

# Are Clinical Practice Guidelines of Low Back Pain interventions of high quality and updated? A systematic review using the AGREE II instrument

**Greta Castellini**

IRCCS Istituto Ortopedico Galeazzi

**Valerio Iannicelli**

Universita Vita Salute San Raffaele

**Matteo Briguglio**

IRCCS Istituto Ortopedico Galeazzi

**Davide Corbetta**

Universita Vita Salute San Raffaele

**Luca Maria Sconfienza**

IRCCS Istituto Ortopedico Galeazzi

**Giuseppe Banfi**

IRCCS Istituto Ortopedico Galeazzi

**Silvia Gianola** (✉ [silvia.gianola@grupposandonato.it](mailto:silvia.gianola@grupposandonato.it))

IRCCS Istituto Ortopedico Galeazzi <https://orcid.org/0000-0003-3770-0011>

---

## Research article

**Keywords:** Low Back Pain, Spine, Clinical Practice Guidelines, Systematic Review, Critical Appraisal, AGREE II, GRADE, Quality of Evidence

**Posted Date:** July 30th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-42082/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on October 22nd, 2020. See the published version at <https://doi.org/10.1186/s12913-020-05827-w>.

# Abstract

**BACKGROUND:** Clinical practice guidelines (CPGs) provide specific recommendations for practice but, due to the increasing number of CPGs developed by multiple organisations over the last years, there are concerns about their quality. The aim is to systematically appraise the CPGs quality for low back pain (LBP) interventions and explore the inter-rater reliability (IRR) among quality appraisers. Time span intended as time from systematic reviews searches to CPGs publication was also assessed.

**METHODS:** We undertook comprehensive searches in Pubmed, Embase, PEDro, TRIP, guideline organisation databases, websites and grey literature from January 2016 to March 2019 to identify all GPCs focus on rehabilitative, pharmacological or surgical interventions for LBP management. Four reviewers independently appraised the selected GPCs by using the Appraisal of Guidelines for Research and Evaluation II-AGREE II tool. Year of CPGs publication and year of search strategies were collected.

**RESULTS:** 21 CPGs met the inclusion criteria and were appraised. Seven (33%) had a broad scope involving surgery, rehabilitative or pharmacological interventions. The guidelines achieved the following scores for each AGREE II item: Editorial Independence (median 67%, interquartile range [IQR] 31 – 84%), Scope and Purpose (median 64%, IQR 22 – 83%), Rigour of Development (median 50%, IQR 21 – 72%), Clarity and Presentation (median 50%, IQR 28 – 79%), Stakeholder Involvement (median 36%, IQR 10 – 74%) and Applicability (median 11%, IQR 0 – 46%). The IRR among assessors was nearly perfect (ICC 0.90; 95% CI 0.88 – 0.91). We observed a median time span of 2 years (range 1-4) however, 38% of CPGs did not report the coverage dates for systematic searches.

**CONCLUSIONS:** We found methodological limitations affecting the CPGs quality. We call for a universal database where all guidelines can be registered and recommendations can be dynamically developed through a living systematic reviews approach ensuring the most updated evidence.

**LEVEL OF EVIDENCE: 1**

**REGISTRATION PROSPERO DETAILS:** CRD42019127619.

## Introduction

Low back pain (LBP) is one the greatest contributor to years lived with disability and it is the first cause of activity limitation as well as absence from work.(1) It is one of the commonest cause of seeking physician office visits with a substantial medical social and economic impact for individuals, families and society due to its high direct and indirect costs.(2) The global burden of this condition has led to the development of numerous guidelines by relevant medical societies or specialized working groups, providing recommendations about its diagnosis and management.(3, 4) Although the principles for developing CPGs are well established, the increased production of CPGs has been associated with a concern about their quality.(5) Several appraisals of guidelines on LBP already exist,(6– 10) but no one did a comprehensive search on both acute and chronic LBP looking at all the possible therapeutic

choices (i.e. rehabilitative, pharmacological or surgical). Furthermore, none considered the most recent published guidelines; as CPGs represent the bridge between relevant scientific literature and clinical decision making, their implementation into clinical practice should be indeed as much as possible updated.(11) It has been documented that 1 out of 5 recommendations included in the clinical guidelines are out of date after 3 years thus, recommendations validity over 3 years might be potentially unreliable. (12) As general rule, CPGs should be reviewed no later than 3 years after completion(13). As well, the National Institute of Clinical Excellence (NICE), the benchmark in guidelines production, declared that “A formal review of the need to update a guideline is usually undertaken by NICE 3 years after its publication”(14). Furthermore, this need is justified by the existing time span between the year of running the systematic search strategy during the production process and the year of publications in systematic reviews.(15) This time span is additionally extended considering that the production and dissemination of guidelines should be based on systematic reviews. Therefore, the selection of guidelines old more than three years would be a methodological bias resulting unethical in the clinical decision making process and mistaken in identifying high quality guidelines with not the most recent-update, available and reliable evidence.(16)

