

# Methodological quality of systematic reviews and clinical trials on women's health published in a Brazilian evidence-based health journal

Cristiane Rufino Macedo, Rachel Riera, Maria Regina Torloni

Brazilian Cochrane Center, São Paulo/SP, Brazil.

**OBJECTIVES:** To assess the quality of systematic reviews and clinical trials on women's health recently published in a Brazilian evidence-based health journal.

**METHOD:** All systematic reviews and clinical trials on women's health published in the last five years in the Brazilian Journal of Evidence-based Health were retrieved. Two independent reviewers critically assessed the methodological quality of reviews and trials using AMSTAR and the Cochrane Risk of Bias Table, respectively.

**RESULTS:** Systematic reviews and clinical trials accounted for less than 10% of the 61 original studies on women's health published in the São Paulo Medical Journal over the last five years. All five reviews were considered to be of moderate quality; the worst domains were publication bias and the appropriate use of study quality in formulating conclusions. All three clinical trials were judged to have a high risk of bias. The participant blinding, personnel and outcome assessors and allocation concealment domains had the worst scores.

**CONCLUSIONS:** Most of the systematic reviews and clinical trials on women's health recently published in a Brazilian evidence-based journal are of low to moderate quality. The quality of these types of studies needs improvement.

**KEYWORDS:** Evidence-Based Medicine; Women's Health; Review; Clinical Trial; Research Design.

Macedo CR, Riera R, Torloni MR. Methodological quality of systematic reviews and clinical trials on women's health published in a Brazilian evidence-based health journal. *Clinics*. 2013;68(4):563-567.

Received for publication on December 4, 2012; First review completed on December 5, 2012; Accepted for publication on December 23, 2012

E-mail: [rachelriera@hotmail.com](mailto:rachelriera@hotmail.com)

Tel.: 55 11 5575-2970

## ■ INTRODUCTION

According to official data from a survey conducted in 2011 by the Brazilian Federal Medical Council, there are 204,563 practicing physicians in the country, 167,225 of whom are specialists. Obstetricians-gynecologists (OB-GYN) account for approximately 12% (22,815) of these specialists, second only to pediatricians (27,232) (1). OB-GYNs are directly responsible for over 5 million pregnancies that occur each year in the country, in addition to the care of over 97 million women in Brazil (2).

The importance of the continued medical education of these physicians and of ensuring their access to the best possible evidence is unquestionable. Articles published in medical journals are important sources of information and medical education for OB-GYNs and clinicians in general. Although systematic reviews (SRs) and clinical trials (CTs)

are considered the highest level of evidence (3), the quality of their methodology is not homogeneous, and these publications should be as rigorously evaluated as other types of studies (4). Thus, readers and users of SRs and CTs should maintain a critical perspective and look carefully at the methodological quality of the existing publications.

SRs involve an exhaustive review of the literature to answer a clearly defined clinical question using a systematic, transparent and explicit methodology to identify, select, critically appraise and synthesize all of the existing evidence (4). Conducting an SR is a complex task, and flaws are possible in this process; these factors lead to variations in the quality of published SRs. A CT is a difficult study and frequently involves a considerable number of researchers and patients to answer a question on treatment or prevention. In an attempt to avoid or minimize bias, a rigorous methodology must be used. However, despite this rigor, bias can compromise findings, and readers must keep this in mind.

There is scarce literature on the quality of Brazilian SRs and CTs in general and, to the best of our knowledge, there have been no previous studies that analyzed the methodological quality of these types of studies on women's health.

**Copyright** © 2013 **CLINICS** – This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

**DOI:** 10.6061/clinics/2013(04)20



Therefore, we set out to critically assess the quality of SRs and CTs on women's health recently published in a Brazilian medical journal.

## MATERIALS AND METHODS

This observational study was performed by researchers of the Brazilian Cochrane Center. Two independent investigators manually reviewed all electronic issues of the São Paulo Medical Journal (SPMJ-Brazilian Journal of Evidence-based Health) published between 2008 and 2012 and available through the SciELO database. All SRs and CTs focused on women's health were eligible for inclusion. The methodological quality and risk of bias of these articles were assessed independently by each of the investigators. The results were compared, and differences in ratings were discussed until a consensus was reached. In the case of disagreement, a third investigator was consulted.

