RBO REVISTA BRASILEIRA DE ORTOPEDIA
www.rbo.org.br/

Updating Article

Prospective non-randomized studies in Orthopaedics and Traumatology: systematic assessment of its methodological quality

Gustavo Soriano Pignataro,[1] *Theophilo Ásfora Lins,*[1] *José Renato Assis Lemos Marques de Oliveira,*[2] *Vinícius Ynoe de Moraes,*[1] *Aldo Okamura,*[1] *João Carlos Belloti,*[3] *Flávio Faloppa*[4]

[1]*Third-year Resident Physician, Department of Orthopedics and Traumatology, Escola Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil.*
[2]*Orthopedist and Second-year Resident Physician in Hand Surgery, Escola Paulista de Medicina, UNIFESP, São Paulo, SP, Brazil.*
[3]*Adjunct Professor, Department of Orthopedics and Traumatology, Escola Paulista de Medicina, UNIFESP, São Paulo, SP, Brazil.*
[4]*Titular Professor, Department of Orthopedics and Traumatology, Escola Paulista de Medicina, UNIFESP, São Paulo, SP, Brazil.*
Work performed in the Discipline of Hand and Arm Surgery, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil.

ARTICLE INFO

Article history:
Received on March 22, 2012
Approved on May 28, 2012

Keywords:
Evidence-Based Medicine
Prospective Studies; Orthopedics
Traumatology
Bibliometrics

ABSTRACT

In surgical interventions, randomization and blinding may be difficult to implement. In this situation, non-randomized prospective studies (EPNR) can generate the best evidence. The objective of this study is to evaluate, by means of the scale proposed by Downs & Black, the quality of EPNR published in our country and to assess the interobserver reproducibility of this scale. EPNR published in Acta Ortopedica Brasileira and Revista Brasileira de Ortopedia until 2011 and prior to 2006 were included. Two of us independently applied the Downs & Black scale. The studies were stratified by period of publication, journal and type of intervention. The scores obtained were considered to assess the reliability of the scale and groups comparison. 59 studies were considered, seven excluded during the assessments. There were no differences between the scores, except for the type of intervention, which showed better methodological quality for studies involving clinical interventions ($p < 0.001$). The correlation coefficient for the Downs & Black score was 0.79 (95% CI 0.65 to 0.88), demonstrating good reliability. EPNR present methodological quality similar when stratified by the periodic publication and publication period. Studies with clinical interventions have better methodological quality. The Downs & Black scale shows good interobserver reproducibility.

Corresponding author at: Rua Borges Lagoa, 778, Vila Clementino, CEP 04038-001, São Paulo, SP, Brazil.
Tel: (11) 55717106.
E-mail: vymoraes@gmail.com

issn
doi: 10.1016/j.rboe.2012.05.005

Introduction

Randomized clinical trials are the gold-standard studies for determining what form of treatment is best.[1-5] However, in situations involving surgical interventions, randomization and/or blinding may not be feasible or may be difficult to implement. This leads to a need to institute treatment grounded in non-randomized studies.[6] Within this spectrum, prospective series provide the greatest volume of published data and these, if conducted with methodological rigor, should not be dismissed.[2,7-10]

Despite the lower decision-making power coming from these studies, worldwide efforts have concentrated on optimizing the methodological rigor of these series, through instituting rigid inclusion criteria and systematic and objective outcome measurements that minimize the bias that may subjectively interfere with evaluations on the results.[6,8,11-14]

Another concern is in relation to the publication quality of the results. Guidelines based on specific study designs (STROBE,[15] MOOSE,[16] STARD,[17] PRISMA[18] and CONSORT[19]) are available and are prerequisites for publishing in journals with significant impact factors.[16,20] The present study took the hypothesis that potentially correctable methodological failures exist in non-randomized studies on surgical treatment. These are capable of improvement through instituting the elementary precepts of evidence-based medicine and following its guidelines.

The objectives of this study were: 1) To critically analyze prospective non-randomized studies that include surgical treatment, in the Brazilian orthopedic literature, by means of descriptive evaluation on a defined summary scale; 2) To determine whether there has been any improvement in production quality and whether there are any differences in the quality of studies, when categorized according to the journal; 3) To assess the interobserver reproducibility of these scores.

