Poor overall quality of clinical practice guidelines for musculoskeletal pain: a systematic review

Ivan Lin,1 Louise K Wiles,2 Robert Waller,3 Roger Goucke,4 Yusuf Nagree,5,6 Michael Gibberd,7 Leon Straker,3 Christopher G Maher,8 Peter P B O’Sullivan3

ABSTRACT

Objectives Undertake a systematic critical appraisal of contemporary clinical practice guidelines (CPGs) for common musculoskeletal (MSK) pain conditions: spinal (lumbar, thoracic and cervical), hip/knee (including osteoarthritis) and shoulder.

Design Systematic review of CPGs (PROSPERO number: CRD42016051653). Included CPGs were written in English, developed within the last 5 years, focused on adults and described development processes. Excluded CPGs were for: traumatic MSK pain, single modalities (eg, surgery), traditional healing/medicine, specific disease processes (eg, inflammatory arthropathies) or those that required payment.

Data sources and method of appraisal Four scientific databases (MEDLINE, Embase, CINAHL and Physiotherapy Evidence Database) and four guideline repositories. The Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument was used for critical appraisal.

Results 4664 records were identified, and 34 CPGs were included. Most were for osteoarthritis (n=12) or low back pain (n=11), most commonly from the USA (n=12). The mean overall AGREE II score was 45% (SD=19.7). Lowest mean domain scores were for applicability (26%, SD=19.5) and editorial independence (33%, SD=27.5). The highest score was for scope and purpose (72%, SD=14.3). Only 8 of 34 CPGs were high quality: for osteoarthritis (n=4), low back pain (n=2), neck (n=1) and shoulder pain (n=1).

INTRODUCTION

Musculoskeletal (MSK) pain conditions are the leading contributor to the burden of disease (BOD) in developed and developing countries and a major reason why people seek healthcare. There have been many clinical practice guidelines (CPGs) developed aiming to improve the quality of care for MSK pain; however, there is increasing evidence that the quality of MSK care is suboptimal, and there are substantial evidence-to-practice gaps. For example, more than one quarter of patients with low back pain are referred for radiological imaging on first presentation with rotator cuff tendinopathy even though it is infrequently warranted, and inappropriate imaging can increase the risk of iatrogenic patient harm. Against best available evidence, 69% and 82% of Australian general practitioners would refer patients for an X-ray or ultrasound (respectively) on first presentation with rotator cuff tendinopathy. Shoulder arthroscopy rates have increased by 55% between 2001 and 2013 despite a lack of supporting evidence. In the US Veterans Health Administration system, 4% of individuals with knee osteoarthritis undergo knee arthroscopy annually, although there is limited clinical benefit. These practises are problematic because at best they represent ineffective, expensive and inefficient care, and at worst, they serve to increase the burden of MSK pain.

One of the foundations of efforts to improve healthcare are CPGs. The Institute of Medicine defines CPGs as ‘statements that include recommendations intended to optimise patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options’. CPGs have the potential to improve healthcare quality in multiple ways: (A) to guide clinicians by informing their decision making in patient care, (B) to identify appropriate standards of care (eg, quality indicators) and therefore identify evidence-to-practice gaps and (C) as a basis for education and continuing professional development. Therefore, CPGs are a vehicle by which to drive improvement in healthcare delivery and reduce the burden of prevalent health conditions such as MSK pain.

For optimal usability and fitness-for-purpose, CPGs need to be contemporary, valid and of high quality. Given the resourcing required for their development, CPGs should be created using systematic, rigorous and transparent processes so that end-users can trust their recommendations. Some of the commonly cited problems with CPGs are: the sheer numbers available (eg, for the same condition), voluminous documents that are not easy to assimilate or use, inconsistently opportunities for end-users to provide formal feedback, lack of detail regarding how evidence was interpreted and weighted to formulate recommendations and having been developed by people with (often undisclosed) professional or commercial conflicts of interest. In practice, this can manifest as confusion and ambiguity about what constitutes ‘recommended care’ and contribute to unwarranted variations in clinical care. To reduce evidence-to-practice gaps in MSK pain care, there is a critical need for consistent, high-quality and trustworthy guidelines.