To assess the quality of SRs, the authors used the AMSTAR tool, which consists of 11 items that are rated as 0 or 1 (5). This tool has good face and content validity for measuring the methodological quality of SRs and requires approximately 10–15 minutes for completion (6). AMSTAR has an acceptable inter-rater agreement of the individual items, with a mean kappa of 0.70 (95% confidence interval: 0.57, 0.83). The intra-class correlation coefficient is 0.84 (95% confidence interval: 0.65, 0.92) (7). Using this tool, we classified the following 11 items for each SR: 1. a priori design; 2. duplicate study selection and data extraction; 3. comprehensive literature search; 4. inclusive publication status; 5. included/excluded studies provided; 6. characteristics of included studies provided; 7. quality assessment of studies; 8. study quality used appropriately in formulating conclusions; 9. appropriate methods used to combine studies; 10. publication bias assessed; and 11. conflict of interest stated. Each of these items was classified as "Yes," "No," "Can't answer" or "Not applicable". We calculated the AMSTAR final score by adding one point for each "Yes" answer and no points for all other answers resulting in summary scores ranging from 0 to 11. For rating the overall quality of the SR, the following categories were used: 0–4 = low-quality SR, 5–8 = moderate-quality SR and 9–11 = high-quality SR (8).

To assess the quality of CTs, the authors used the Risk of Bias Table, which was developed by the Cochrane Collaboration and is available in the Cochrane Handbook (9). This tool consists of seven domains: i) sequence generation, ii) allocation concealment, iii) blinding of participants and personnel, iv) blinding of outcome assessors, v) incomplete outcome data, vi) selective reporting and vii) other sources of bias. These domains are classified as "Yes" (i.e., low risk of bias), "Unclear" (i.e., uncertain risk of bias) or "No" (i.e., high risk of bias). As recommended by the Cochrane Handbook, the overall classification of each CT was based on the rating of the first four domains (9). A study was classified as having a high risk of bias when at least one of the answers to these four items was "No." When at least one of the answers to these four items was "Unclear," the trial was classified as being at an unclear or moderate risk of bias.

## RESULTS

The SPMJ publishes six editions per year, with an average of 11 articles per edition. Between the beginning of 2008 and the third edition of 2012, a total of 196 articles were published, of which 61 were related to women's health,

including five SRs and three CTs. All SRs were on gynecological topics: teriparatide for osteoporosis in postmenopausal women (10), lapatinib for advanced or metastasized breast cancer (11), comparative evaluation of digital mammography and film mammography (12), colposcopic triage methods for grade 3 cervical intraepithelial neoplasia (CIN3) after a cytopathological diagnosis of a low-grade squamous intraepithelial lesion (13) and risk of persistent high-grade squamous intraepithelial lesion after an electro-surgical excision with positive margins (14). Three of the reviews focused on treatment (10,11,14), and two focused on diagnosis (12,13). Three of these SRs presented meta-analyses of their results (10,13,14).

The CTs were on ropivacaine plus clonidine for labor analgesia (15), upper limb rehabilitation after breast cancer mastectomy with preservation of the medial pectoral nerve (16) and pelvic floor muscle training *versus* hypopressive exercises for pelvic organ prolapse (17).

### Description of the evidence presented in systematic reviews

**Teriparatide in postmenopausal women with osteoporosis.** The authors analyzed five randomized trials involving 3,504 women and concluded that compared to placebo, the intermittent administration of 20 or 40 µg of teriparatide reduced new vertebral and non-vertebral fractures and improved whole-body and lumbar bone mineral density without serious adverse effects. Teriparatide (40 µg) was more effective than alendronate (10 mg/day) in increasing whole-body, femoral and lumbar bone mineral density but was similar to alendronate regarding the occurrence of new fractures (10).

**Lapatinib for the treatment of advanced or metastasized breast cancer.** The authors identified only one trial that fulfilled the selection criteria, which included 324 women. The review concluded that the combination of lapatinib plus capecitabine was more effective than capecitabine monotherapy for reducing the risk of cancer progression. However, the authors emphasized the need for more randomized clinical trials to assess the effectiveness of lapatinib alone or in association with other drugs as first- or second-line treatments for advanced breast cancer (11).

### Comparative evaluation of digital *versus* film mammography

This review included 11 studies and involved 190,322 digital and 638,348 film mammographies. The authors concluded that digital mammography was slightly more effective than film mammography in terms of cancer detection rates. There were no significant differences in recall rates between the two diagnostic methods, and the characteristics of the tumors were similar in patients screened by either type of mammography (12).

