
material in Appendix 1: Medial Branch Blocks; Other Facet Joint Related Procedures; Spinal Cord Stimulation; Intradiccal Procedures; Discography; Diagnostic Selective Nerve Root Block; Prolotherapy; IDET; Trigger point injections; Botulinum toxin injections; Grouped Treatment Categories and Miscellaneous.

Organization of Recommendations:

The gathered recommendations were then organized within each treatment category based on AGREE II recommendation scores (described in detail in our first paper10). CPGs that received “yes” only votes were included in the high-quality group, CPGs that received a mix of “yes” and “yes with modification” votes were included in the moderate quality group, and any CPG that received a “no” vote were in the low-quality group.

The basis for and the language used to describe the strength of recommendations in the CPGs were variable. For the purposes of this review, recommendations of similar strength were grouped together using a crosswalk table. [Table 1] We created this crosswalk table by cross-matching different CPGs’ recommendation language and descriptions to the reference standard terminology of the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology and descriptions. We utilized the strength of evidence, the quality of studies, and the wording of written recommendations in each respective CPG. This is a table that we created specifically for this review and has not been previously validated or published. Therefore, in order to validate it, several months after its initial creation, we repeated the process of creating the crosswalk table using the same principles detailed above, for a test sample of 11 out of 21 CPGs. Percent agreement between the two crosswalk tables was then calculated to determine internal consistency, as per prior literature stating that the values from 75-90% demonstrate an acceptable level of agreement when using absolute percent agreement.47

Our crosswalk table used the GRADE system as the reference standard because GRADE is the most widely adopted tool (over 100 global organizational endorsements) for grading quality of evidence and for creating recommendations around the world.32 We followed the GRADE system recommendation wording and descriptions as our reference standard, dividing between “strong” and “weak”.33

A summary narrative was then drafted pertaining to each category of interventional treatment, with statements qualified by the strength of recommendations and the quality of the CPG from which the recommendations were extracted.
RESULTS
The crosswalk table was validated by calculating percent agreement between the original creation and the recreation (of a sample of 11 out of 21 CPGs) as a measure of internal consistency (Appendix I, Table 2). Percent agreement between the two tables was 78.1%, demonstrating satisfactory internal consistency.47

SYNTHESIS AND DISCUSSION of EXTRACTED RECOMMENDATIONS
The 21 CPGs reviewed included recommendations for interventional management of not only axial LBP, but also for adjacent clinical conditions including disc herniation/radiculopathy and spinal stenosis. Because the intended objective of this review was to extract the recommendations pertaining to axial, non-radicular sub-acute and chronic LBP only (including LBP described as non-specific, axial, or non-radicular) the synthesis of recommendations presented and discussed below are restricted only to axial, non-radicular sub-acute and chronic LBP. Recommendations pertaining to radiculopathy were not included. The vast majority of recommendations were not extracted as they pertained to radiculopathy were in the ESI category.

Epidural Steroid Injections (ESI):
For the management of chronic LBP, from the high-quality CPG group, there was one strongly-against recommendation from CPG #813, one recommendation weakly in support of caudal and lumbar interlaminar epidural injections from CPG #130, three inconclusive recommendations (CPGs #618 and #1121), and no recommendations strongly in support of ESIs. The only recommendation from a moderate-quality CPG (#1531) was strongly-against the use of ESIs for the treatment of non-radicular LBP. The two recommendations from a low-quality CPG (#1820) were both weakly in support of the use of ESIs.

Radiofrequency Procedures (RF):
For RF procedures, two recommendations were strongly-against, two weakly-against, six inconclusive, and 14 weakly-for. The specified diagnoses included chronic LBP, facet related LBP, and discogenic LBP and specified targets included lumbar, sacroiliac joint (SI), coccygeal, and ramus communicans targets, but several recommendations did not specify diagnosis or anatomical region.

Diagnosis-specific: For the recommendations pertaining to chronic LBP, there were two weakly-against and three weakly-for, all from high-quality CPGs. The weakly-for recommendations stated that RF should only be performed after positive response to medial branch blocks (MBB) (CPG #225 and #1224). One weakly-against recommendation qualified their recommendation that with a successful diagnostic block and optimal technique RF might yield better results.

