Journal of Clinical Epidemiology 116 (2019) 36−38

ORIGINAL ARTICLE

Poor quality patient reported outcome measures bias effect estimates in orthopaedic randomized studies

Joel J. Gagnier[a,b,*], Bradley C. Johnston[c]

[a]Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI, USA
[b]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA
[c]Department of Community Health & Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada

Accepted 16 July 2019; Published online 30 July 2019

Abstract

Objectives: The objective was to assess the potential for biased treatment effects associated with patient-reported outcome measures (PROMs) of varying psychometric quality in randomized clinical trials (RCTs) for rotator cuff disease (RCD).

Study Design and Setting: We searched for RCTs published in the past 5 years (January 2011 to December 2016) in the top five 2015 impact factor orthopedic journals. We accepted RCTs including human participants with RCD, published in English, and using PROMs specific to RCD. We extracted data on study design, sample size, risk of bias for RCTs, quality of PROM used, estimates of effect, and associated measures of variance. PROMs were given numerical ratings of psychometric quality from a prior publication. Continuous measures of effect were transformed by dividing the effect estimate by the standard deviation. Multilevel linear regression analyses were performed to determine whether PROM quality was associated with the magnitude of effect.

Results: Overall, we included 72 RCTs reporting 174 separate outcomes. Mean sample size was 66.8 (95% CI 62.30 to 71.27), mean risk of bias score across all studies was 7.00/10 (95% CI 6.72 to 7.29), psychometric quality summary scores ranged from −2 to 10, and the standardized mean effect estimate was 0.47 (95% CI −0.17 to 1.11). Regression revealed that higher-quality PROMs had smaller estimates of effect (β = −0.32; 95% CI −0.51 to −0.13; $P = 0.001$). We also found that a longer follow-up period predicted slightly increased effect estimates (β = 0.08; 95% CI 0.02 to 0.13; $P = 0.007$).

Conclusions: PROMs with poor or unknown psychometric properties overestimate treatment effects in clinical research of RCD by 68.4% (β −0.32/standardized mean effect 0.47). To our knowledge, this is the first empirical evidence that variations in the quality of PROMs bias treatment effect estimates. Researchers and clinicians using data from PROMs must be cautious to explore the quality of that measure so as to not mislead decision-making resulting from biased outcomes.

Keywords: Bias; Patient reported outcomes; Clinical trials; Randomized trials; Orthopaedics; Musculoskeletal conditions; Psychometrics

1. Introduction

Patient-reported outcome measures (PROMs) provide self-reported, subjective data from the patients themselves, without interpretation by clinicians or others, and their use has increased [1−6]. There are, however, a wide array of heterogeneous PROMs used in clinical research, and evidence is mounting that many of these measures have poor or unknown psychometric properties [7−16]. The lack of psychometric evidence for PROMs used in clinical research calls into question the data derived from these outcome measures and any decision-making based on such findings. The purpose of this study was to assess how varying PROM quality influences treatment effect estimates in randomized trials of patients with rotator cuff disease (RCD) and secondarily to identify variables that modify this relationship.

2. Methods

This study involved five steps [17−21]:

1. Identification of PROMs used in RCD and assess their psychometric properties.
2. Identification of a sample of randomized trials among patients with RCD using PROMs in step 1.
3. Data extraction.

Conflict of interest: Neither author has any conflict of interest associated with the content of this manuscript.

* Corresponding author. Tel.: 519-999-4903.
E-mail address: jgagnier@umich.edu (J.J. Gagnier).

https://doi.org/10.1016/j.jclinepi.2019.07.012
0895-4356/

What is new?

Key findings

- PROMs with poor or unknown psychometric properties overestimate treatment effects in clinical research of rotator cuff disease by 68.4% (β −0.32/ standardized mean effect 0.47).

What this adds to what was known?

- To our knowledge, this is the first empirical evidence that variations in the quality of PROMs bias treatment effect estimates.

What is the implication and what should change now?

- Researchers and clinicians using data from PROMs must be cautious to explore the quality of that measure so as to not mislead decision-making resulting from biased outcomes.

4. Risk of bias assessment.
5. Statistical analyses: Mixed effects linear regression with elimination and a sensitivity analysis. STATA/ MP 14.2 was used.

3. Results

A total of 72 randomized clinical trials (71 with multiple outcomes) reporting 174 separate outcomes were included. The mean sample size was 66.8 (95% CI 62.30 to 71.27), mean risk of bias score was 7.00/10 (95% CI 6.72 to 7.29), and mean follow-up was 9.7 months (95% CI 7.6 to 11.7). Fig. 1 shows a scatterplot and line of best fit for the standardized treatment effect estimate for each PROM psychometric summary score value.