**Colposcopic triage methods for detecting CIN3 after a cytopathological diagnosis of low-grade squamous intraepithelial lesions.** Three studies involving a total of 1,766 women fulfilled the selection criteria and were included in the review. The authors concluded that there is currently no scientific evidence to support the hypothesis that colposcopic triage using oncogenic human papilloma virus (HPV)-DNA testing to detect CIN3 is better than repeated cytological tests for women with low-grade squamous intraepithelial lesions aged 35 years and older (13).



**Risk of persistent high-grade squamous intraepithelial lesions after electrosurgical excisional treatment with positive margins.** This review included four studies with a total of 1,209 women. The authors concluded that the risk of residual disease one year after electrosurgical excisional treatment was approximately 11-fold higher in cases with positive margins compared with cases with negative margins. Patients with positive margins had a 29.4% absolute risk of residual lesions during the first year and a 6% risk during the second year after the procedure. The authors emphasized that to reduce the risks of residual disease, attention should be given to correct indications, appropriate surgical procedures, correct processing of the excised specimen and appropriate choice of treatment. This treatment choice should be individualized for each case. The authors also recommended that additional studies were needed to determine the best strategy for following up these patients, particularly during the first year after excision (14).

### Description of evidence from clinical trials

**Labor analgesia with ropivacaine added to clonidine.** The authors randomized 32 women in labor to epidural analgesia with either 15 ml of ropivacaine 0.125% or 15 ml of ropivacaine 0.0625% plus 75 µg clonidine. The authors then assessed maternal pain and neonatal effects. They concluded that the pain score, sensory block level, duration of epidural analgesia and Apgar scores did not differ significantly between the two groups. However, infants of mothers who received only ropivacaine had better neurological and adaptative capacity scores (15).

**Preservation of the medial pectoral nerve following mastectomy due to breast cancer: impact on upper limb rehabilitation.** The authors of this study randomized 30 women with breast cancer to undergo modified mastectomy with either preservation or sectioning of the medial pectoral nerve. They assessed pectoral muscle strength and mass after day 43. The women who underwent nerve preservation had significantly higher muscle strength than those who underwent nerve sectioning. No differences in muscle mass or in abduction and flexion capacities of the homolateral shoulder were identified between the groups (16).

**Pelvic floor muscle training and hypopressive exercises for treating pelvic organ prolapse in women.** The authors randomized 58 women with grade II pelvic prolapse to either pelvic floor muscle training, hypopressive exercise or a control group. At baseline and at 12 weeks after the intervention, the authors used ultrasound to assess the cross-sectional area of the levator ani muscle. They reported a significant increase in this measurement in both intervention groups but not in the control group. The authors concluded that physiotherapy is effective in increasing the cross-sectional area of the levator ani muscle in women with pelvic organ prolapse and that both modalities of physiotherapy are equally effective (17).

### Critical appraisal of included studies

As depicted in Table 1, all five SRs mentioned an a priori design, a comprehensive literature search and an inclusive publication status and provided characteristics of included studies and declared whether conflicts of interest were present. On the other hand, none of the reviews considered or mentioned the quality of included trials at formulating their conclusions nor did they assess publication bias. Only two SRs (10,12) provided a list of included/excluded studies

**Table 1 - Methodological quality of systematic reviews focusing on women's health published in the São Paulo Medical Journal between the beginning of 2008 and 2012.**

Systematic review	AMSTAR item*											Overall Quality
	1	2	3	4	5	6	7	8	9	10	11	
Trevisani et al. (10)	1	1	1	1	1	1	0	0	0	0	1	7 Moderate
Riera et al. (11)	1	0	1	1	0	1	1	0	0	0	1	6 Moderate
Iared et al. (12)	1	1	1	1	1	1	0	0	1	0	1	8 Moderate
Correa et al. (13)	1	0	1	1	0	1	1	0	0	0	1	6 Moderate
Oliveira et al. (14)	1	0	1	1	0	1	0	0	1	0	1	6 Moderate

\*AMSTAR (a measurement tool to assess systematic reviews) items are: 1. a priori design; 2. duplicate study selection and data extraction; 3. comprehensive literature search; 4. inclusive publication status; 5. list of included/excluded studies provided; 6. characteristics of included studies provided; 7. quality assessment of studies; 8. study quality used appropriately in formulating conclusions; 9. appropriate methods used to combine studies; 10. publication bias assessed; and 11. conflict of interest stated.