Based on this context, we aim to critically appraise the most recent update evidence-based CPGs for LBP interventions using the AGREE (Appraisal of Guidelines Research and Evaluation) II instrument, the gold standard for critical appraisal of guidelines.(17, 18) Furthermore, we investigate the inter-rater reliability of the AGREE II and the time span as coverage years between search strategy launch and publication date of guidelines.

## Materials And Methods

The reporting of this systematic review fulfils the Preferred Reporting Items for Systematic Reviews and Meta-Analyses.(19, 20) The PRISMA checklist is in **Supplement Digital Content 1**.

No ethics committee approval is needed. The protocol is registered on PROSPERO database (CRD42019127619).

## Inclusion and exclusion criteria

According to the World Health Organization, we defined a CPG as a document containing “systematically developed evidence-based statements that assist providers, patients, policy makers and other stakeholders to make informed decisions on health care and public health policy”.(21)

We included a CPG if: (i) the systematic process evaluated the recommendations; (ii) it was focused on rehabilitation, pharmacological or surgical therapeutic interventions for LBP management; (iii) the full-text was published in the last four years (2016–2020). We considered the most up-to-date version and its supplementary documents. No language restriction.

CPGs were excluded whether: (i) not primarily focused on LBP, such as national/international guidelines in which LBP was briefly mentioned in the context of a more comprehensive disease evaluation; (ii) not issued by national and international societies (e.g., designed for local use); (iii) declaration of recommendations was based exclusively on consensus statement or systematic reviews or commentary editorials related to published CPGs; (iv) focusing on interventions other than therapeutic (e.g., prevention, diagnosis); (v) based on population subgroups (e.g. pregnancy), specific causes (e.g. spondyloarthritis) or mixed/generic population (e.g. musculoskeletal chronic pain).

## Information sources and search strategy

We systematically searched Pubmed, Embase, PEDro and TRIP databases using the adapted terms and keywords derived from scoping search outlined in the search strategy. We also checked guideline organisation databases (e.g., National Institute for Clinical Excellence) and guideline websites (e.g., eGuidelines). **Supplement Digital Content 2** shows the whole search strategy. Two reviewers with a solid background in clinical epidemiology ran the search strategy in March 2019 and updated in January 2020. Furthermore, we looked for grey literature through Google Scholar and we screened the reference list for further eligible CPGs.

## Clinical practice guidelines selection

Search results were uploaded to Endnote software and duplicates were removed.<sup>(22)</sup><sup>(23)</sup> Independently, two reviewers screened titles and abstracts on the basis of the eligibility criteria. Full-texts were retrieved for abstracts with insufficient information or in case of disagreement between the two reviewers. When disagreement persisted, a third reviewer was consulted. We used Rayyan software (<https://rayyan.qcri.org/>) to manage screening and selection<sup>(24)</sup>. We reported reasons for study exclusion.

## Clinical practice guidelines appraisal

Four researchers (MB, GC, SG, VI) independently appraised each included CPG with the AGREE II instrument and recorded with a self-chronometer the time needed for each assessment. Researchers received training in the use of AGREE II. The appraisers completed the first global rating item on a 7-point scale (1 = lowest possible quality, 7 = highest possible quality) and the second global rating item to recommend guidelines for use in practice, with the three options of 'yes', 'yes, with modifications', and 'no'. One author (VI) calculated the standardised domain scores for each of the six domains as recommended by AGREE II.<sup>(17, 25)</sup> We then collected the following general data from each CPG: i) authors and year of publication; ii) ex novo, update or adoption/adolopment CPG status; iii) continent of origin; iv) organization/society/association, funding source, conflict of interest. We also extracted content information such as target population, target interventions (ie. surgery, physical therapy, pharmaceuticals, educational / behavioural, alternative medicine), rating methods for the quality of evidence (e.g., the Grading of Recommendations Assessment, Development and Evaluation - GRADE), presence of a multidisciplinary panel, and patient involvement.