Method

This secondary analysis took into consideration for assessment studies that had previously been published in the Brazilian orthopedic literature: Acta Ortopédica Brasileira (AOB) and Revista Brasileira de Ortopedia (RBO). All studies published up to 2011 (2006-2011) and up to 2005, in each journal, were selected.

Inclusion criteria

The studies included were prospective non-randomized studies that reported on interventions relating to orthopedic and/or traumatological diseases.

Exclusion criteria

Retrospective studies, accuracy studies, studies assessing reproducibility, prospective randomized studies and studies presenting difficulty or uncertainty in identifying their temporal nature (prospective versus retrospective) were excluded.

Intervention

The studies were selected from computerized records: issue by issue, through identifying studies that fitted within the inclusion criteria. Doubts that arose during this process were resolved by reaching a consensus between two of the investigators (V.Y.M. and A.O.). After studies had been identified, two examiners (A.O. and T.A.) independently gathered the qualitative characteristics of the studies: journal (AOB or RBO), year of publication (period from 2006 to 2011 or 2005 and earlier), length of follow-up (months), subspecialty (knee, foot, hand, traumatology and others) and type of intervention (clinical or surgical).

The systematic quality assessment was done by means of the Downs and Black score.[21] This tool has the aim of evaluating methodological quality and was specially designed to take into consideration randomized and non-randomized studies. It contains 27 scorable items. This scale is available in an indexed published paper (jech.bmj.com/content/52/6). Divergences and doubts relating to the assessments were resolved by means of reaching a consensus, mediated by a third investigator (J.C.B.).

Statistical analysis

For the descriptive evaluation, means and their standard deviations were taken into consideration. After the studies had been categorized, the scores were compared by means of the Mann-Whitney U test. To evaluate the reliability, with the aim of assessing the interobserver concordance between the evaluations, the intraclass correlation was used,[22] and this was displayed using Bland-Altman plots. To interpret the degree of concordance for this method, values greater than 0.65 were considered to be satisfactory.

Results

Fifty-two studies were included, and seven were then excluded through consensual assessment among the present authors, because of doubts relating to their design. The evaluation on the study characteristics is presented in Tables 1 and 2.[23]

There were no differences in Downs & Black scores for any of the categories analyzed, except in relation to the type of intervention, for which the scores were seen to be higher for studies that presented non-surgical interventions (Tables 3, 4 and 5). The intraclass correlation[22] between the examiners was seen to be satisfactory (ICC = 0.79; 95% CI = 0.65-0.88) (Fig. 1).

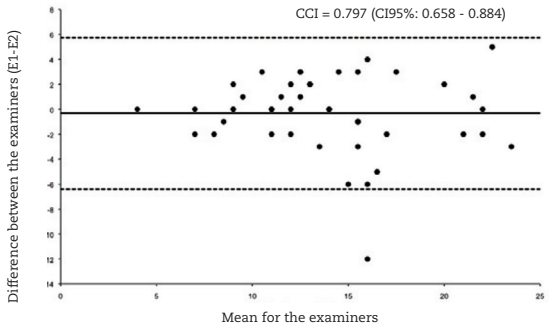

Fig. 1 - Intraclass correlation – interobserver concordance (E1 vs. E2) – Bland-Altman plot.

Table 1 - Characteristics of the studies included.

Variable	Frequency	%
Journal		
Acta Ortopédica Brasileira	26	59.1
Revista Brasileira de Ortopedia	18	40.9
Year stratified		
2005 and earlier	30	68.2
2006-2011	14	31.8
Subject		
Hand surgery	10	22.7
Knee	8	18.2
Foot	6	13.6
Traumatology	4	9.1
Others	16	36.4
Intervention		
Clinical	8	18.2
Surgical	36	81.8
Total	**44**	**100**

Table 2 - Characteristics of the studies included.

Variable	Mean	SD	N
Sample size	55.16	98,85	44
Length of follow-up (months)*	20.85	24,19	32

* For 12 studies, it was not possible to include this information.

Table 3 - Downs and Black score for examiner[23] – grouped analysis.

Variable	Category	Mean	SD	N	p
Journal	Acta Ortopédica Brasileira	13.88 5	5.30	26	0.886
	Revista Brasileira de Ortopedia	13.44	3.91	18	
Year stratified	2005 and earlier	12.97	4.36	30	0.146
	2006-2011	15.29	5.28	14	
Intervention	Clinical	19.50	3.85	8	< 0.001
	Surgical	12.42	3.90	36	

Downs & Black score E1.