Musculoskeletal pain CPGs for the most common MSK pain conditions have never been appraised. One recent systematic review focused on chronic MSK pain, but most CPGs in this review were for generic chronic pain or the use of opioids. Therefore, the aim of our review was to describe the characteristics, methods used for development and quality of contemporary CPGs for MSK pain using a systematic critical appraisal approach.
Box  Clinical practice guidelines (CPGs) selection criteria

**Inclusion criteria**
- Published between January 2011 and September 2016.
- Created for one of: spinal pain (lumbar, thoracic and cervical spine), hip/knee pain including hip/knee osteoarthritis or shoulder pain.
- Relating to assessment and treatment (ie, the processes of care in a clinical management plan).
- For adult populations (aged >18 years).
- Published in the English language or a complete English language version was available.
- Details of CPD development processes were available (ie, methods were described in sufficient detail).
- Were based on an original body of work (ie, not solely an adaptation or systematic review of existing guidelines).

**Exclusion criteria**
- CPGs related to a single treatment modality including; surgery, massage, manipulation or pharmacology.
- CPGs related to traditional healing/medicine (eg, traditional Indigenous medicine).
- CPGs for pain arising from musculoskeletal (MSK) tissues, related to a specific disease process that requires a specific clinical care pathway, including osteoporosis, frozen shoulder, rheumatoid arthritis and other inflammatory arthropathies, infection and cancer.
- CPGs for traumatic MSK pain only (eg, whiplash).
- CPGs that address recommendations for the system/organisation of care.
- CPGs requiring payment to access.

**METHODS**

We undertook a systematic review of contemporary CPGs for three of the most common MSK pain conditions: spinal pain (lumbar, thoracic and cervical spine), hip/knee pain including hip/knee osteoarthritis and shoulder pain. This study was part of a wider project examining MSK pain care recommendations in primary care and emergency care for non-traumatic MSK pain. Therefore, we included CPGs relevant to these care settings (box). We defined a CPG as being identified by the authors as such and consistent with the definition of the Institute of Medicine. We excluded CPGs for acute MSK pain conditions due to trauma (eg, acute whiplash) and pain arising from musculoskeletal tissues caused by a specific disease process that requires a specific clinical care pathway (eg, rheumatoid arthritis or other inflammatory arthropathies) (box). This article forms the first stage of a larger review registered with the International Prospective Register of Systematic Reviews (PROSERO number: CRD42016051653).

**Review team**

The review team included three academic and practising physiotherapists (IL, RW and PO), two MSK pain researchers (CGM and LS), an indicator development researcher (LW), a specialist emergency care physician (YG), a senior medical officer (MG) and a pain medicine physician (RG).

**Search strategy**

The search was guided by a reference librarian. We undertook a systematic search of scientific databases (MEDLINE (including the Cochrane library), Embase, the Cumulative Index to Nursing and Allied Health Literature - CINAHL, and the Physiotherapy Evidence Database - PEDro). We also searched four online guideline repositories: Guidelines International Network (G-INA), National Health and Medical Research Council (NHMRC), the National Guideline Clearinghouse of the Agency for Health Care Research and Quality (USA) and the National Institute for Health and Care Excellence (NICE).

The database search combined key words and Medical Subject Headings related to CPGs (eg, exp guideline/OR clinical guideline”.mp”) and the MSK pain conditions of interest (eg, exp Osteoarthritis/OR exp Back Pain) (online supplementary file 1). The search range was January 2011–September 2016.

**Study selection**

Search results were imported into a series of Endnote libraries and duplicates identified. Article titles/abstracts were screened by a single reviewer (IL). Following title/abstract screening, relevant CPGs were imported to the Cochrane systematic review software (Veritas Health Innovation, Melbourne, Australia. Available at www.cochrane.org) for management. Two reviewers independently screened the full texts (IL and LW). Final inclusion of articles was agreed on by consensus. While undertaking searches, we identified one CPG for the assessment an management of low back pain in draft form, scheduled for publication in September 2016. As we had identified the draft version during our searches, we decided to include this in our review following its publication.

**Data extraction and appraisal**

Data were extracted to a purpose-designed spread sheet. Extracted variables included: the title/topic, developer, type of developer, first author (if applicable), accompanying documents, number of pages and country of origin. The country of origin was described as ‘Europe’ if there were authors from multiple European countries and ‘international’ if two or more authors were from different continents. Accompanying documents were sourced where relevant (eg, methodological reports). CPG developers were contacted by email to request further information if it was not readily available.