## Data synthesis

We used descriptive statistics to summarize the characteristics of CPGs deemed eligible for inclusion. Data were summarized as frequency number (percentage) or median and interquartile range (IQR). We calculated a quality score for each of the six domains of CPGs by using the formula presented in the AGREE II User's Manual.(25) In addition, the appraisers added notes and completed the two global rating items at the end of each AGREE II assessment. The first global rating item asked appraisers to rate the overall quality of the guideline on a 7-point scale (1 = lowest possible quality and 7 = highest possible quality). Domain scores were calculated by summing up all the appraisers' scores of the individual items in a domain and by scaling the total as a percentage of the maximum possible score for that domain, automatically generated on the platform My AGREE PLUS. (26)

The second global rating item asked whether the appraiser would recommend the guideline for use in practice, with options of 'Yes', 'Yes, with modifications', and 'No'. Agreement among four appraisers was measured by using the intra-class correlation coefficient (ICC) with 95% Confidence Interval (CI). We judged the degree of agreement according to Landis and Koch(27): slight 0.01 to 0.2; fair 0.21 to 0.4; moderate 0.41 to 0.6; substantial 0.61 to 0.8; and almost perfect 0.81 to 1. We referred to a P value < 0.05 for a statistical significance. All tests were two-sided.(27) All data analyses were performed using STATA version 15.

## Results

### Search results

The systematic search generated 2502 citations; further 30 citations have been retrieved from the grey literature. A total of 70 CPGs and related documents underwent full-text screening, 25 of which met the inclusion criteria. Four of them are in awaiting assessment (Fig. 1). Finally, we appraised 21 CPGs with AGREE II (**Supplement Digital Content 3**).

[Figure 1]

### Characteristics of CPGs

**Table 1** reported main characteristics of the 21 CPGs. Ten CPGs addressed multiple interventions (47.6%). Rating of evidence quality was planned in 76% of guidelines and reported in 67%. More than half of the CPGs (52%) had a multidisciplinary panel and less than half (38%) reported patient involvement (**Supplement Digital Content 4**).

### [Table 1]

**Table 1.** Characteristics of included CPGs

### AGREE II domains assessment

Overall, the highest rating AGREE II domain was Editorial Independence (median 67%, interquartile range [IQR] 31–84%) followed by Scope and Purpose (median 64%, IQR 22–83%), Rigour of Development (median 50%, IQR 21–72%), Clarity and Presentation (median 50%, IQR 28–79%), Stakeholder Involvement (median 36.1%, IQR 10–74%), and Applicability (median 11%, IQR 0–46%). In the overall guideline assessment, the median of the overall quality item was 42% (IQR 15–67%) and the most frequent recommendation regarding the use of the guideline was “No” (Table 2).

NICE guideline(28) shown the highest quality (96%) for Educational/behavioural, physical therapy, pharmaceuticals interventions as well, the KCE (83%) guideline(29) cover the same interventions plus surgery with a minimum time span (1 and 2 years, respectively) for searching evidence (Supplement Digital Content 4).

## [Table 2]

Table 2. Overall domain assessment of included CPGs.

## Inter-rater reliability and time for AGREE II appraisal

Inter-rater agreement was nearly perfect (ICC 0.90; 95% CI 0.88–0.91). Guidelines appraisal took an average of 42 minutes to complete (95% CI 35–50).

## Time span in publication

Overall, the 38.1% did not report the dates of systematic search strategy, whereas less than half of CPGs (47.6%) reported a median time of 2 years (IQR 1–4) of time span. Only a half of them provided a search within one year from the guideline publication (Table 1).

## Discussion

This article reports the results of the quality appraisal with the AGREE II of the most recent CPGs for LBP interventions retrieved through a systematic search through electronic databases and guidelines websites and published from January 2016 to January 2020.

A key findings of this research was the quality variability of LBP CPGs across all six AGREE II domains; the *Domain 6 - Editorial Independence* and the *Domain 1 - Scope and Purpose* obtained the highest average scores (> 60%) whereas the *Domain 5 - Applicability* obtained the lowest (< 15%).

The overall quality in LBP CPGs was judged to be low and the most frequent judgement for the guideline recommendation was “No” (n = 15 out of 21).

The *Domain 1 - Scope and Purpose* addresses the overall aim of the guideline, the clinical question, and the target population. The highest compliance of our sample with this item (63.9%) could be mainly due to their focus on LBP, which is the most prevalent musculoskeletal condition for which guidelines should be needed considering the consequent years lived with disability in most countries.(30)

The *Domain 2 - Stakeholder Involvement* addresses the degree to which the guideline represents the views of its intended users. Less than one third of our CPGs accomplish the AGREE II requirements lacking the participation of patients and their advocate.