For facet mediated LBP, there were 2 inconclusive and five weakly-for recommendations. Two of the weakly-for recommendations required failed prior conservative treatment, and suspected - medial branch involvement. One of the inconclusive recommendations was specifically about percutaneous intradiscal RF.

For discogenic pain one strongly-against, one inconclusive, and one weakly-for recommendation were identified. The strongly-against recommendation and weakly-for were both from a low-quality CPG (#1617). After failed conservative treatment, RF to the dorsal ramus was weakly supported and disc RF was strongly recommended against.

Target-specific: RF to the ramus communicans was weakly in support. For RF targeting the lumbar region, one strongly-against recommendation from a low-quality CPG (#1617), and one
weak-for from a moderate-quality CPG (#10\textsuperscript{15}) were identified. For SI joint, one weak-for recommendation from a low-quality CPG (#16\textsuperscript{17}) stated that after failed SI joint intraarticular steroid injection, cooled RF or RF can be considered. RF targeting coccygeal pain was inconclusive.

Unspecified (No Target or Diagnosis Specified): 2 inconclusive, 5 weakly-for, and 1 strongly-for recommendations were noted. CPG #7\textsuperscript{18} had 4 separate weakly-for recommendations that stated that larger lesions could yield better outcomes, repeat RFA can be useful without needing repeat MBBs. From a high-quality CPG#1, one recommendation was weakly in support of pulsed RF\textsuperscript{30}, and a strongly-for recommendation supported confirmatory diagnostic facet nerve blocks.

**Sacroiliac Joint Injections (SI):**

This section included one weakly-against, three inconclusive, four weakly-for, and one strongly-for recommendation. All were from high-quality CPGs except one inconclusive recommendation from a moderate-quality CPG and one weakly-for recommendation from a low-quality CPG (#16\textsuperscript{17}). The recommendations discussed three diagnostic/treatment modalities related to SI procedures: cooled RF neurotomy/ablation, SI blocks, and intra-articular (IA) injections. The recommendations for cooled RF procedures were both weakly-for, and from high-quality CPGs (#13\textsuperscript{10} and #6\textsuperscript{18}). Both recommendations stated that the procedure is recommended after initial diagnosis with SI joint injection/block. One inconclusive (CPG #9\textsuperscript{19}, moderate quality) and one strongly-for (CPG #13\textsuperscript{10}, high quality) recommendation for SI blocks were noted. One weakly-against, two inconclusive, and two weakly-for recommendations for IA injections were found. One recommendation (high-quality CPG #8\textsuperscript{13}) was weakly-against and one (low-quality CPG #16\textsuperscript{17}) was weakly-for IA steroids.

**DISCUSSION:**

In this systematic review, we extracted and summarized the recommendations pertaining to interventional management of LBP from 21 LBP CPGs. Based on our appraisal of the quality of the 21 CPGs using the AGREE II instrument, we were able to stratify them into high, moderate, and low-quality groups. We created a crosswalk table to match the various evidentiary terms used in the 21 CPGs to the GRADE terminology, which allowed us to further stratify recommendations based on strength of recommendation, within each interventional treatment category. Other recent reviews that included critical appraisal of LBP CPGs did not primarily focus on interventional management, utilize a crosswalk table to match evidentiary terms or compile the relevant recommendations and create a summative narrative.\textsuperscript{9, 34-36}

The number of CPGs included in our review was the same as one recent overview study,\textsuperscript{34} but significantly higher than other similar overview studies.\textsuperscript{9, 35, 36} This may be because we cast a wider net in our search to include CPGs that not only exclusively focused on interventional management of LBP but also LBP guidelines that were intended for primary care audiences that reviewed overall management of LBP including interventional management recommendations.

There were several contradictory recommendations within some treatment categories. For example, with ESIs, CPG #1 (high-quality) stated “for axial or discogenic pain, either lumbar interlaminar or caudal epidural injections are recommended”\textsuperscript{30} and CPG #18 (low-quality) stated “lumbar ESIs are an option for short-term relief of chronic low-back pain without radiculopathy in patients with degenerative disease of the lumbar spine” and “caudal ESIs are an option for decreasing low-back pain of greater than 6 weeks' duration, without radiculopathy, in patients
with degenerative disease of the lumbar spine”20, whereas CPG #8 (moderate-quality) posits “there is no evidence for the effectiveness of epidural corticosteroids in patients with non-radicular, nonspecific low back pain”13 and CPG #15 (moderate-quality) says “for the long-term reduction of radicular low back pain, non-radicular low back pain, or spinal stenosis, we recommend against offering spinal epidural steroid injections”.31

