Brazilian Journal of Physical Therapy 27 (2023) 100563

Brazilian Journal of Physical Therapy

https://www.journals.elsevier.com/brazilian-journal-of-physical-therapy

SHORT COMMUNICATION

Compared to what? An analysis of comparators in trials informing the National Institute of Clinical Excellence (NICE) low back pain guideline

Geronimo Bejarano[a,*], **Ben Csiernik**[b], **Joshua R. Zadro**[c], **Giovanni E. Ferreira**[c]

[a] *Brown University, Providence, RI, United States*
[b] *Ontario Tech University, Oshawa, Ontario, Canada*
[c] *Institute for Musculoskeletal Health, The University of Sydney and Sydney Local Health District, New South Wales, Australia*

Received 24 January 2023; received in revised form 4 October 2023; accepted 5 November 2023
Available online 14 November 2023

KEYWORDS
Clinical practice guidelines;
Comparator bias;
Low back pain;
Sciatica

Abstract
Background: Low back pain (LBP) is the leading cause of disability globally. Clinical practice guidelines (CPGs) have been developed in hopes of encouraging evidence-based care for LBP. However, poor quality of trials that underpin CPGs can lead to misleading recommendations for LBP.
Objectives: To categorize the comparator used in trials included in the National Institute of Clinical Excellence (NICE) LBP CPG and describe the proportion and association of suboptimal comparators with NICE recommendation.
Methods: We conducted a cross-sectional analysis to describe the proportion of trials included in the NICE LBP CPG that used a suboptimal comparator. If comparators used an ineffective treatment, a treatment of unknown effectiveness, or no or minimal treatment then they were considered suboptimal.
Results: We included 408 trials and analyzed 580 comparators used in the trials. 30.9% of the comparators used in the trials were suboptimal. Trials testing invasive treatments (32.4%) had the highest proportion of suboptimal comparators followed by non-surgical (32.3%) and pharmacological (19.0%) treatments. Trials using suboptimal treatments were less likely to have their treatment recommended (odds ratio: 0.68; 95% CI: 0.47, 0.98) for use by NICE.
Conclusion: There is a concerning proportion of suboptimal comparators used in LBP trials that may be misleading CPG recommendations, funding allocation decisions, and ultimately clinical practice. Efforts to increase the use of optimal comparators in LBP trials are urgently needed to better understand what treatments should be recommended.

* Corresponding author at: Brown University, 121 S Main St, Providence, RI, 02903, United States.
E-mail: geronimo_bejarano_cardenas@brown.edu (G. Bejarano).

https://doi.org/10.1016/j.bjpt.2023.100563

Introduction

Low back pain (LBP) is the leading cause of disability worldwide with an increasing social and economic burden.[1] Suboptimal care results in a large economic burden and an increased risk of transition from acute to chronic LBP.[1] The World Health Organization defines clinical practice guidelines (CPGs) 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."[2] Dissemination of the recommendations provided by CPGs are encouraged to increase use of effective care. Furthermore, recommendations provided by CPGs are used to inform research and health funding allocation decisions. The National Institute of Clinical Excellence (NICE) LBP CPG has been graded as high quality while including a broad scope of treatment options ranging from invasive to non-surgical treatments.[3,4]

While the quality and reporting of the NICE LBP CPG is considered to be high,[3] little is discussed about the quality of evidence underpinning its recommendations — particularly on the impact of trial design, such as the comparators used in the trials that inform the CPG. Testing a treatment against a suboptimal comparator may produce inaccurate results by overestimating the treatment effect, this is known as comparator bias.[5] The use of studies with comparator bias could result in misleading recommendations in CPGs, which ultimately impacts clinical care, health funding allocation decisions, and research.

To our knowledge, there has been no analysis assessing the quality of comparators in trials included in CPGs for LBP. Therefore, the aim of this study was to categorize the comparator used in trials included in the NICE LBP CPG, determine the proportion of trials that used a suboptimal comparator, and assess the relationship between using a suboptimal comparator and the recommendation made by the NICE LBP CPG.

Methods

We conducted a cross-sectional analysis of trials that informed recommendations in the NICE LBP CPG.[6] We chose the NICE CPG as it has been consistently graded as high-quality and includes a broad range of treatment from invasive (e.g., fusion surgery, disc replacement, etc.) to non-invasive (e.g., pharmacological, manual therapy, exercise, etc.) interventions.[3] We reported this study per the STROBE guidelines.[7]

We developed a standardized data extraction form which was independently tested and piloted by two authors (GB and BC) on a random sample of 205 (50%) trials followed by discussion between the two authors to get familiar with the coding framework. One author (GB) extracted data from the remaining trials (n = 203) using the agreed coding framework. We extracted the following data from trials: title, author, journal, publication year, treatment, comparator (e.g., exercise, manual therapy, NSAIDs, paracetamol, opioids, etc.), comparator effectiveness (effective, ineffective, or unknown effectiveness), type of comparator (type of treatment, no treatment, minimal treatment, usual care, or placebo) and whether the treatment was recommended for or against by NICE. Examples of minimal treatment include booklets/advice or treatments defined as 'minimal treatment' by trial authors themselves.

We used evidence from systematic reviews, preferentially Cochrane reviews, published in the past 10 years to classify comparators as effective, ineffective, or unknown effectiveness. We chose Cochrane reviews as our primary benchmark for evidence quality as their quality is typically higher than non-Cochrane reviews.[8] We classified comparators in the included trials as optimal or suboptimal. Optimal comparators were treatments for which there was evidence of effectiveness or placebo comparators. Suboptimal comparators were treatments known to be ineffective, those with unknown effectiveness, or when participants in the trials received no or minimal treatment. We classified comparators including multiple treatments as optimal if at least one component of the intervention was effective as judged by both authors extracting the data. If a comparator was named usual care, we classified the comparator by the actual treatments delivered when study reporting was sufficient. The unit of analysis of this study is the number of control groups, not trials, as many trials had more than one control group.

We described the proportion of suboptimal comparators in the overall sample and stratified by type of treatment (non-surgical, pharmacological, or invasive). We also investigated whether treatments compared to suboptimal comparators were more likely to be recommended for or against by the NICE CPG using logistic regression. The recommendation made by NICE (for/against) for each treatment was the independent variable. Output from the logistic regression was reported as odds ratio (95% CI). The analysis was conducted in Stata version 17.0 (StataCorp LLC, College Station, TX).

Results

The NICE LBP CPG included 424 unique trials to inform their treatment recommendations, and we were able to retrieve full-texts from 408 trials and 580 comparators There were 284 (69.6%) non-surgical, 80 (19.6%) invasive, and 44 (10.8%) pharmacological trials. Overall, 179 (30.9%) comparators were classified as suboptimal. The largest proportion of suboptimal comparators were in invasive trials with 33 (32.4%) suboptimal comparators, followed by 134 (32.3%) in non-surgical, and 12 (19.0%) in pharmacological trials (Table 1). There were 106 (28.0%) suboptimal comparators used in trials of treatments recommended by NICE. Overall, trials that had a suboptimal comparator were less likely to have their treatment recommended (odds ratio 0.68, 95% CI: 0.47, 0.98).

Discussion

We found that almost a third of comparators used in trials that informed the NICE LBP CPG are suboptimal. Trials of invasive treatments most often used suboptimal

Table 1 Proportion and association of suboptimal and optimal comparators in treatment categories by NICE recommendation.

Treatment category	Recommended For (*n* = 379) n (%)	Recommended Against (*n* = 201) n (%)	Total n (%)	Suboptimal Comparator, Odds Ratio (95 % CI)
Invasive			102 (17.6)	0.17 (0.07, 0.41)
Suboptimal	10 (16.7)	23 (54.8)	33 (32.4)	
Optimal	50 (83.3)	19 (45.2)	69 (67.6)	
Pharmacological			63 (10.9)	*
Suboptimal	0 (0.0)	12 (22.6)	12 (19.0)	
Optimal	10 (100.0)	41 (77.4)	51 (81.0)	
Non-surgical			415 (71.5)	0.81 (0.51, 1.28)
Suboptimal	96 (31.1)	38 (35.8)	134 (32.3)	
Optimal	213 (68.9)	68 (64.2)	281 (67.7)	
Overall			580 (100.0)	0.68 (0.47, 0.98)
Suboptimal	106 (28.0)	73 (36.3)	179 (30.9)	
Optimal	273 (72.0)	128 (63.7)	401 (69.1)	

* Denotes that logistic regression was not able to be performed due to no pharmacological treatments that were recommended by NICE being informed by trials that used a suboptimal comparator.

comparators. Interestingly, using suboptimal comparators reduced the likelihood of the treatment being recommended for use by NICE. To the best of our knowledge, this is the first study assessing comparator quality in trials testing treatments for LBP, and more broadly, musculoskeletal conditions. A 2019 study found that 17% of trials of oncological drugs approved by the United States Food and Drug Administration used suboptimal comparators.[9] This is similar to the proportion of suboptimal comparators in pharmacological trials in our study (19.0%). However, we found a greater proportion of suboptimal comparators overall (30.9%). This may be because they assessed trials approved by the United States Food and Drug Administration, which does not include trials denied due to quality.

Strengths of this study include using a large sample size of trials (*n* = 408) with various comparators (*n* = 580) and using a systematic data extraction form to assess the quality of the comparators. Using the NICE LBP CPG, which is high-quality and comprehensive, to investigate the proportion of LBP trials that use a suboptimal comparator can also be considered a strength.[3] A limitation of our study is that by only investigating the body of evidence that informs the NICE LBP CPG, we have missed trials not included in the NICE LBP CPG including those published after the 2016 NICE LBP CPG publication. Evaluating all trials of treatments for LBP would not be feasible due to the large number of published trials. However, the trials that we did assess are part of a distinct group of studies that directly shaped the recommendations of a high-quality CPG used globally.

The proportion of trials in LBP using suboptimal comparators is concerning as they may overestimate treatment effects, and negatively influence CPG recommendations, funding allocation decisions, and clinical practice.[5] Our findings show that a substantial proportion of trials underpinning the recommendations in favor of treatments that are now considered first-line care use suboptimal comparators. Trials with suboptimal comparators do not contribute to advancing care and waste resources. Future trials must use optimal comparators commonly used in clinical practice that are appropriately methodologically designed for the specific research question.[10]

Acknowledgements

JRZ and GEF are supported by a National Health Medical and Research Council (NHMRC) Emerging leader fellowship (APP1194105 and APP2009808).

References

1. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. *Lancet*. 2018;391 (10137):2356–2367. https://doi.org/10.1016/S0140-6736(18)30480-X.
2. Organization WH. *WHO Handbook for Guideline Development*. World Health Organization; 2014. Accessed June 28, 2021. https://apps.who.int/iris/handle/10665/145714.
3. Castellini G, Iannicelli V, Briguglio M, et al. Are clinical practice guidelines for low back pain interventions of high quality and updated? A systematic review using the AGREE II instrument. *BMC Health Serv Res*. 2020;20(1):970. https://doi.org/10.1186/s12913-020-05827-w.
4. Gianola S, Bargeri S, Cinquini M, Iannicelli V, Meroni R, Castellini G. More than one third of clinical practice guidelines on low back pain overlap in AGREE II appraisals. Research wasted? *BMC Med Res Methodol*. 2022;22(1):184. https://doi.org/10.1186/s12874-022-01621-w.
5. Mann H, Djulbegovic B. Comparator bias: why comparisons must address genuine uncertainties. *J R Soc Med*. 2013;106 (1):30–33. https://doi.org/10.1177/0141076812474779.
6. National Guideline Centre (UK). *Low Back Pain and Sciatica in Over 16s: Assessment and Management*. National Institute for Health and Care Excellence (UK); 2016. Accessed June 11, 2021. http://www.ncbi.nlm.nih.gov/books/NBK401577/.
7. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61 (4):344–349. https://doi.org/10.1016/j.jclinepi.2007.11.008.

8. Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *J Clin Epidemiol*. 2009;62(10):1021—1030. https://doi.org/10.1016/j.jclinepi.2008.09.018.
9. Hilal T, Sonbol MB, Prasad V. Analysis of control arm quality in randomized clinical trials leading to anticancer drug approval by the US Food and Drug Administration. *JAMA Oncol*. 2019;5(6):887—892. https://doi.org/10.1001/jamaoncol.2019.0167.
10. Williams CM, Skinner EH, James AM, Cook JL, McPhail SM, Haines TP. Comparative effectiveness research for the clinician researcher: a framework for making a methodological design choice. *Trials*. 2016;17(1):406. https://doi.org/10.1186/s13063-016-1535-6.