Each CPG was independently appraised by three reviewers (IL, RW and LW) using the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument. The AGREE II instrument was developed by the AGREE Collaboration as a generic instrument to evaluate the development and reporting of all CPGs. Its implementation is supported by a user manual, training tools and a web-based platform to complete AGREE II appraisals online. Two overall assessment scores are assigned based on the score of 23 core items grouped under six domains: scope and purpose, stakeholder involvement, rigour of development, clarity of presentation, applicability and editorial Independence. Each item is ranked on a seven-point scale (1: strongly disagree to 7: strongly agree). The AGREE II is a valid and reliable tool for use with any practice guideline in any disease area and is the most widely used guideline appraisal instrument.

Prior to appraisal, reviewers completed two training exercises available on the AGREE Enterprise website. Reviewers met after appraising a test CPG, and again after 10 CPGs were completed to review scoring and as a ‘quality check’ of interpretation of the instrument.

**ANALYSIS**

Using AGREE PLUS on the AGREE II website, scores for each domain were calculated as a percentage, by summing all scores...
of the individual items in a domain and by scaling the total as a percentage of the maximum possible score for that domain. The final ranked item of the AGREE II instrument, in which the overall quality of the guideline is rated (1–7), was calculated manually as a percentage of the maximum possible score. Data were entered and analysed using SPSS (IBM SPSS Statistics V.24.0). Means and SD for each item (1–7 scale) and overall domain score (percentage) were calculated. Inter-rater agreement was determined using intraclass correlation coefficients (ICC) with two-way random effects model. We calculated ICCs for each domain and overall rating scores. We classified level of reliability according to Cicchetti (1994) as poor (ICC <0.40), fair (ICC 0.40–0.59), good (ICC 0.60–0.74) or excellent level of agreement (ICC 0.75–1.00).29

The AGREE II developers do not provide cut-off scores for high/low quality CPGs. Consistent with previous research,30–32 CPGs were rated as higher quality when domain scores, in three domains we believed were most important for validity24,33 were equal to or greater than 50% of the maximum possible score. The domains of interest were: rigour of development (domain 3), editorial independence (domain 6) and stakeholder involvement (domain 2).

RESULTS
Searches identified 4664 discrete records, from which 34 CPGs were eventually selected for inclusion (figure 1 – Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart). Twenty-three of the 34 selected CPGs provided recommendations for osteoarthritis (n=12), or low back pain (n=11) (online supplementary file 2).

Characteristics of included CPGs
With one exception (Malaysia34), all CPGs were developed by high-income countries or collaborations involving high-income countries. More than one-third of CPGs were from the USA (n=12), including seven for low back pain (online supplementary file 2). International collaborations accounted for nearly one-fifth of our sample (n=6), of which three were for osteoarthritis. Multiple CPGs had been developed in Canada (n=5), the Netherlands (n=2), UK (n=2) and North America (developers from Canada and the USA, n=2).

Most CPGs were developed by medical societies (n=18) and government agencies (n=7). Other developers included research groups—usually university based (n=3), expert panels (n=3) and ‘other’—that included non-profit organisations or did not specify the developer type (n=3). Medical societies focused on

Figure 1  PRISMA flow diagram. CPGs, clinical practice guidelines; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
In addition, the ESCEO produced a CPG and algorithm for the and lumbar disc herniation with radiculopathy. Similarly, the NICE produced CPGs for osteoarthritis and low back pain. Against Rheumatism, American Academy of Orthopedic CPGs were: four for osteoarthritis, from the European League The potential resource implications of applying the recommen- The views and preferences of the target population (patients, monitoring and/or auditing criteria', 2.2 (SD=1.2) for item 5: 'The guideline presents monitoring and/or auditing criteria'. 2.2 (SD=1.2) for item 5: 'The views and preferences of the target population (patients, public and so on) have been sought. 2.3 (SD=1.3) for item 20: ‘The potential resource implications of applying the recommendations have been considered.’ The highest individual score was 5.6 (SD=0.84) for item 1: ‘The overall objective(s) of the guideline is (are) specifically described’. Of the 34 CPGs, 8 were of high quality (table 3). High-quality CPGs were: four for osteoarthritis, from the European League Against Rheumatism, American Academy of Orthopedic Surgeons, Osteoarthritis Research Society International and NICE; 2 for low back pain, from NICE and the Council on Chiropractic Guidelines and Practice Parameters; 1 for neck pain, from the Ontario Protocol for Traffic Injury Management Collaboration; and 1 for shoulder pain, from the University of New South Wales. The two CPGs from NICE had scores greater than 70% in all domains, which was substantially higher than other CPGs.