The *Domain 3 - Rigour of Development* revealed an extreme variability across CPGs scoring from 0 to 90.1%. Only an half of CPGs have an acceptable rigour of development, a low score of this domain is worrying, as *Domain 3* has been identified as stronger predictor of quality of the AGREE instrument.(5) Indeed, a regression analysis showed a statistically significant strongest influence of *Domain 3* on overall guideline quality.(31) Among the items addressed by the *Domain 3*, the one regarding the systematic search might be considered of great importance (i.e., “Item 7: Systematic methods were used to search for evidence”) because CPGs have the duty to present the most updated evidence. We found that less than half of CPGs did not reported the time coverage of systematic search or, when reported, it ranged from 1 to 4 years before CPGs publication. Another important item in *Domain 3* (e.g., “Item 12: There is an explicit link between the recommendations and the supporting evidence”) found that two-third of CPGs in our sample adequately planned and judged the body of the evidence (e.g, GRADE). However, the application of system for grading the evidence (i.e., GRADE) not always guarantee the inclusion of the most updated evidence in an acceptable time span: thus, reliability should be cautiously evaluated.

The validity of each recommendation, and consequently of the CPG, is actually determined by the methodological quality and transparency of its development and by the “living evidence” on which it is based. As suggested by Garcia et al. waiting more than 3 years to review a guideline is potentially too long and, in this case, recommendations could be outdated even at the time of guideline publication. (12) This critical issue has been answered by the living clinical practice guidelines concept (32), which draws inspiration from the already established model of living systematic reviews, where the evidence is continuously updated and incorporated as soon as available in the literature, through a process of continuous surveillance. (33)

On this view, the AGREE II should put emphasis on timing, rating a high-quality CPGs if they conducted the search within 2 years of completion of the review.(34)

The *Domain 4 - Clarity of Presentation* reflects the adequacy in the reporting of recommendations and different options for management and it was satisfactory in only half of our CPGs. This can be related to the purpose of AGREE II: the current version makes no distinction between quality of reporting and quality of conduct of a CPG. Despite good reporting, the methodological conduct underlying a guideline can be weak.(35) Quality of conduct and reporting should be judged separately as for all other study design.(36, 37) For instance, in systematic reviews, the PRISMA and the AMSTAR assess quality of reporting and quality of conduct, respectively. (38)

The *Domain 5 - Applicability* was the poorest scoring domain reporting results similar to other conditions. (5, 8, 39–41) Development and implementation of guidelines are erroneously considered as separate activities.(5)

The *Domain 6 - Editorial Independence*, has high compliance in most of the CPGs. Considering the high social-economic global burden of low back pain and the relative need for care, CPGs must reported the presence and management conflict of interests.

## Strength, limitations

Our appraisal has several strengths. We performed an exhaustive search including explicit eligibility criteria and independent duplicate assessment of eligibility. We involved four reviewers for the appraisal with the inter-rater reliability nearly perfect. While all appraisers were trained in the use of AGREE II, it should be acknowledged that the appraisers shared a similar background (methodology and rehabilitation), which may partially explain the highest overall agreement. Indeed, our team included clinical experts and methodologists with wide experience in clinical epidemiology including systematic reviews and CPGs. Even after the same training, guideline appraisers from different disciplines may interpret the items and scoring system differently.(42) Furthermore, it is possible that the appraisers, basing the assessment on their own experience, paid more attention to assessing the quality of reporting than the quality of conduct and vice versa. We analysed a reliable subset of CPGs restricted to LBP in order to ensure consistency of appraisal while avoiding discrepancies in item judgements due to different clinical contexts (e.g., AGREE II assessed CPGs in oncology differently from orthopaedics). We focused on the most recent versions of the guidelines in order to offer stakeholders, policy makers, clinicians and patients with the last evidence about effectiveness of interventions. However, selection of CPGs was challenging, since the definition of guidelines is not universally established. In the literature, there is confusion between what is the meaning of consensus and what is an evidence-based CPG. The rigour of methods and panel of experts have to be simultaneously considered in a CPG, but the current definition does not explicate these elements.