Scale for item score: 1 = "Yes," 0 = "No," "Can't answer" or "Not applicable." The following categories were used to rate the overall quality of the reviews: score of 0-4 = low quality; 5-8 = moderate quality; and 9-11 = high quality (8).

and clearly stated that duplicate study selection and data extraction had been performed. Only two SRs assessed the quality of included CTs: Riera et al. (11) used the Cochrane risk of bias table (9), and Iared et al. (12) used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (18). Finally, only two SRs used an appropriate method to combine studies (12,14).

As presented in Table 2 and summarized in Figure 1, three CTs were classified as having a high risk of bias. One CT (15) did not provide information on sequence generation (selection bias), and two (16,17) did not report outcome assessors (detection bias) or blind the participants and personnel (performance bias).

## DISCUSSION

One of the most frequent methodological flaws of the SRs was the failure to assess the quality of the studies. Authors of SRs should grade the quality of their recommendations and the strength of the evidence presented, which inevitably depend on the quality of the original studies included in the review. Out of the five published reviews, only two (11,13) provided quality assessments of the primary studies, and none mentioned this assessment in their conclusions. Another frequent flaw was the failure to assess publication bias, which was not investigated by any of the five reviews. However, it should be noted that all of the reviews that received a "zero" on the publication bias AMSTAR item received this score because it was impossible to assess their publication bias. Several SRs had no meta-analyses, and in those studies where meta-analyses were performed, the graphics included less than ten CTs. In this case, funnel plot analyses to investigate publication bias are not recommended by the Cochrane Handbook instructions (9).

Authoritative sources such as the Cochrane Handbook (9) and PRISMA (19) guidelines emphasize that duplicate study selection and data extraction are important for minimizing the risk of bias in the selection of studies and the risk of errors while transcribing data from the original studies.





**Table 2 - Risk of bias of clinical trials on women's health published in the São Paulo Medical Journal between 2008 and 2012.**

Bias/Study	Nakamura 2008 et al. (15)	Gonçalves 2009 et al. (16)	Bernardes 2012 et al. (17)
Was the allocation sequence adequately generated?	No	Yes	Yes
Was allocation adequately concealed?	Unclear	Yes	Unclear
Was knowledge of the allocated intervention adequately prevented during the study? (patients and personnel)	Unclear	Yes	No
Was knowledge of the allocated intervention adequately prevented during the study? (outcome assessors)	Unclear	No	No
Were incomplete outcome data adequately addressed?	Yes	Yes	Yes
Are reports of the study free of suggestions of selective outcome reporting?	Yes	Yes	Yes
Was the study free of other problems that could put it at a high risk of bias?	Yes	Yes	Yes

"Yes" = low risk of bias; "Unclear" = unclear risk of bias (moderate risk); "No" = high risk of bias. According to the Cochrane recommendations (9), the answers to the first four items should be analyzed when performing the final classification of the study. When at least one of the answers to these items is "No," the study is classified as having a "High Risk of Bias"; when at least one of the answers to these items is "Unclear," the study is classified as having a "Moderate or Unclear Risk of Bias."

However, only two of the SRs published in the SPMJ (10,12) followed this recommendation. The lack of a list with excluded studies could be due in part to editorial policies and the need to limit the number of words per article. Only two SRs used appropriate methods to combine studies (12,14). However, it should be noted that the other three SRs were graded as "zero" because it was impossible to perform meta-analyses due to a lack of similar studies. This conservative approach to assessing the quality of published SRs has been used by other investigators performing similar evaluations (20).

The most frequent methodological flaws of the three CTs on women's health were a lack of sequence generation information, allocation concealment and patient blinding and personnel and/or outcomes assessors, all of which are considered critical bias risk domains. Although there are several tools to assess the risk of bias of CTs, including the Jadad scale (21) and the Delphi list (22), we opted for the Cochrane tool, which is widely used and internationally validated (9).

An implicit limitation of our study is that it assessed the reporting quality of the SRs and CTs published in a Brazilian evidence-based health journal and not necessarily the actual methodological quality of these studies. If we had contacted the original authors and asked for missing or unclear methodological details, it is possible that their studies could have been upgraded. Similarly, if the peer reviewers and editors of the journal had asked the researchers to address missing information before publishing their manuscripts, the final reporting quality of these eight studies would have likely been higher.