### Appraisal of CPGs: inter-rater reliability
Inter-rater reliability was fair for scope and purpose (domain 1) and good for clarity of presentation (domain 4). Reliability was excellent for all other domains and overall AGREE II score (table 1).

### Appraisal of CPGs: quality
The mean overall score for all CPGs was 45.1% (SD=19.7) (table 2). Overall, the lowest domain score was for applicability (domain 5) with a mean score of 26.3% (SD=19.5). Editorial independence (domain 6) was the next lowest score with a mean of 32.5% (SD=27.5). The highest overall score was for scope and purpose (domain 1) with a mean of 72.4% (SD=14.3), and then clarity of presentation (domain 4) with a mean of 59.1% (SD=17.7).

The lowest mean scores for individual items (on a 1–7 scale) were 2.1 (SD=1.3) for item 21: ‘The guideline presents monitoring and/or auditing criteria’, 2.2 (SD=1.2) for item 5: ‘The views and preferences of the target population (patients, public, etc) have been sought’ and 2.3 (SD=1.3) for item 20: ‘The potential resource implications of applying the recommendations have been considered.’ The highest individual score was 5.6 (SD=0.84) for item 1: ‘The overall objective(s) of the guideline is (are) specifically described’.

### Appraisal of CPGs: inter-rater reliability
The mean overall score for all CPGs was 45.1% (SD=19.7) (table 2). Overall, the lowest domain score was for applicability (domain 5) with a mean score of 26.3% (SD=19.5). Editorial independence (domain 6) was the next lowest score with a mean of 32.5% (SD=27.5). The highest overall score was for scope and purpose (domain 1) with a mean of 72.4% (SD=14.3), and then clarity of presentation (domain 4) with a mean of 59.1% (SD=17.7).

The lowest mean scores for individual items (on a 1–7 scale) were 2.1 (SD=1.3) for item 21: ‘The guideline presents monitoring and/or auditing criteria’, 2.2 (SD=1.2) for item 5: ‘The views and preferences of the target population (patients, public, etc) have been sought’ and 2.3 (SD=1.3) for item 20: ‘The potential resource implications of applying the recommendations have been considered.’ The highest individual score was 5.6 (SD=0.84) for item 1: ‘The overall objective(s) of the guideline is (are) specifically described’.

Of the 34 CPGs, 8 were of high quality (table 3). High-quality CPGs were: four for osteoarthritis, from the European League Against Rheumatism, American Academy of Orthopedic Surgeons, Osteoarthritis Research Society International and NICE; 2 for low back pain, from NICE and the Council on Chiropractic Guidelines and Practice Parameters; 1 for neck pain, from the Ontario Protocol for Traffic Injury Management Collaboration; and 1 for shoulder pain, from the University of New South Wales. The two CPGs from NICE had scores greater than 70% in all domains, which was substantially higher than other CPGs.