## Future spin for research

At the time of its publication, a CPG can be already old not reflecting the most recent evidence. Indeed, time can influence the reliability at two points: (a) during the conduction of systematic reviews for the production of the body of the evidence needed during CPG development; (b) between the finalization of a CPG and its publication. In order to avoid waste of efforts resulting in CPGs duplicate or unreliable new-born already old CPG, we call for a universal database where all guidelines can be registered and updated. An example might be the registers of RCTs (e.g., WHO or clinicaltrials.gov) and SRs (e.g., PROSPERO) but for CPGs. In this way, a “living and dynamic” development of recommendations can be better recognized identifying the most recent literature. (43)

## Abbreviations

A Measurement Tool to Assess systematic Reviews (AMSTAR); Appraisal of Guidelines for Research and Evaluation II (AGREE II); Confidence Interval (CI); Clinical Practice Guidelines (CPGs); Grading of Recommendations Assessment, Development and Evaluation (GRADE); Inter-Quartile Range (IQR); Intra-Class Correlation (ICC); Low Back Pain (LBP); National Institute of Clinical Excellence (NICE);

Physiotherapy Evidence Database (PEDro); Preferred Reporting Intervention for Systematic Review and Meta-analysis (PRISMA); Turning Research into Practice (TRIP); World Health Organization (WHO).

## Declarations

**Ethics approval and consent to participate:** Not applicable.

**Consent for publication:** Not applicable.

**Competing interests:** The authors declare no competing interests.

**Funding:** The work was supported by the Italian Ministry of Health “Linea 3 – Valutazione della qualità delle attuali linee guida in ortopedia e in riabilitazione” L3042. The funding sources had no controlling role in the study design, data collection, analysis, interpretation or report writing.

**Availability of data and materials:** All data generated or analysed during this study are included in this published article with all Additional materials. Raw data are stored at the following link:

[https://osf.io/xwbu2/?view\\_only=d3aa81b467874b468bd1207d96df7376](https://osf.io/xwbu2/?view_only=d3aa81b467874b468bd1207d96df7376)

### Authors' Contributions:

Concept development providing idea for the research: SG, CG

Design planning the methods to generate the results: SG, CG, VI

Supervision: SG, GC

Data collection: SG, CG, VI, MB

Analysis and interpretation (statistics, evaluation and presentation of the results): VI, SG, GC, DC

Writing: VI, SG, GC

Critical review (revised manuscript for intellectual content; this does not relate to spelling and grammar checking): DC, GB, LMS

## References

1. Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Annals of the rheumatic diseases*. 2014;73(6):968-74.
2. Vrbancic TS. [Low back pain—from definition to diagnosis]. *Reumatizam*. 2011;58(2):105-7.
3. O'Connell NE, Cook CE, Wand BM, Ward SP. Clinical guidelines for low back pain: A critical review of consensus and inconsistencies across three major guidelines. *Best practice & research Clinical*

- rheumatology. 2016;30(6):968-80.
4. O'Sullivan K, O'Keefe M, O'Sullivan P. NICE low back pain guidelines: opportunities and obstacles to change practice. *British journal of sports medicine*. 2017;51(22):1632-3.
  5. Alonso-Coello P, Irfan A, Sola I, Gich I, Delgado-Noguera M, Rigau D, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Quality & safety in health care*. 2010;19(6):e58.
  6. van Tulder MW, Tuut M, Pennick V, Bombardier C, Assendelft WJ. Quality of primary care guidelines for acute low back pain. *Spine*. 2004;29(17):E357-62.
  7. Bouwmeester W, van Enst A, van Tulder M. Quality of low back pain guidelines improved. *Spine*. 2009;34(23):2562-7.
  8. Doniselli FM, Zanardo M, Manfre L, Papini GDE, Rovira A, Sardanelli F, et al. A critical appraisal of the quality of low back pain practice guidelines using the AGREE II tool and comparison with previous evaluations: a EuroAIM initiative. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*. 2018;27(11):2781-90.
  9. Dagenais S, Tricco AC, Haldeman S. Synthesis of recommendations for the assessment and management of low back pain from recent clinical practice guidelines. *The spine journal : official journal of the North American Spine Society*. 2010;10(6):514-29.
  10. Meroni R, Piscitelli D, Ravasio C, Vanti C, Bertozzi L, De Vito G, et al. Evidence for managing chronic low back pain in primary care: a review of recommendations from high-quality clinical practice guidelines. *Disability and rehabilitation*. 2019:1-15.
  11. Gurgel RK. Updating Clinical Practice Guidelines: How Do We Stay Current? *Otolaryngology–head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery*. 2015;153(4):488-90.
  12. Martinez Garcia L, Sanabria AJ, Garcia Alvarez E, Trujillo-Martin MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2014;186(16):1211-9.
  13. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? *Jama*. 2001;286(12):1461-7.
  14. (NICE) TNIfHaCE. Developing NICE guidelines: the manual. Process and methods, Published: 31 October 2014, niceorguk/process/pmg20

©. 2014.

15. Yoshii A, Plaut DA, McGraw KA, Anderson MJ, Wellik KE. Analysis of the reporting of search strategies in Cochrane systematic reviews. *Journal of the Medical Library Association : JMLA*. 2009;97(1):21-9.