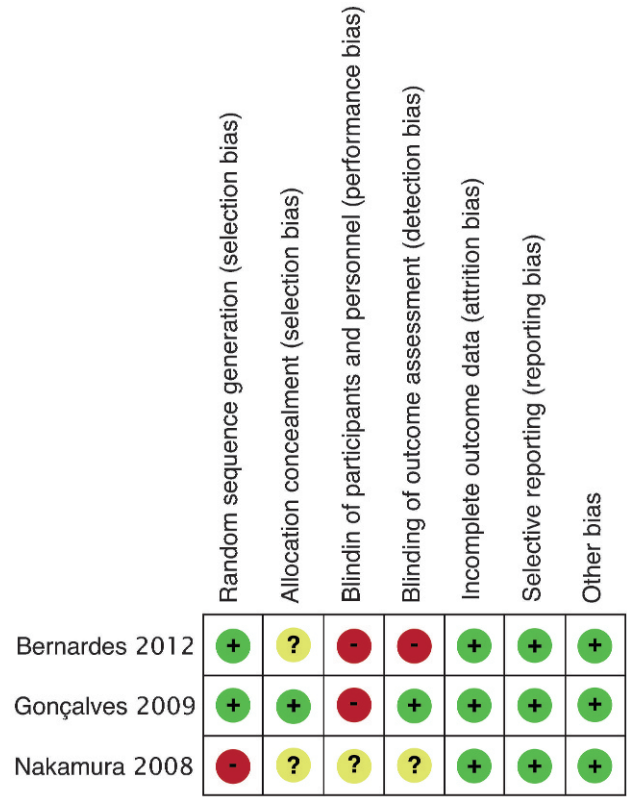
Due to the conclusions of this study, the "Instructions to Authors" section of the SPMJ has been modified and improved. Future authors who submit manuscripts for potential publication in the SPMJ are now required to follow internationally accepted guidelines, such as the PRISMA (19), CONSORT (23), STARD (24), MOOSE (25) or STROBE (26) recommendations, depending on the design of their study.

In summary, our findings indicate that most of the SRs and CTs on women's health recently published in a Brazilian evidence-based health journal are of low to moderate quality. As a result of this study, changes in the "Instructions to Authors" section have been made, and higher standards have been adopted for future volumes of this journal. To help improve the standards of our journals and to ensure that our readers are consulting studies of high methodological quality, we encourage other Brazilian scientific journals to perform a similar critical appraisal of the quality of the studies that they publish.

**Conflicts of Interest:** Torloni MR and Riera R are authors of two studies evaluated in this manuscript.

**AUTHOR CONTRIBUTIONS**

Macedo CR was responsible for evaluating the quality of the included studies and reviewing the manuscript. Riera R was responsible for the selection and evaluation of the included studies and for reviewing the



**Figure 1 - Summary of the risk of bias for clinical trials on women's health published in the São Paulo Medical Journal between 2008 and 2012.**



manuscript. Torloni MR was responsible for the selection of studies and drafting the manuscript. All authors declare that they have participated sufficiently in the work to take public responsibility for appropriate portions of the content.

## REFERENCES

- Scheffer M, Biancarelli A, Cassenote A. Demografia Médica no Brasil: dados gerais e descrições de desigualdades. Conselho Regional de Medicina do Estado de São Paulo e Conselho Federal de Medicina, 2011. 117p. Available from: <http://portal.cfm.org.br/images/stories/pdf/demografiamedicanobrasil.pdf>.
- IBGE. Censo demográfico 2010. Available from: <http://serieestatisticas.ibge.gov.br/series.aspx?vcodigo=POP101&sv=32&t=populacao-sexo-populacao-presente-residente>.
- Centre for Evidence Based Medicine. Oxford Centre for Evidence-based Medicine - Level of Evidence. Available from: <http://www.cebm.net/index.aspx?o=1025>.
- Torloni MR, Riera R. Design and level of evidence of studies published in two Brazilian medical journals recently indexed in the ISI Web of Science database. *Sao Paulo Med J.* 2010;128(4):202-5, <http://dx.doi.org/10.1590/S1516-31802010000400005>.
- Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7:10, <http://dx.doi.org/10.1186/1471-2288-7-10>.
- Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic review. *J Clin Epidemiol.* 2009; 62(10):1013-20.
- Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS One.* 2007;2(12):e1350, <http://dx.doi.org/10.1371/journal.pone.0001350>.
- Mikton C, Butchart A. Child maltreatment prevention: a systematic review of reviews. *Bull World Health Organ.* 2009;87(5):353-61, <http://dx.doi.org/10.2471/BLT.08.057075>.
- Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).
- Trevisani VFM, Riera R, Imoto AM, Saconato H, Atallah AN. Teriparatide (recombinant human parathyroid hormone 1-34) in postmenopausal women with osteoporosis: systematic review. *Sao Paulo Med J.* 2008;126(5):279-84.
- Riera R, de Soárez PC, Puga MES, Ferraz MB. Lapatinib for treatment of advanced or metastasized breast cancer: systematic review. *Sao Paulo Med J.* 2009;127(5):295-301, <http://dx.doi.org/10.1590/S1516-31802009000500009>.
- Iared W, Shigueoka DC, Torloni MR, Velloni FG, Ajzen SA, Atallah AN, et al. Comparative evaluation of digital mammography and film mammography: systematic review and meta-analysis. *Sao Paulo Med J.* 2011;129(4):250-60, <http://dx.doi.org/10.1590/S1516-31802011000400009>.
- Corrêa FM, Russomano FB, Oliveira CA. Colposcopic triage methods for detecting cervical intraepithelial neoplasia grade 3 after cytopathological diagnosis of low-grade squamous intraepithelial lesion: a systematic review on diagnostic tests. *Sao Paulo Med J.* 2012;130(1):44-52, <http://dx.doi.org/10.1590/S1516-31802012000100008>.
- Oliveira CA, Russomano FB, Júnior SCG, Corrêa FM. Risk of persistent high-grade squamous intraepithelial lesion after electrosurgical excisional treatment with positive margins: a meta-analysis. *Sao Paulo Med J.* 2012;13(2):119-25, <http://dx.doi.org/10.1590/S1516-31802012000200009>.
- Nakamura G, Ganem EM, Módolo NSP, Rugolo LMSS, Castiglia YMM. Labor analgesia with ropivacaine added to clonidine: a randomized clinical trial. *Sao Paulo Med J.* 2008;126(2):102-6.
- Gonçalves AV, Teixeira LC, Torresan R, Alvarenga C, Cabello C. Randomized clinical trial on the preservation of the medial pectoral nerve following mastectomy due to breast cancer: impact on upper limb rehabilitation. *Sao Paulo Med J.* 2009;127(3):117-21.
- Bernardes BT, Resende APM, Stüpp L, Oliveira E, Castro RA, di Bella ZIKJ, et al. Efficacy of pelvic floor muscle training and hypopressive exercises for treating pelvic organ prolapse in women: randomized controlled trial. *Sao Paulo Med J.* 2012;130(1): 5-9, <http://dx.doi.org/10.1590/S1516-31802012000100002>.
- Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003;3:25, <http://dx.doi.org/10.1186/1471-2288-3-25>.
- 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 healthcare interventions: explanation and elaboration. *BMJ.* 2009;339:b2700, <http://dx.doi.org/10.1136/bmj.b2700>.
- Seo HJ, Kim KU. Quality assessment of systematic reviews or metaanalyses of nursing interventions conducted by Korean reviewers. *BMC Medical Research Methodology.* 2012;12(1):129, <http://dx.doi.org/10.1186/1471-2288-12-129>.
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17(1):1-12, [http://dx.doi.org/10.1016/0197-2456\(95\)00134-4](http://dx.doi.org/10.1016/0197-2456(95)00134-4).
- Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol.* 1998;51(12):1235-41, [http://dx.doi.org/10.1016/S0895-4356\(98\)00131-0](http://dx.doi.org/10.1016/S0895-4356(98)00131-0).
- Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c332, <http://dx.doi.org/10.1136/bmj.c332>.
- Rennie D. Improving reports of studies of diagnostic tests: the STARD initiative. *JAMA.* 2003;289(1):89-90, <http://dx.doi.org/10.1001/jama.289.1.89>.
- Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283(15):2008-12.
- STROBE Statement. Strengthening the reporting of observational studies in epidemiology. What is strobe? Available from: <http://www.strobe-statement.org/>. Accessed in 2012 (Oct 15).