### High-quality CPGs
High-quality CPGs were from the UK (n=2), USA (n=2), international (n=1), Europe (n=1), Australia (n=1) and Canada (n=1). The developing groups included medical societies (n=4), government bodies (n=2), an expert panel (n=1) and research collaboration (n=1). Five of the eight high-quality CPGs stated that the development group included members with expertise in CPG development, such as methodologists or representatives from CPG departments. In two high-quality CPGs that did not specifically include members with CPG expertise, clinical epidemiologists and/or health economists and library scientists were involved in guideline development. The University of New South Wales shoulder CPG engaged a CPG development
## Table 3  CPG AGREE II domain scores and quality assessment (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteoarthritis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAO – knee osteoarthritis</td>
<td>94</td>
<td>57</td>
<td>87</td>
<td>78</td>
<td>29</td>
<td>81</td>
<td>78</td>
<td>High</td>
</tr>
<tr>
<td>ACR – hand, hip and knee osteoarthritis</td>
<td>69</td>
<td>33</td>
<td>53</td>
<td>67</td>
<td>7</td>
<td>22</td>
<td>44</td>
<td>Low</td>
</tr>
<tr>
<td>EULAR – hip and knee osteoarthritis</td>
<td>80</td>
<td>67</td>
<td>56</td>
<td>76</td>
<td>19</td>
<td>50</td>
<td>61</td>
<td>High</td>
</tr>
<tr>
<td>ESCOE – knee osteoarthritis</td>
<td>65</td>
<td>26</td>
<td>37</td>
<td>61</td>
<td>21</td>
<td>67</td>
<td>44</td>
<td>Low</td>
</tr>
<tr>
<td>ESCOE – knee osteoarthritis with updated algorithm of pharmacological interventions</td>
<td>63</td>
<td>20</td>
<td>10</td>
<td>63</td>
<td>11</td>
<td>0</td>
<td>28</td>
<td>Low</td>
</tr>
<tr>
<td>MAESTRA – physiotherapy hip/knee osteoarthritis</td>
<td>78</td>
<td>63</td>
<td>55</td>
<td>83</td>
<td>46</td>
<td>11</td>
<td>61</td>
<td>Low</td>
</tr>
<tr>
<td>NICE – osteoarthritis</td>
<td>80</td>
<td>80</td>
<td>90</td>
<td>94</td>
<td>88</td>
<td>78</td>
<td>89</td>
<td>High</td>
</tr>
<tr>
<td>OARSI – knee osteoarthritis</td>
<td>74</td>
<td>63</td>
<td>55</td>
<td>70</td>
<td>4</td>
<td>78</td>
<td>67</td>
<td>High</td>
</tr>
<tr>
<td>Ottawa – osteoarthritis low back pain</td>
<td>89</td>
<td>56</td>
<td>58</td>
<td>50</td>
<td>28</td>
<td>28</td>
<td>56</td>
<td>Low</td>
</tr>
<tr>
<td>Ottawa – patient education osteoarthritis</td>
<td>72</td>
<td>54</td>
<td>53</td>
<td>39</td>
<td>13</td>
<td>6</td>
<td>33</td>
<td>Low</td>
</tr>
<tr>
<td>VA/DoD – hip and knee osteoarthritis</td>
<td>85</td>
<td>50</td>
<td>60</td>
<td>78</td>
<td>25</td>
<td>3</td>
<td>56</td>
<td>Low</td>
</tr>
<tr>
<td>Low back pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APTA – low back pain</td>
<td>78</td>
<td>50</td>
<td>45</td>
<td>67</td>
<td>29</td>
<td>3</td>
<td>44</td>
<td>Low</td>
</tr>
<tr>
<td>CCGPP – low back pain</td>
<td>67</td>
<td>54</td>
<td>60</td>
<td>39</td>
<td>25</td>
<td>61</td>
<td>44</td>
<td>High</td>
</tr>
<tr>
<td>BPS – low back pain/radiculopathy</td>
<td>56</td>
<td>48</td>
<td>8</td>
<td>57</td>
<td>21</td>
<td>25</td>
<td>17</td>
<td>Low</td>
</tr>
<tr>
<td>Cheng et al – low back pain</td>
<td>41</td>
<td>15</td>
<td>26</td>
<td>39</td>
<td>17</td>
<td>0</td>
<td>17</td>
<td>Low</td>
</tr>
<tr>
<td>Colorado DLE WC – low back pain</td>
<td>70</td>
<td>33</td>
<td>63</td>
<td>48</td>
<td>31</td>
<td>22</td>
<td>28</td>
<td>Low</td>
</tr>
<tr>
<td>ICSI – low back pain</td>
<td>78</td>
<td>46</td>
<td>41</td>
<td>72</td>
<td>51</td>
<td>72</td>
<td>56</td>
<td>Low</td>
</tr>
<tr>
<td>NASS – lumbar disc herniation with radiculopathy</td>
<td>72</td>
<td>33</td>
<td>51</td>
<td>48</td>
<td>10</td>
<td>33</td>
<td>39</td>
<td>Low</td>
</tr>
<tr>
<td>NASS – degenerative lumbar spinal stenosis</td>
<td>76</td>
<td>26</td>
<td>48</td>
<td>44</td>
<td>8</td>
<td>39</td>
<td>39</td>
<td>Low</td>
</tr>
<tr>
<td>NASS – degenerative lumbar spondylolisthesis</td>
<td>81</td>
<td>30</td>
<td>51</td>
<td>44</td>
<td>25</td>
<td>42</td>
<td>39</td>
<td>Low</td>
</tr>
<tr>
<td>NICE – low back pain and sciatica in over 16s</td>
<td>89</td>
<td>78</td>
<td>85</td>
<td>93</td>
<td>83</td>
<td>72</td>
<td>89</td>
<td>High</td>
</tr>
<tr>
<td>TOP – low back pain</td>
<td>72</td>
<td>31</td>
<td>17</td>
<td>74</td>
<td>19</td>
<td>0</td>
<td>33</td>
<td>Low</td>
</tr>
<tr>
<td>Shoulder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOA – subacromial pain</td>
<td>78</td>
<td>26</td>
<td>28</td>
<td>48</td>
<td>11</td>
<td>3</td>
<td>22</td>
<td>Low</td>
</tr>
<tr>
<td>Eubank et al – rotator cuff</td>
<td>74</td>
<td>37</td>
<td>33</td>
<td>46</td>
<td>25</td>
<td>22</td>
<td>39</td>
<td>Low</td>
</tr>
<tr>
<td>Hopman et al – rotator cuff</td>
<td>87</td>
<td>74</td>
<td>73</td>
<td>78</td>
<td>50</td>
<td>56</td>
<td>67</td>
<td>High</td>
</tr>
<tr>
<td>I.S.Mu.L.T – rotator cuff tears</td>
<td>43</td>
<td>17</td>
<td>19</td>
<td>35</td>
<td>10</td>
<td>0</td>
<td>28</td>
<td>Low</td>
</tr>
<tr>
<td>WSDL&amp;I – shoulder</td>
<td>74</td>
<td>44</td>
<td>20</td>
<td>37</td>
<td>24</td>
<td>8</td>
<td>22</td>
<td>Low</td>
</tr>
<tr>
<td>Neck</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCA – neck pain</td>
<td>63</td>
<td>37</td>
<td>55</td>
<td>50</td>
<td>10</td>
<td>31</td>
<td>44</td>
<td>Low</td>
</tr>
<tr>
<td>Colorado DLE WC – cervical spine injury</td>
<td>81</td>
<td>37</td>
<td>57</td>
<td>52</td>
<td>35</td>
<td>25</td>
<td>39</td>
<td>Low</td>
</tr>
<tr>
<td>OPTIMA – neck pain</td>
<td>94</td>
<td>67</td>
<td>80</td>
<td>91</td>
<td>38</td>
<td>50</td>
<td>67</td>
<td>High</td>
</tr>
<tr>
<td>SIMFER – neck pain</td>
<td>46</td>
<td>30</td>
<td>15</td>
<td>37</td>
<td>7</td>
<td>0</td>
<td>22</td>
<td>Low</td>
</tr>
<tr>
<td>Knee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barton et al – P-F pain</td>
<td>69</td>
<td>24</td>
<td>56</td>
<td>61</td>
<td>25</td>
<td>69</td>
<td>61</td>
<td>Low</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE – musculoskeletal injuries</td>
<td>37</td>
<td>28</td>
<td>17</td>
<td>39</td>
<td>11</td>
<td>9</td>
<td>17</td>
<td>Low</td>
</tr>
</tbody>
</table>