16. Moher D. Reporting guidelines: doing better for readers. *BMC medicine*. 2018;16(1):233.
17. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2010;182(18):E839-42.
18. Brouwers MC, Kerkvliet K, Spithoff K, Consortium ANS. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *Bmj*. 2016;352:i1152.
19. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*. 2009;151(4):264-9, W64.
20. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj*. 2009;339:b2700.
21. WHO WHO. WHO handbook for guideline development, 2nd ed. World Health Organization. <http://www.who.int/iris/handle/10665/145714>. 2014.
22. Eapen BR. EndNote 7.0. *Indian journal of dermatology, venereology and leprology*. 2006;72(2):165-6.
23. Bramer WM, Milic J, Mast F. Reviewing retrieved references for inclusion in systematic reviews using EndNote. *Journal of the Medical Library Association : JMLA*. 2017;105(1):84-7.
24. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Systematic reviews*. 2016;5(1):210.
25. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. Development of the AGREE II, part 2: assessment of validity of items and tools to support application. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2010;182(10):E472-8.
26. Makarski J, Brouwers MC, Enterprise A. The AGREE Enterprise: a decade of advancing clinical practice guidelines. *Implementation science : IS*. 2014;9:103.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
28. Low Back Pain and Sciatica in Over 16s: Assessment and Management. National Institute for Health and Care Excellence: Clinical Guidelines. London2016.
29. Van Wambeke P, Desomer A, Ailliet L, Berquin A, Demoulin C, Depreitere B, et al. Low back pain and radicular pain: assessment and management. Good Clinical Practice (GCP) Brussels: Belgian Health Care Knowledge Centre (KCE). 2017. KCE Reports 287. D/2017/10.273/36.
30. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1789-858.
31. Hoffmann-Esser W, Siering U, Neugebauer EA, Brockhaus AC, Lampert U, Eikermann M. Guideline appraisal with AGREE II: Systematic review of the current evidence on how users handle the 2 overall assessments. *PloS one*. 2017;12(3):e0174831.

32. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Annals of internal medicine*. 2007;147(4):224-33.
33. Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, et al. Living systematic review: 1. Introduction-the why, what, when, and how. *J Clin Epidemiol*. 2017;91:23-30.
34. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *Bmj-Brit Med J*. 2017;358.
35. Jarl G, Hellstrand Tang U, Norden E, Johannesson A, Rusaw DF. Nordic clinical guidelines for orthotic treatment of osteoarthritis of the knee: A systematic review using the AGREE II instrument. *Prosthetics and orthotics international*. 2019:309364619857854.
36. Chen Y, Yang K, Marusic A, Qaseem A, Meerpohl JJ, Flottorp S, et al. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. *Annals of internal medicine*. 2017;166(2):128-32.
37. Huwiler-Muntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. *Jama*. 2002;287(21):2801-4.
38. Pussegoda K, Turner L, Garritty C, Mayhew A, Skidmore B, Stevens A, et al. Identifying approaches for assessing methodological and reporting quality of systematic reviews: a descriptive study. *Systematic reviews*. 2017;6(1):117.
39. Acuna SA, Huang JW, Scott AL, Micic S, Daly C, Brezden-Masley C, et al. Cancer Screening Recommendations for Solid Organ Transplant Recipients: A Systematic Review of Clinical Practice Guidelines. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2017;17(1):103-14.
40. Tong A, Chapman JR, Wong G, de Bruijn J, Craig JC. Screening and follow-up of living kidney donors: a systematic review of clinical practice guidelines. *Transplantation*. 2011;92(9):962-72.
41. Acuna-Izcaray A, Sanchez-Angarita E, Plaza V, Rodrigo G, de Oca MM, Gich I, et al. Quality assessment of asthma clinical practice guidelines: a systematic appraisal. *Chest*. 2013;144(2):390-7.
42. Marciano NJ, Merlin TL, Bessen T, Street JM. To what extent are current guidelines for cutaneous melanoma follow up based on scientific evidence? *International journal of clinical practice*. 2014;68(6):761-70.
43. Akl EA, Meerpohl JJ, Elliott J, Kahale LA, Schunemann HJ, Living Systematic Review N. Living systematic reviews: 4. Living guideline recommendations. *J Clin Epidemiol*. 2017;91:47-53.