Table 4 - Downs and Black score for examiner (E2) – categorized analysis.

Variable	Category	Mean	SD	N	p
Journal	Acta Ortopédica Brasileira	13.88	5.24	26	0.774
	Revista Brasileira de Ortopedia	14.22	4.65	18	
Year stratified	2005 and earlier	13.47	5.12	30	0.244
	2006-2011	15.21	4.53	14	
Intervention	Clinical	20.00	3.21	8	< 0.001
	Surgical	12.69	4.26	36	

Downs & Black score E2.

Table 5 - Comparison between different categories – Number of patients included.

Sample size	Category	Mean	SD	N	p
Journal	Acta Ortopédica Brasileira	44.62	54.59	26	0.793
	Revista Brasileira de Ortopedia	70.39	141.13	18	
Intervention	Clinical	132.88	202.85	8	0.056
	Surgical	37.89	45.68	36	

Downs & Black score E2.

Discussion

Our results demonstrate that within our setting there is no distinction of methodological quality with regard to the journal, and that there has not been any improvement in publication quality with the passage of time. This study used a tool that has frequently been used in the literature and which has been shown to be reproducible for orthopedic and traumatological studies.

In stratifying the studies according to the type of intervention (clinical or surgical), we demonstrated that the scores were higher (thus inferring better methodological quality) in studies in which the main intervention was strictly non-surgical. This can be explained by the ease of instituting blinding techniques and creating comparison groups in such studies, which are factors that the Downs & Black tool takes into account.

The internal validity of our study is based on the evaluative power of the specific scale developed by Downs & Black. This scale presents the virtue of being a universal tool for critical assessment of clinical studies in general and it presents acceptable reproducibility and internal consistency.[21] It has been shown to be easy to apply and reproducible in our experience. Thus, it seems to us to be a feasible tool for use in assessments similar to this one. Nonetheless, some authors

have recommended that it should be used cautiously, especially when it is intended to group studies according to their methodological quality, as is done in systematic reviews.[24-26]

Critical analysis on our own scientific production is little disseminated within orthopedics and traumatology. However, some studies have reported on the status of this scientific production and have identified the most frequent types of studies, in the light of evidence-based medicine,[4,5,27,28] or have attempted to compare Brazilian production with what has been published in indexed international journals.[29] The present study has the virtue of including a validated tool for assessing the quality of the scientific method, which makes this approach unique within our setting.

However, there is no agreement in the literature regarding the real capacity of these scales for categorizing the quality of these studies, especially for systematic reviews of the literature.[25] Future studies should also consider other descriptive scales or should aim to propose a checklist that would be feasible in the light of orthopedic and traumatological scientific realities. Studies published in the English language have demonstrated concern for measuring the quality of published papers, so as to provide guidelines for future studies.[24,30,31] This is the scenario that the present analysis addresses.

Conflicts of interest

The authors declare that there was no conflict of interests in conducting this study.