*First author given where there is no stated organisation.

AAOS, American Academy of Orthopedic Surgeons; ACR, American College of Rheumatology; AGREE, Appraisal of Guidelines for Research and Evaluation; APTA, Orthopaedic Section of the American Physical Therapy Association; BPS, British Pain Society; CCA, Canadian Chiropractic Association; CCGPP, Council on Chiropractic Guidelines and Practice Parameters; Colorado DLE WC, Colorado Department of Labour and Employment Workers Compensation; CPGs, clinical practice guidelines; DOA, Dutch Orthopaedic Association; ESCEO, European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis; EULAR, European League Against Rheumatism; I.S.Mu.L.T., Italian Society of Muscle Ligaments and Tendons; ICSI, Institute for Clinical Systems Improvement; KNGF, Royal Dutch Society of Physiotherapy; MAESTRA, Malaysia Health Technology Assessment Section Medical Development Division, Ministry of Health; NASS, North American Spine Society; NICE, National Institute for Health and Care Excellence; OARSI, Osteoarthritis Research Society International; OPTIMA, Ontario Protocol for Traffic Injury Management Collaboration; Ottawa, Ottawa Panel; P-F, patellofemoral; SIMFER, the Italian Society of Physical and Rehabilitation Medicine; TOP, Toward Optimized Practice; VA/DoD, Department of Veterans Affairs and the Department of Defence; WSDL&I, Washington State Department of Labour and Industries.
consultancy during the development process. Seven of the 26 low-quality CPGs included experts in CPG development.