## Tables

**Table 1.** Characteristics of included CPGs

REF	Publication year	Country	Status	Topic	Publication dates of systematic search strategy
ACP	2017	USA	Update	Educational / behavioural, physical therapy, pharmaceuticals	2008 - 2015
AIM	2019	USA	Update	Surgery	Not reported
AOA	2016	USA	Update	Physical therapy	2003 - 2014
ASIPP	2019	USA	New	Pharmaceuticals	Not reported
BMA	2018	Brazil	New	Educational / behavioural	Not reported
CAMM	2016	China	New	Alternative medicine	Not reported
CCGI	2018	Canada	New	Educational / behavioural, physical therapy	2015 - 2017
CCGPP	2016	USA	Adoption	Physical therapy	2009 - 2014
CPLA	2018	Latin America	Adoption	Physical therapy, pharmaceuticals	2004 - 2014
DSA	2016	Netherlands	New	Surgery	1990 - 2011
GSCI	2018	International	New	Surgery	Not reported
ICSI	2018	USA	Update	Educational / behavioural, physical therapy, pharmaceuticals	2000 - 2017
KCE	2017	Belgium	Adoption	Educational / behavioural, physical therapy, pharmaceuticals, surgery	2010 - 2015
KIOM	2017	Korea	New	Physical therapy, pharmaceuticals, alternative medicine	start date not reported - 2015
KSSS	2017	Korea	Adoption	Educational / behavioural, physical therapy, pharmaceuticals	2000 - 2016
L&I	2016	USA	Update	Surgery	Not reported
NICE	2016	UK	Update	Educational / behavioural, physical therapy, pharmaceuticals	2013 - 2015
PSP	2017	Poland	New	Physical therapy	Not reported
PSSS	2016	Poland	New	Surgery	Not reported
TOP	2017	USA	Update	Educational / behavioural, physical therapy, pharmaceuticals, surgery	2010 - 2014
VADoD	2017	USA	Update	Educational / behavioural, physical therapy, pharmaceuticals	2006 - 2016

For full CPG's reference lists and acronyms see *Supplement Digital Content 3*.

Table 2. Overall domain assessment of included CPGs.

Clinical Practice Guideline	Scope and Purpose	Stakeholder Involvement	Rigour of Development	Clarity of Presentation	Applicability	Editorial Independence	OVERALL		
							Personal rating	I would recommend? (MODE)	
NICE	2016	98,61%	95,83%	90,10%	100,00%	65,63%	93,75%	95,83%	Yes
CCGI	2018	90,28%	84,72%	88,54%	90,28%	73,96%	85,42%	87,50%	Yes
KCE	2017	88,89%	77,78%	83,85%	69,44%	86,46%	95,83%	83,33%	Yes
ACP	2017	97,22%	75,00%	76,56%	80,56%	22,92%	93,75%	75,00%	Yes
VADoD	2017	84,72%	90,28%	80,21%	87,50%	41,67%	60,42%	70,83%	Yes, with mod.
ICSI	2018	80,56%	72,22%	68,23%	86,11%	51,04%	79,17%	62,50%	Yes, with mod.
TOP	2017	69,44%	48,61%	66,15%	77,78%	71,88%	75,00%	58,33%	No
KIOM	2017	72,22%	40,28%	47,92%	59,72%	1,04%	83,33%	45,83%	No
CAMM	2016	63,89%	44,44%	35,94%	63,89%	0,00%	29,17%	45,83%	No
GSCI	2018	52,78%	11,11%	25,00%	70,83%	11,46%	75,00%	41,67%	No
ASIPP	2019	66,67%	25,00%	57,29%	33,33%	6,25%	81,25%	41,67%	No
AOA	2016	76,39%	36,11%	52,60%	25,00%	33,33%	58,33%	41,67%	No
CPLA	2018	26,39%	13,89%	27,60%	27,78%	23,96%	93,75%	33,33%	No
CCGPP	2016	43,06%	33,33%	53,13%	47,22%	16,67%	66,67%	29,17%	No
DSA	2016	63,89%	56,94%	50,52%	50,00%	4,17%	45,83%	29,17%	No
BMA	2018	22,22%	0,00%	34,90%	29,17%	0,00%	29,17%	20,83%	No
KSSS	2017	15,28%	2,78%	17,71%	20,83%	0,00%	29,17%	8,33%	No
PSSS	2016	20,83%	2,78%	12,50%	29,17%	0,00%	62,50%	8,33%	No
AIM	2019	18,06%	8,33%	6,77%	27,78%	0,00%	0,00%	4,17%	No
PSP	2017	22,22%	0,00%	1,56%	19,44%	0,00%	33,33%	4,17%	No
L&I	2016	12,50%	13,89%	0,00%	29,17%	0,00%	0,00%	0,00%	No

