How good are Low Back Pain guidelines? A critical appraisal of the quality of clinical practice guidelines using the AGREE II tool

Valerio Iannicelli [1], Greta Castellini [2,3], Matteo Briguglio [3], Davide Corbetta [1,4], Luca Maria Sconfienza[2,3], Giuseppe Banfi[1,3], Silvia Gianola[3]

1 Vita-Salute San Raffaele University, Milan, Italy
2 University of Milan, Milan, Italy
3 IRCCS Istituto Ortopedico Galeazzi, Milan, Italy
4 IRCCS San Raffaele Hospital, Milan, Italy

Background

Clinical practice guidelines (CPGs) provide to clinicians specific recommendations for practice, but there are concerns about their quality due to an increasing number in the last years. We aimed to systematically assess the methodological quality of guidelines for the management of low back pain (LBP).

Methods

Comprehensive searches in Pubmed and Embase, TRIP database, National/International databases and guidelines websites were undertaken up to 23 January 2019 to identify all CPGs assessing recommendations on rehabilitation, pharmacological or surgical therapeutic interventions for LBP management. In order to guarantee the most update evidence coming from CPGs, only those published since 2016 have been included. Two authors independently screened and selected the records according to the eligibility criteria. The primary outcome was the evaluation of quality of guidelines reporting. Four reviewers independently evaluated eligible guidelines by using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) assessment tool. According to the AGREE II User's Manual, the six domains of the instrument are independent and calculated according to AGREE's formula. Domain scores were categorized as good (≥ 80%), acceptable (60–79%), low (40–59%), or very low (< 40%). Finally, we defined high quality when all the domains scored > 60%, average quality when 3, 4 or 5 domains scored > 60%, and low quality when ≤ 2 domains scored > 60%. The secondary outcome was the inter-rater reliability between raters calculated using the intraclass correlation coefficient.

Results

The systematic search generated 1195 database citations and 30 grey literature links. A total of 56 guidelines and related documents underwent full-text screening, of which 11 met the inclusion criteria for AGREE II assessment. Intraobserver agreement was good (overall intraclass correlation coefficient of 0.882, 95% confidence interval 0.857–0.904). Overall, the highest rating AGREE II domains were Editorial Independence' (mean 75,4%, range 33,3 – 95,8%), 'Scope and Purpose' (mean 69,8%, range 20,8 – 98,6%) and 'Clarity and Presentation' (mean 67,9%, range 19,4 – 100%). The lowest rating domains were 'Rigour of Development' (mean 60,8%, range 1,96 – 90,1%), 'Stakeholder Involvement' (mean 56,4%, range 0 – 95,8%) and 'Applicability' (mean 39,2%, range 0 – 86,5%). The National Institute for Health and Care Excellence guideline performed best with respect to AGREE II criteria; only two other CPGs warranted high scores on all domains. All other guidelines received scores of under 60% on one or more domains (Table 1).

Conclusions

In the overall evaluation, three guidelines were considered 'high quality', four 'medium quality', and four 'low quality'. The results of this review can help researchers and Italian policymakers to select and adopt the highest quality GPCs for LBP management in the GPCs national system (Sistema Nazionale Linee Guida - SNLG).

Unità di Epidemiologia Clinica
IRCCS Istituto Ortopedico Galeazzi
greta.castellini@grupposandonato.it;
silvia.gianola@grupposandonato.it

Table 1. Quality assessment of all included CPGs.

	REF	Scope and Purpose	Stakeholder Involvement	Rigour of Development	Clarity of Presentation	Applicability	Editorial Independence	OVERALL		
								Personal rating	I would recommend? (MODE)	Global assessment
1	Bussières et al.	90,3%	84,7%	88,5%	90,3%	74,0%	85,4%	87,5%	Yes	High quality
2	ICSC	80,6%	72,2%	68,2%	86,1%	51,0%	79,2%	62,5%	Yes, with mod.	Medium quality
3	VA/DoD	84,7%	90,3%	80,2%	87,5%	41,7%	60,4%	70,8%	Yes, with mod.	Medium quality
4	Hee Jun et al.	72,2%	40,3%	47,9%	59,7%	1,0%	83,3%	45,8%	No	Low quality
5	Kassolik et al.	22,2%	0,0%	1,6%	19,4%	0,0%	33,3%	4,2%	No	Low quality
6	Alberta	69,4%	48,6%	66,1%	77,8%	71,9%	75,0%	58,3%	No	Medium quality
7	Qaseem et al.	97,2%	75,0%	76,6%	80,6%	22,9%	93,8%	75,0%	Yes	Medium quality
8	KCE	88,9%	77,8%	83,9%	69,4%	86,5%	95,8%	83,3%	Yes	High quality
9	Latka et al.	20,8%	2,8%	12,5%	29,2%	0,0%	62,5%	8,3%	No	Low quality
10	Globe et al.	43,1%	33,3%	53,1%	47,2%	16,7%	66,7%	29,2%	No	Low quality
11	NICE	98,6%	95,8%	90,1%	100,0%	65,6%	93,8%	95,8%	Yes	High quality