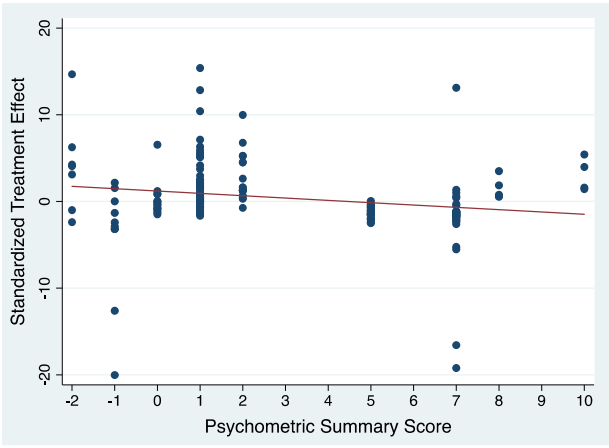

Fig. 1. Scatterplot for treatment effect estimates by psychometric summary score.

Table 1. Mixed effects linear regression: final model

Covariate	Beta coefficient (95% CI)	*P*-value
Psychometric summary score	−0.32 (−0.52 to −0.12)	0.002
Follow-up (mo)	0.08 (0.02 to 0.13)	0.007

Following the step-wise elimination procedure, only two variables remained in the final model (see Table 1)—psychometric summary score and follow-up months (N = 174 [adjusted for 71 clusters]; Model P-value < 0.0001). The ratio of the psychometric summary score influence on the standardized mean effect estimate was −0.582 (−0.32/0.47)—meaning, with each 1-point increase in psychometric summary score, the effect estimate decreased by 68.4%, after controlling for the follow-up period. In the sensitivity analysis, adjusting for a lack of reporting had no influence on the findings.

4. Discussion

We found that in randomized trials among patients with RCD using PROMs, the effect estimates were exaggerated on average by 68.4% among PROMs with poor or unknown psychometric properties. Specifically, with each one-unit increase in the quality of the PROM, the treatment effect estimates decrease by 0.32 units, which can change the summary score from a large-effect to a moderate-effect size or from a moderate-effect to a small-effect size. That is, PROMs with questionable psychometric properties significantly bias (i.e., inflate) the estimates of treatment effect. To our knowledge, this is the first empirical evidence that variations in the psychometric properties of disease-specific PROMs bias treatment effect estimates in clinical research.

CRediT authorship contribution statement

Joel J. Gagnier: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Bradley C. Johnston:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing.

References

[1] Cella DF. Measuring quality of life in palliative care. Semin Oncol 1995;22(2 Suppl 3):73–81.

[2] Johnston BC, Patrick DL, Busse JW, Schünemann HJ, Agarwal A, Guyatt GH. Patient-reported outcomes in meta-analyses–Part 1: assessing risk of bias and combining outcomes. Health Qual Life Outcomes 2013;11:109.

[3] FDA guidance on patient reported outcomes (PROs). Available at http://www.ispor.org/workpaper/FDA%20PRO%20Guidance.pdf. Accessed October 15, 2015.
[4] Basch E. The missing voice of patients in drug safety reporting. N Engl J Med 2010;362:865−9.
[5] Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. BMJ 1998;317:1191−4.
[6] Soreide K, Soreide AH. Using patient reported outcome measures for improved decision making in patients with gastrointestinal cancer-the last clinical frontier in surgical oncology? Front Oncol 2013;3:157.
[7] Huang H, Grant JA, Miller BS, Gagnier JJ. A systematic review of the psychometric properties of patient-reported outcome instruments for use in patients with rotator cuff disease. Am J Sports Med 2015;43(10):2572.
[8] Gagnier JJ, Mullins M, Huang W. Patient reported outcome questionnaires for total knee arthroplasty: a systematic review of psychometric properties and methodological quality. J Arthroplasty 2017;32(5):1688−97.
[9] Jia Y, Huang H, Gagnier JJ. Patient reported outcome questionnaires for foot and ankle conditions: a systematic review of psychometric properties and methodological quality. Qual Life Res 2017;26(8):1969−2010.
[10] Bryan S, Whitehurst D. Patient-reported outcome measurement (PROMs) in British Columbia: What has been achieved so far and where next?. Quality Forum 2013, Vancouver Coastal Health Research Institute 2013.
[11] Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. Oxford, UK: Oxford Press; 2008.
[12] DeVellis RF. Scale development: theory and applications. 3rd ed. Thousand Oaks, CA: SAGE Publications; 2012.
[13] Terwee CB, Bot SDM, de Boer MR, van der Windt DAWN, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34−42.
[14] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737−45.
[15] Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomized controlled trials of treatments for schizophrenia. Br J Psychiatry 2000;176:249−52.
[16] Mercieca-Bebber RL, Perreca A, King M, Macann A, Whale K, Soldati S, et al. Patient-reported outcome in head and neck and thyroid cancer randomized controlled trials: a systematic review of completeness of reporting and impact on interpretation. Eur J Cancer 2016;56:144−61.
[17] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539−49.
[18] Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651.
[19] Schellingerhout JM, Verhagen AP, Heymans MW, Heymans MW, Koes BW, de Vet HC, et al. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. Qual Life Res 2012;21:659.
[20] Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 (updated March 2011). The Cochrane Collaboration. Available at www.cochrane-handbook.org. Accessed June 1, 2019.
[21] Gagnier JJ, Morgenstern H, Chess L. Meta-analysis of neuromuscular interventions for prevention of anterior cruciate ligament injuries. Am J Sports Med 2013;41:1952−62.