For full CPG's reference lists and acronyms see **Supplement Digital Content**

## List Of Supplemental Digital Content

**Supplement Digital Content 1.** Prisma checklist

**Supplement Digital Content 2.** Literature search strategy

**Supplement Digital Content 3.** List of included CPGs appraised with AGREE II

**Supplement Digital Content 4.** Additional Characteristics of included CPGs

## Figures

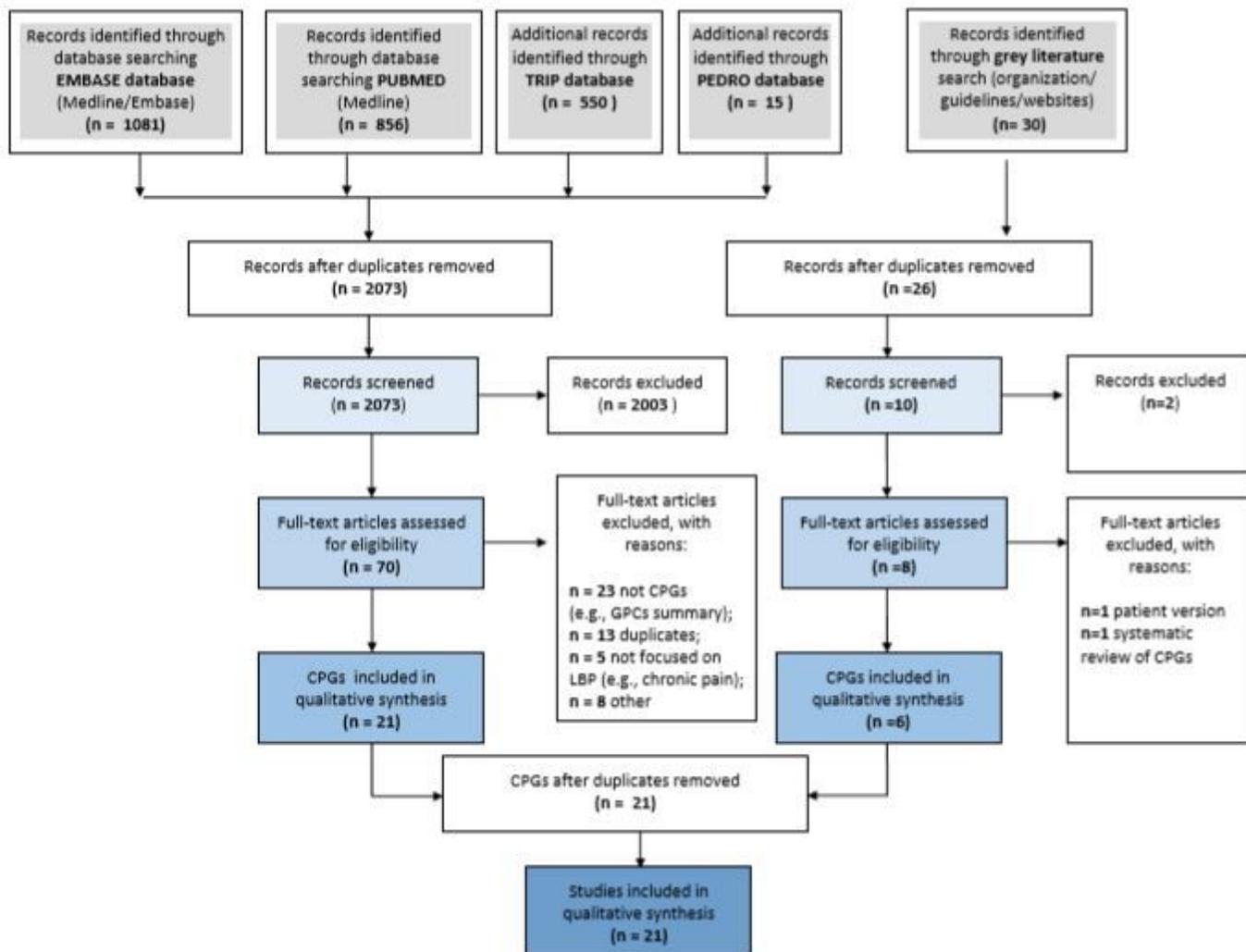


Figure 1

Flow diagram of CPG selection process

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarydigitalcontent3AdditionalGPCscharac.xlsx](#)
- [SupplementDigitalContent2ListofincludedCPGs.docx](#)
- [Supplementaldigitalcontent1LiteratureSearchstrategy.docx](#)
- [SupplementaryPRISMA2009checklist.doc](#)