Seven CPGs were informed by the AGREE II or AGREE instrument (one CPG also using a quality instrument from Australia’s NHMRC). Two high-quality, and five low-quality CPGs were informed by the AGREE II or AGREE instrument.

DISCUSSION

There is substantial variation in the quality of MSK pain CPGs. The overall quality of MSK pain CPGs is generally poor, with only eight out of 34 CPGs rated as high quality. This and other factors, including CPG replication, and inconsistencies in the way MSK pain conditions are defined, contribute to inefficiencies and wasted effort in CPG development. Further limitations were an unequal distribution of CPGs by conditions and country of development, and a lack of attention to aspects of the development process. Consolidation of CPG development efforts and greater attention to the development process is needed.

High-quality CPGs were from a range of countries and developed by diverse groups. The quality of CPGs may be improved by increasing international collaborations during development. However, only one of the high-quality CPGs included in our review was the result of international collaboration. Six of the eight high-quality CPGs involved dedicated CPG development expertise, supporting the value of including CPG methodologists within development teams.

Consistent with other work, we applied a cut-off threshold of 50% in three AGREE II domains to differentiate high-quality and low-quality CPGs. Even though our criteria were less stringent than others, a large proportion of CPGs were of poor quality. CPGs require considerable resources to develop. Expending resources on low-quality CPGs that have, based on the development processes, invalid care recommendations is wasteful and confusing to users. An additional issue is duplication of CPGs and that many of the duplicated CPGs are of poor quality. For example, there were 11 low back pain CPGs identified in our search, and 9 of these were judged as poor quality. Directing resources towards development of fewer, higher quality and less ‘redundant’ CPGs would be helpful in reducing inefficient resource use and user confusion. One recommendation is to increase collaboration in CPG development through networks such as the WHO, the G-I-N or the Cochrane Collaboration. An alternative is for smaller organisations with fewer resources and less development expertise to adopt or adapt existing high-quality CPGs to suit their needs.

Another problem was that CPG developers used inconsistent terminology to define MSK pain conditions. For example, three CPGs from the NASS were for: ‘degenerative lumbar spinal stenosis’, ‘degenerative lumbar spondylolisthesis’ and ‘lumbar disc herniation with radiculopathy’. Defining a CPG by structural ‘pathology’ is problematic because many changes, including spinal stenosis and spondylolisthesis, are common in asymptomatic individuals and poorly associated with pain and disability. Other higher quality low back pain CPGs classified these conditions as non-specific low back pain. Similarly, in shoulder pain, there were CPGs for ‘rotator cuff tears’, ‘rotator cuff syndrome’, ‘rotator cuff pathology’ and ‘sub acromial pain syndrome’. Consistent and contemporary terminology to define MSK pain conditions, irrespective of developer/professional group, is needed to reduce inefficiencies and CPG replication.

Most published CPGs are for osteoarthritis and then low back pain. Other common MSK conditions are underaddressed. Only four CPGs addressed neck pain, even though neck pain is the fourth leading cause of disability globally, higher than the burden attributable to hip and knee osteoarthritis (ranked 11th). There were no CPGs for thoracic spine pain. Thoracic spine pain has a point prevalence as high as 72% in young females and 1 month prevalence estimates of 15.8%–34.8% (depending on cohort age and pain definition). There was only one CPG for non-osteoarthritis-related knee pain.

Among the English-language CPGs we reviewed, there was an uneven distribution by geographical region. More than double the number of CPGs were developed in the USA compared with the next most common region (international collaborations). A high proportion of CPGs developed in the USA were of low quality. The high number of USA-developed CPGs may reflect the medicolegal healthcare environment in which CPGs are used to evaluate the performances of providers in malpractice suits, resource availability or healthcare priorities.

There is a lack of attention and/or reporting in aspects of CPG development. The main problems we identified were a lack of attention to guideline applicability (domain 5), limited involvement of patients/consumers in the development process (domain 2) and low editorial independence (domain 6). Limited attention to these areas is consistent with reviews of CPGs across a broad range of health conditions and is a fundamental issue in CPG development. Poor CPG applicability and lack of editorial independence is a consistent problem, despite improvements in the overall quality of CPGs across diverse areas of health.