There are several possible explanations for these discrepancies. These include differences in a) CPG focus/scope, b) creating group/organization and c) inherent differences in the body of evidence that informed each CPG (i.e., which studies-observational studies, randomized controlled trials (RCTs) or systematic reviews were included). Evaluation of quality of evidence of the source articles that informs the CPGs is baked into the AGREE II tool (Item 9—“The strengths and limitations of the body of evidence are clearly described” and Item 12—“There is an explicit link between the recommendations and the supporting evidence”, Domain 3-Rigor of Development). However, although this domain (one of six domains) describes and assesses how evidence was searched for and how it informed eventual recommendations, bias and inaccuracies in interpretation of such evidence in each CPG may not be adequately appraised using this tool.37

In terms of scope and focus, there was variability across the CPGs. For instance, CPG #9 pertained to the treatment of LBP through the lens of primary care19 (CPG title “Evidence-Informed Primary Care Management of Low Back Pain”) while CPG #18 was from a set of guidelines primarily focused on the surgical treatment of LBP20 (CPG title “Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 13: Injection therapies, low-back pain, and lumbar fusion”). While both CPGs met our study criteria and contained relevant recommendations pertinent to interventional treatment of LBP, they did not contain any overlapping recommendations pertaining to the same treatment category. The “scope and focus” domain covers the CPG’s description of objectives, health questions covered and populations to whom the CPG is meant to apply. As a result, CPGs addressing recommendations of the same clinical condition may still have different recommendations due to differences in their “scope and focus”.37

Differences in composition of the CPG developers could also have been a factor in the recommendation discrepancies. CPG developers were either professional societies or governmental organizations. Differences in recommendations could have been secondary to differences in purpose and objectives, focus and target audience for each type of developer. For instance, one of the government organizations (the National Institute for Health and Care Excellence (UK)), who developed CPG #2, included healthcare commissioners and other administrative level professionals as part of their target audience.25 This contrasts with a professional society creating a CPG, such as the British Pain Society, who created CPG #19, who targeted their CPG to those with direct patient contact.23 It is conceivable that a CPG created by a governmental organization, composed of content experts with fewer potential conflicts of interest, and a targeted primary readership of healthcare administrators, could have different recommendations than a professional society, composed of content experts with higher conflicts of interest, and with a CPG targeted to a primary readership of clinical practitioners, although CPG #2 and #19 did not contain any overlapping recommendations pertaining to the same treatment category.

Inconclusive or weak recommendations in the collected CPGs are not entirely unexpected given that majority of the source articles that inform the CPGs are RCTs, which rarely report treatment
effect with large effect sizes, due to several factors including heterogeneity of patient populations or clinical conditions studied. This pattern of CPGs with several inconclusive recommendations is not unique to pain management and has been noted in topic areas such as spinal cord injury, critical care, and oncology. Similarly, Cochrane reviews and systematic reviews that are also largely based on RCTs also frequently report inconclusive recommendations. Inconclusive recommendations therefore need to be judiciously interpreted and applied to the specific context of the individual patient’s care.

Disparate search strategy and criteria for inclusion and exclusion that formed the evidence-basis of CPGs could also have influenced recommendation contradictions. Recommendations based on different bases of evidence are likely to be different. For instance, CPG #6 had an inconclusive recommendation, “there is insufficient evidence to make a recommendation for or against the use of interlaminar epidural steroid injections in patients with low back pain” based on one observational study. This contrasts with another recommendation which was strongly against use of ESIs from CPG #15 which stated “for the long-term reduction of radicular low back pain, non-radicular low back pain, or spinal stenosis, we recommend against offering spinal epidural steroid injections”. CPG #15’s recommendation was based on 4 RCTs and 4 systematic reviews and did not include the observational study included in CPG #6. Therefore, differences in the body of evidence that informed each CPG were likely a major factor that contributed to recommendation contradictions. CPG#6’s inclusion criteria were more restrictive than CPG#15’s resulting in the exclusion of the RCTs and SRs that informed CPG #15’s recommendations in this treatment category.

These contradictory recommendations can make it challenging for interventional pain practitioners to apply these guidelines to their daily practice and may contribute to the wide variations observed in clinical practice. We are hopeful however that our attempt at creating the crosswalk table using the widely accepted GRADE as our reference standard and recommendations summary tables (Tables 1 & 2), stimulates future efforts at harmonizing CPG recommendation drafting standards.

CPGs are intended to guide clinical care, by supporting use of treatments that have strong evidence basis and discouraging use of treatments that have strong evidence of no efficacy or potential for harm. Historically, CPGs have had minimal effect on practice patterns, for a variety of factors including but not limited to economic, social, cultural, and educational factors. We are hopeful that information contained in 3rd order reviews (systematic review of CPGs that synthesize systematic reviews) such as this paper may be more meaningful and impactful to effect evidence-based changes in individual clinician’s practice pattern.

The inconsistencies in the CPG recommendations also help identify knowledge gaps that need to be filled by future research that may better clarify the evidence-informed utilization of interventional procedures in treating chronic pain. Better identification of CPGs that are best aligned with the clinical question and outcomes of interest to the clinician, and increased adherence to the applicable CPG’s recommendations is likely to improve health outcomes and reduce healthcare utilization and costs.

Limitations:
This systematic review used a rigorous study design and methodology to search the literature and identify 21 CPGs that met our criteria. Despite this rigor, this review had several limitations. Firstly, the study criteria may have resulted in missing some CPGs that addressed interventional
management of LBP as a part of a larger population. It is possible that if LBP management was discussed in other adjacent clinical conditions than those examined in this study, we may not have identified those CPGs. Secondly, grouping CPGs into high, moderate, and low quality was based on arbitrary thresholds (AGREE II was used to appraise CPGs and determine quality scores, but the thresholds for high, moderate, and low-quality groups were arbitrary), and may not be reflective of the confidence/trust in their recommendations. This quality-based grouping, which was performed to enable quality-sorting and discussing the recommendations was not based on any previously published criteria. CPGs are inherently limited by recommendations that are reflective of research that is a few years old and may not include pragmatic clinical trials that are more reflective of actual clinical practice. Therefore, this study, a systematic review of such CPGs, is also likely subject to the same limitations. Finally, the crosswalk table used in this study was created solely for this study and has not been previously validated. We used our best judgment to match evidentiary terms used in the CPGs to the GRADE terminology. We validated this crosswalk table by calculating internal consistency between the original and a recreated version of the crosswalk table which yielded a percent agreement of 78.1%, demonstrating internal consistency. However, because this table has not been externally validated, the strength of recommendation may have been mis-characterized. The goals of this paper were to provide a synthesis and summary of the recommendations from the 21 CPGs appraised in our previously published paper. We did not attempt to review the source articles that informed the CPGs or resolve discrepancies in the recommendations made in those CPGs. While those objectives are worth pursuing, they were beyond the scope of this study.

CONCLUSIONS:
The AGREE II quality appraisals of CPGs on interventional management of LBP performed in this study were of unclear utility in guiding their clinical implementation. Recommendations on interventional management of non-radicular LBP extracted from the 21 CPGs included in our review, covered a wide range of treatment categories. Most of the gathered recommendations were either weakly-for, weakly-against or inconclusive, with several recommendations within each treatment category contradicting each other. We recommend further efforts at standardizing CPG development and dissemination standards including harmonization of recommendation drafting standards, so that developer bias is minimized and the clinical utility and impact of CPGs is improved.
References


22. EBHC KT (SA) - Chair of Evidence-based Healthcare and Knowledge translation CoM, King Saud University. KSUMC Evidence-Based Clinical Practice Guideline for Management of Persistent non-specific low back pain (KSUMC CPG 1st Edition 2013): Adapted from source


40. Harvey, L. Randomised controlled trials do not always give the results we want but that doesn’t mean we should abandon randomised controlled trials. Spinal Cord 53, 251 (2015). https://doi.org/10.1038/sc.2015.59


Table 1: Crosswalk Table

<table>
<thead>
<tr>
<th>CP G #</th>
<th>“Strongly Against” Defined As</th>
<th>“Weakly Against” Defined As</th>
<th>“Strongly For” Defined As</th>
<th>“Weakly For” Defined As</th>
<th>Inconclusive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>&quot;are recommended&quot; (with good evidence)</td>
<td>&quot;are recommended&quot; (with limited or fair evidence)</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>“Do not offer”</td>
<td>“Do not routinely offer”</td>
<td>n/a</td>
<td>&quot;consider _____&quot;, &quot;only perform _____ after&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>“Strong recommendation against” (DHA guideline)</td>
<td>“Weak/conditional recommendation against” (DHA guideline)</td>
<td>n/a</td>
<td>&quot;should only be offered to&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>4</td>
<td>“Do not offer”</td>
<td>n/a</td>
<td>n/a</td>
<td>&quot;some evidence&quot;, &quot;limited evidence&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>5</td>
<td>Strong recommendation (for GRADE recommendations)</td>
<td>Weak recommendation (for GRADE recommendations)</td>
<td>strong recommendation</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>6</td>
<td>Grade A recommendation (good evidence)</td>
<td>Grade B recommendation (fair evidence) or Grade C recommendation (poor quality evidence)</td>
<td>Grade A recommendation (good evidence)</td>
<td>Grade B recommendation (fair evidence) or Grade C recommendation (poor quality evidence)</td>
<td>Insufficient evidence (I)</td>
</tr>
<tr>
<td>7</td>
<td>Grade D recommendation + High level of certainty</td>
<td>Grade D recommendation + Moderate level of certainty</td>
<td>Grade A + high level of certainty</td>
<td>Grade B or C + moderate or low level of certainty</td>
<td>insufficient</td>
</tr>
<tr>
<td>8</td>
<td>“Do not recommend” + Level A (strong evidence)</td>
<td>“Do not recommend” + Level B (moderate evidence)</td>
<td>n/a</td>
<td>&quot;consider _____&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>9</td>
<td>“Do Not Do” recommendation</td>
<td>n/a</td>
<td>n/a</td>
<td>&quot;may be useful&quot;</td>
<td>&quot;inconclusive&quot;</td>
</tr>
<tr>
<td>10&lt;sup&gt;15&lt;/sup&gt;</td>
<td>n/a</td>
<td>n/a</td>
<td>strong or moderate-strong strength of evidence</td>
<td>moderate or weak strength of recommendation</td>
<td>n/a</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>11&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Grade D recommendation + Good strength of evidence</td>
<td>Grade D recommendation + Fair strength of evidence</td>
<td>&quot;strong recommendation&quot;, grade A</td>
<td>&quot;moderate recommendation&quot;, &quot;weak recommendation&quot;, grade B</td>
<td>&quot;unable to estimate&quot;, grade C or I</td>
</tr>
<tr>
<td>12&lt;sup&gt;24&lt;/sup&gt;</td>
<td>“Do not offer” (GRADE system)</td>
<td>“Do not routinely offer” (GRADE system)</td>
<td>n/a</td>
<td>&quot;consider _____&quot;, &quot;only do _____ if ____&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>13&lt;sup&gt;28&lt;/sup&gt;</td>
<td>GRADE “low quality” + individual search terms in each database’s search engine</td>
<td>GRADE “low quality” + individual search terms in each database’s search engine</td>
<td>SoR: strong, &quot;strongly recommend&quot;</td>
<td>SoR: weak, &quot;weakly recommend&quot;</td>
<td>SoR: inconclusiv e</td>
</tr>
<tr>
<td>14&lt;sup&gt;29&lt;/sup&gt;</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>15&lt;sup&gt;31&lt;/sup&gt;</td>
<td>“We recommend against offering this option …”</td>
<td>“We suggest not offering this option …”</td>
<td>n/a</td>
<td>&quot;we suggest offering&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>16&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Evidence Score IV (negative implication for practice)</td>
<td>n/a</td>
<td>&quot;is recommended&quot;</td>
<td>&quot;can be&quot;, &quot;may be&quot;</td>
<td>n/a</td>
</tr>
<tr>
<td>17&lt;sup&gt;22&lt;/sup&gt;</td>
<td>“Do not offer” or “Do not refer” recommendation s</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>18&lt;sup&gt;20&lt;/sup&gt;</td>
<td>Grade A recommendation (good evidence)</td>
<td>Grade B recommendation (fair evidence) or Grade C recommendation (poor quality evidence)</td>
<td>Grade A recommendation (good evidence)</td>
<td>Grade B recommendation (fair evidence) or Grade C recommendation (poor quality evidence)</td>
<td>n/a</td>
</tr>
<tr>
<td>19&lt;sup&gt;23&lt;/sup&gt;</td>
<td>n/a</td>
<td>n/a</td>
<td>&quot;are recommended&quot;</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>20²⁶</td>
<td>Grade A recommendation (NHMRC guideline)</td>
<td>Grade B recommendation or Grade C recommendation (NHMRC guideline)</td>
<td>n/a</td>
<td>n/a</td>
<td>&quot;insufficient evidence&quot;</td>
</tr>
<tr>
<td>21¹⁴</td>
<td>“Not recommended”</td>
<td>n/a</td>
<td>&quot;recommended&quot;</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

The column headers reflect terminology for strength of recommendation used in the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology and descriptions, which was adopted as our standard. Each row reflects the terminology of the appraised CPG’s strength of recommendation matched to the closest strength terminology in the adopted standard. When no matching terminology was identified, the corresponding cell was marked as not applicable or “n/a”.

If a clinical practice guideline (CPG) did not differentiate between “strongly against” and “weakly against,” we categorized all their recommendations as “strongly against.”

---

²⁶NHMRC: National Health and Medical Research Council
¹⁴NHMRC: National Health and Medical Research Council

"Recommended" indicates a recommendation that is supported by strong evidence.
"Not recommended" indicates a recommendation that is not supported by strong evidence.
"Insufficient evidence" indicates that there is not enough evidence to make a recommendation.
"n/a" indicates that no matching terminology was identified.
Table 2: Recommendations by CPG Quality and Crosswalk Grade of Recommendation Strength

<table>
<thead>
<tr>
<th>Tx. Grp</th>
<th>GRADE Strength</th>
<th>Strongly Against</th>
<th>Weakly Against</th>
<th>Inconclusive</th>
<th>Weakly For</th>
<th>Strongly For</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidural Steroid Injections</td>
<td>#16-2</td>
<td></td>
<td>#8</td>
<td>#11</td>
<td>#6-2</td>
<td>#18-2</td>
</tr>
<tr>
<td></td>
<td>#16</td>
<td></td>
<td>#11-3</td>
<td>#9</td>
<td>#11-3</td>
<td>#6</td>
</tr>
<tr>
<td></td>
<td>#18</td>
<td></td>
<td>#16-3</td>
<td>#10</td>
<td>#16-3</td>
<td>#16-3</td>
</tr>
<tr>
<td>Radiofrequency Procedures</td>
<td>#8</td>
<td></td>
<td>#11-3</td>
<td>#9</td>
<td>#2-2</td>
<td>#16-3</td>
</tr>
<tr>
<td></td>
<td>#9</td>
<td></td>
<td>#16-4</td>
<td>#10</td>
<td>#10</td>
<td>#16-3</td>
</tr>
<tr>
<td></td>
<td>#16</td>
<td></td>
<td>#18</td>
<td>#12-3</td>
<td>#16-3</td>
<td>#16-3</td>
</tr>
<tr>
<td>Sacroiliac Joint Injections</td>
<td>#11</td>
<td></td>
<td>#2-2</td>
<td>#6</td>
<td>#2-2</td>
<td>#6-2</td>
</tr>
<tr>
<td></td>
<td>#9-2</td>
<td></td>
<td>#11</td>
<td>#6-2</td>
<td>#6-2</td>
<td>#16</td>
</tr>
</tbody>
</table>

Treatment categories are listed in rows.

Strength of recommendations are listed in columns (Columns 1-5 are: Strongly-Against, Weakly-Against, Inconclusive, Weakly-For and Strongly-For).

CPG Quality: no highlight: high quality CPG, light-grey highlight: moderate quality CPG, dark-grey highlight: low quality CPG

Each cell lists the CPG list number followed by the number of recommendations in that treatment category and strength of recommendation (if more than 1) (e.g., middle column [Inconclusive], 1st row: #6-2 indicates that there were two recommendations from CPG #6 that were inconclusive).
Figure Legends

Figure 1: PRISMA Flow Chart. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart of search results

Figure 2: Search Strategy. Search strategy used in this systematic review
**Figure 1:**

PRISMA 2009 Flow Diagram

- Records identified through database searching (n = 714)
- Additional records identified through other sources (n = 769)
- Records after duplicates removed (n = 578)
- Records screened (n = 578)
- Full-text articles assessed for eligibility (n = 79)
- Studies included in qualitative synthesis (n = 21)
- Full-text articles excluded, with reasons (n = 64)
  - Not CPGs (n = 22)
  - Pertaining to a specific back pathology (n = 8)
  - Not the most recent version of the guideline (n = 7)
  - Duplicates (n = 6)
  - Not in English (n = 4)
  - Not intervention focused (n = 4)
  - Unavailable for a variety of reasons (for instance, one of the parent organizations had gone bankrupt and disbanded) (n = 4)
  - Focused on the acute phase (n = 3)
  - Surgical focus (n = 2)
  - Triplicates (n = 2)
  - Imaging focused (n = 1)
  - Focused on one specialty (n = 1)
- Studies included in quantitative synthesis (meta-analysis) (n = 21)


For more information, visit [www.prisma-statement.org](http://www.prisma-statement.org).
Figure 2:

Ovid MEDLINE search strategy:

1. low back pain or ((low* adj3 back adj3 pain*) or (low* adj3 back adj3 ache*) or (low* adj3 backache*) or (low* adj3 back adj3 pain*) or (low* adj3 back adj3 ache*) or (low* adj3 backache*) or (lumbar adj3 pain*) or (lumbar adj3 pain*) or (lumbar adj3 pain*) or (lumbar adj3 pain*) or (lumbar adj3 pain*) or (lumbar adj3 spine adj3 syndrome*) or (lumbar giesia or lumbargia or (lumbar adj3 pain*) or (lumbar adj3 spine adj3 syndrome*) or (lumbarosacral adj3 pain*) or (lumbarosacral adj3 syndrome*) or (lumbarosacral spine adj3 strain*)).mp.
2. Injections, Spinal/ or ((spinal adj3 injection*) or (intraspinal adj3 injection*) or (intrathecal adj3 injection*) or (spinal adj3 infusion*) or (spinal adj3 administration) or (intraspinal adj3 administration) or (intra-spinal adj3 administration)).mp.
3. Injections, Epidural/ or ((epidural adj3 administration) or (epidural adj3 application*) or (epidural adj3 dose*) or (epidural adj3 infiltration*) or (epidural adj3 infusion*) or (epidural adj3 medication*) or (epidural adj3 treatment*) or (extradural adj3 administration) or (extradural adj3 infusion*) or (extradural adj3 injection*) or (peridural adj3 administration) or (peridural adj3 injection*)).mp.
4. radiofrequency ablation/ or catheter ablation/ or ((radiofrequency adj3 ablation) or (radio adj3 frequency adj3 ablation) or (radio-frequency adj3 ablation) or RFA).mp.
5. nerve block/ or ((autonomic adj3 block*) or (nerve adj3 block*) or (conduction adj3 block*) or (neurogenic adj3 block*) or (medial adj3 branch adj3 block*)).mp.
6. Injections, Intra-Articular/
7. Injections, Intra-Articular/ or ((intraarticular adj3 administration) or (intraarticular adj3 delivery) or (intraarticular adj3 injection*) or (intraarticular adj3 infusion*) or (intraarticular adj3 medication*) or (intraarticular adj3 treatment) or (intra-articular adj3 administration) or (intra-articular adj3 delivery) or (intra-articular adj3 injection*) or (intra-articular adj3 infusion*) or (intra-articular adj3 medication*) or (intra-articular adj3 treatment) or (intracoxal adj3 administration) or (joint adj3 infusion*) or (joint adj3 injection*) or (intrasynovial adj3 administration) or (intrasynovial adj3 injection*)).mp.
8. exp rhizotomy/ or rhizotomy*.mp.
9. glucocorticoids/ or (glucocorticoid* or glyocorticoid*).mp.
10. Steroids/ or (steroid* or cyclosteroid*).mp.
11. spine/ or (spine or spinal or vertebra* or (columna adj3 dorsal) or (dorsal adj3 column)).mp.
12. zygaphyseal joint/ or (zygaphyseal or zygaphysial or zygaphyseis or facet*).mp.
13. sacroiliac joint/ or (sacroiliac* or ileosacral or iliocacral or transforamin*).mp.
14. Injections/ or (injection* or injectable* or shot* or intervention* or procedure*).mp.
15. 11 or 12 or 13
16. 14 and 15
17. Guideline/ or Guidelines as Topic/ or Practice Guideline/ or guideline*.mp.
18. 1 and 16 and 17