REFERENCES

1. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med. 1997;126(5):376-80.
2. Poolman RW, Kerkhoffs GM, Struijs PA, Bhandari M. International evidence-based orthopedic surgery working G. Don't be misled by the orthopedic literature: tips for critical appraisal. Acta Orthop. 2007;78(2):162-71.
3. Brighton B, Bhandari M, Tornetta P 3rd, FELSON DT. Hierarchy of evidence: from case reports to randomized controlled trials. Clin Orthop Relat Res. 2003;(413):19-24.
4. Moraes VY, Belloti JC, Moraes FY, Galbiatti JA, Palácio EP, Santos JBG, et al. Hierarchy of evidence relating to hand surgery in Brazilian orthopedic journals. Sao Paulo Med J. 2011;129(2):94-8.
5. Moraes VY, Moreira CD, Tamaoki MJS, Faloppa FV, Belloti JOC. Ensaios clínicos randomizados na ortopedia e traumatologia: avaliação sistemática da evidência nacional. Rev Bras Ortoped. 2010;45(6):601-5.
6. Bhandari M, Giannoudis PV. Evidence-based medicine: what it is and what it is not. Injury. 2006;37(4):302-6.
7. Pemberton J, Kraeva J, Bhandari M. Hierarchy of evidence: a simple system for orthopaedic research? J Long Term Eff Med Implants. 2007;17(2):95-102.
8. Viveiros H, Mignott T, Bhandari M. Evidence-based orthopaedics: is it possible? J Long Term Eff Med Implants. 2007;17(2):87-93.
9. Poolman RW, Petrisor BA, Marti RK, Kerkhoffs GM, Zlowodzki M, Bhandari M. Misconceptions about practicing evidencebased orthopedic surgery. Acta Orthop. 2007;78(1):2-11.
10. Poolman RW, Sierevelt IN, Farrokhyar F, Mazel JA, Blankevoort L, Bhandari M. Perceptions and competence in evidencebased medicine: are surgeons getting better? A questionnaire survey of members of the Dutch Orthopaedic Association. J Bone Joint Surg Am. 2007;89(1):206-15.
11. Petrisor BA, Bhandari M. Principles of teaching evidencebased medicine. Injury. 2006;37(4):335-9.
12. Van Oldenrijk J, Sierevelt IN, Schafroth MU, Poolman RW. Design considerations in implant-related randomized trials. J Long Term Eff Med Implants. 2007;17(2):153-63.
13. Degen RM, Hodgins JL, Bhandari M. The language of evidence based medicine: answers to common questions? Indian J Orthop. 2008;42(2):111-7.
14. Hoppe DJ, Bhandari M. Evidence-based orthopaedics: a brief history. Indian J Orthop. 2008;42(2):104-10.
15. Vandenbroucke JP, Von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. Plosmed. 2007;4(10):e297.
16. Mundi R, Chaudhry H, Singh I, Bhandari M. Checklists to improve the quality of the orthopaedic literature. Indian J Orthop. 2008;42(2):150-64.
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem. 2003;49(1):7-18.
18. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009;62:e1-e34.
19. Moher D, Schulz K, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. Lancet. 2001;357(9263):1191-4.
20. Mundi R, Chaudhry H, Sharma R, Schemitsch E, Bhandari M. What is the quality of the orthopaedic literature? J Long Term Eff Med Implants. 2007;17(2):103-9.
21. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998;52(6):377-84.
22. Goldfarb CA, Borrelli Jr. J, Lu M, Ricci WM. A prospective evaluation of patients with isolated orthopedic injuries transferred to a level I trauma center. J Orthop Trauma. 2006;20(9):613-7.
23. Bhandari M, Sprague S, Schemitsch EH, International Hip Fracture Research Collaborative. Resolving controversies in hip fracture care: the need for large collaborative trials in hip fractures. J Orthop Trauma. 2009;23(6):479-84.
24. Bhandari M, Morshed S, Tornetta P 3rd, Schemitsch EH. Design, conduct, and interpretation of nonrandomized orthopaedic studies: a practical approach. (All) evidence matters. J Bone Joint Surg Am. 2009;91(Suppl 3):1.
25. Simunovic N, Sprague S, Bhandari M. Methodological issues in systematic reviews and meta-analyses of observational studies in orthopaedic research. J Bone Joint Surg Am. 2009;91(Suppl 3):87-94.
26. Norris SL, Atkins D. Challenges in using nonrandomized studies in systematic reviews of treatment interventions. Annals Internal Med. 2005;142(12Pt2):1112-9.

27. De Moraes VY, Gracitelli GC, Rezende FC, Da Silva MSP, Viola DCM, Garcia RJ. Endoprostheses-related complications in patients with bone tumors of the knee. Acta Ortop Bras. 2011;19(1):49-51.
28. Riera R. Designs of studies published in two Brazilian journals of orthopedics and sports medicine, recently indexed in the ISI Web of Science. São Paulo Med J. 2009;127(6):355-8.
29. Amatuzzi MLL. Análise da evolução qualitativa de publicações em ortopedia e traumatologia: comparação entre a Revista Brasileira de Ortopedia e o Journal of Bone and Joint Surgery. Rev Bras Ortop. 2004;39(9):527-35.
30. Bhandari M, Tornetta P 3rd. Evidence-based orthopaedics: a paradigm shift. Clin Orthop Relat Res. 2003;(413):9-10.
31. Dahm P, Poolman RW, Bhandari M, F, Baum J, Kosiak B, et al. Perceptions and competence in evidence-based medicine: a survey of the American Urological Association Membership. J Urol. 2009;181(2):767-77.