Poor applicability is a barrier to the uptake of CPG recommendations into practice. All CPGs we reviewed were in written, ‘hard copy’ formats. CPG developers should consider newer, emerging methods to improve user uptake, awareness and ease of use. These include mobile technologies, for example, smartphone apps (summaries of NICE CPGs were available as apps), digital guideline platforms for rapid review and update of guidelines/recommendations and ‘living’ documents such as Wikis and other collaborative writing applications. In addition to increasing uptake, Wiki platforms have the potential to increase patient/consumer input as codevelopers and respond rapidly to new evidence as it becomes available.

There is also the opportunity to link MSK pain CPGs to healthcare quality initiatives targeting practitioners and consumers. This could include the Choosing Wisely campaign, or, in Australia, the Australian Commission on Safety and Quality in Health Care Atlas of Healthcare Variation.

Developers should consider the AGREE II criteria when developing and publishing CPGs, highlighted in particular by the low scores found for editorial independence (domain 6). Editorial independence is an important domain for CPG quality, and achieving a high score should be relatively straightforward as this only requires the inclusion of two statements. Poor scores for editorial independence could mean there are conflicts of interest/competing interests or, if this is not the case, has been reported in a way that does not enable high scoring against the AGREE II criteria. Despite the fact that seven CPGs were informed by the AGREE II or AGREE instruments, five of these were rated poor quality. Improved use of quality instruments, such as the AGREE II, during CPG development/reporting is needed.

The two CPGs that had the highest AGREE II scores were developed by NICE for osteoarthritis and low back pain and were the only CPGs with an overall score greater than 80%. Based on the quality of reporting the NICE CPGs should be favoured by healthcare clinicians, managers and policy makers.
What is already known?

- There is poor uptake/implementation of evidence in musculoskeletal pain care.
- High-quality clinical practice guidelines (CPGs) are an important vehicle to improve care.

What are the new findings?

- Most musculoskeletal pain CPGs are of poor quality; only 8 of 34 met high-quality standards.
- CPG developers should (1) focus on applicability to enhance uptake into care, (2) make better use of quality instruments and (3) involve CPG methodologists.

Limitations

The AGREE II reflects methodological processes, not necessarily content, and scores may reflect the quality of reporting rather than methodological quality. However, the AGREE II has been extensively validated and is a benchmark for assessing CPG quality. The next stage in our project is a content analysis of recommendations found in higher quality CPGs. In the current study, CPGs were appraised by three authors and ideally four should be used. In addition, two reviewers were academic physiotherapists (IL and RW), and the third was an indicator development researcher also with a background in physiotherapy (LW). Potentially, the appraisal may reflect the perspectives of reviewers. This potential limitation was addressed by including an interprofessional author group. As always, the search strategy may have failed to identify all relevant documents; however, for comprehensiveness, our search strategy was guided by a reference librarian. Only English language CPGs were reviewed, and high-quality non-English language CPGs may have been excluded. For practical reasons, one reviewer undertook initial screening of titles/abstracts; however, ideally, there are multiple reviewers.

SUMMARY AND RECOMMENDATIONS

The overall quality of MSK pain CPGs is poor. There is duplication of CPGs for osteoarthritis and low back pain, an under-representation for neck and knee pain and no CPGs for thoracic pain. MSK pain CPG developers should:

- First consider carefully if a new CPG is needed, or if existing high-quality CPGs could be adopted or adapted.
- Focus on under-addressed MSK pain conditions, such as neck and thoracic pain.
- Involve team members with methodological expertise.
- Use a CPG quality tool in the development and reporting processes, especially addressing applicability, the involvement of patients/consumers and editorial independence.
- Use contemporary, widely accepted terminology for MSK pain conditions.

MSK pain CPG users should:

- Be critical of current MSK pain CPG quality.
- Seek and use higher quality CPGs.

Acknowledgements

We acknowledge Anne Smith for statistical assistance.

Contributors

All authors were involved in the conception, design and interpretation of data. IL, LW and RW performed the data analysis and initial interpretation. IL was responsible for initial writing and drafting of the article which was reviewed by all authors. All authors revised critically for important intellectual content and approved the final version to be submitted.

Funding

IL is funded by an Australian National Health and Medical Research Council Early Career Fellowship (APP1090403). CGM's fellowship and major project is funded by Australia's National Health and Medical Research Council (APP1113022 and APP1113532). LW works on a project funded by a National Health and Medical Research Council Program Grant (APP1054146).

Competing interests

None declared.

Provenance and peer review

Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES
