

Review

Evaluating the Quality of Randomized Controlled Trials that Examine the Efficacy of Natural Health Products: A Systematic Review of Critical Appraisal Instruments

Anne Marie Whelan, Tannis M. Jurgens and Lindsay Lord

College of Pharmacy, Dalhousie University, Dalhousie University, 5968 College Street, Halifax NS, Canada B3H 3J5

The purpose of this project was to conduct a systematic review to identify instruments designed to evaluate the quality of randomized controlled trials (RCTs) of natural health products (NHPs). Instruments were examined for inclusion of items assessing methods, identity and content of the NHP, generalizability of results and instructions for use. Online databases, websites, textbooks and reference lists were searched to identify instruments. Relevance assessment and data extraction of articles were completed by two investigators and disagreements were settled by the third investigator. Data were analyzed using descriptive statistics. Of the 4442 citations identified, 29 were potentially relevant with 16 meeting the criteria for inclusion. None of the instruments stated they were validated; content in the four areas of interest varied considerably. The most common items included randomization sequence generation (100%), blinding (100%), allocation concealment (75%) and participant flow (75%). Only nine of the NHP instruments included at least one item to appraise the specific content of the NHP. The CONSORT Statement for Herbal Interventions most closely addressed the four areas of interest; however, this instrument was specific for herbs. There is a need for the development of a validated instrument for assessment of the quality of RCTs that would be useful for herbs as well as other NHPs.

Keywords: Checklists – herbs – quality assessment

Introduction

Natural sourced medicines, such as herbal products, have become increasingly popular among consumers who search for ‘natural’ ways to maintain their health. In Canada, natural sourced medicines are called natural health products (NHPs) and are classified by Health Canada to include all herbal remedies, traditional medicines, homeopathic medicines, probiotics, amino acids, essential fatty acids, vitamins and minerals that are appropriate for over the counter selection (1).

Consumers often ask health care providers for guidance on the safe and effective use of NHPs.

When providing recommendations to consumers, health care providers are accustomed to following principles of evidence based medicine, which includes, in part, considering data from research reports that have been assessed on their quality of reporting of research methodology, design and analysis of data (2). Increasingly, randomized controlled trials (RCTs), considered the ‘gold standard’ when evaluating the efficacy of conventional medicines (3), are being used to evaluate the efficacy of NHPs (4). Therefore, it is essential for health care providers and their patients that reports of RCTs of NHPs be critically appraised with the same rigor as is applied to the assessment of the quality of RCTs of conventional medications.

For reprints and all correspondence: Dr Anne Marie Whelan, B.Sc. (Pharm), Pharm D, College of Pharmacy, Dalhousie University, 5968 College Street, Halifax NS, Canada B3H 3J5.
Tel: 902-494-3503; Fax: 902-494-1396;
E-mail: anne.marie.whelan@dal.ca

© 2008 The Author(s).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.0/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Two kinds of critical appraisal instruments have been developed specifically to aid the evaluation of the quality of evidence produced by RCTs of conventional medications (5,6). Many of these instruments focus on evaluating the methodological quality of the trial, examining how the design, conduct and analysis of the RCT minimize biases (e.g. selection, measurement) (7). Other instruments focus on the quality of the report itself; examining the information provided regarding the design, conduct and analysis of the results (8).

A critical appraisal instrument useful for evaluating the quality of evidence provided by RCTs of NHPs would include all components found in critical appraisal instruments for conventional medications, as well as questions to assess the quality of reporting of data that is specific to NHPs. Medicines from natural sources are predisposed to significant variation in chemical content as a consequence of how they are grown and/or processed. The resulting qualitative and/or quantitative variation in chemical content of NHPs can have a direct effect on their pharmacological effects and subsequently on the outcome of the RCT in which they are being evaluated (9,10). Adequate reporting of the identity and chemical content of the specific NHP that is being evaluated in the RCT is critical if the results of the trial are to be applicable to other brands or preparations of the NHP. Other factors that are unique to NHPs and can affect the quality of the results of an RCT, may be the appearance, smell and taste of the NHP. Therefore correctly blinding an RCT by creating a placebo that is identical in every way to the test NHP is very important. Thus, to adequately appraise the results of an RCT of an NHP, the optimal instrument would include items to assess the identity of the NHP (and its placebo) being evaluated. It would also include questions to help the user determine if the results of the RCT could be applied to other patients and to other products made from the same plant or other natural source. This ability to help the user determine the generalizability of the trial results would facilitate the translation of the research results for application to practice situations (11). It is also important that instructions for how to use the critical appraisal instrument are provided to help ensure that the user is interpreting the questions or items in the instrument correctly (12).

The purpose of this project was to conduct a systematic review to identify and analyze instruments designed for critically appraising RCTs that evaluate the efficacy of NHPs. Instruments were examined to determine if they adequately addressed four components: quality of methods, details of the identity and content of the NHP tested, generalizability of results to other patients and other preparations of the NHP and instructions for use of the instrument. Additionally they were examined to determine if the instruments had been validated. Thus, the first objective of this project was to identify

published critical appraisal instruments designed to assess the quality of RCTs evaluating the efficacy of NHPs. The second objective was to systematically review the appraisal instruments to document the inclusion of items in each instrument that examined the following: the methodological quality of the trial; the identity and content of the NHP and placebo; and the applicability of trial results to clinical practice. The third objective was to determine how each appraisal instrument had been developed, if it had been validated and if instructions for use were provided.

Methods

Retrieval of Articles

Database searches were performed (PubMed, Embase, Web of Science, CINAHL, IPA and Cochrane Library) up to September 2006 to locate potentially relevant articles describing the critical appraisal of RCTs of NHPs. Search terms were: critical appraisal, scale, checklist, instrument, tool, form, randomized controlled trial, natural health product, herbs, botanical, plant preparation, dietary supplement, complementary therapy, evidence based medicine, quality control, quality assessment and standards. Searches were limited to English language articles, with no restrictions on publication dates. Textbooks, files of the investigators and websites of evidence-based medicine (13) were examined for additional instruments. Finally, bibliographies of relevant references were examined for additional citations.

Relevance Assessment

All potentially relevant articles were retrieved and screened to determine relevance to study objectives by two investigators (T.J. and L.L.), using a form developed specifically for this purpose. Articles were deemed relevant if they met at least one of the following inclusion criteria: they contained instruments used for the appraisal of RCTs of NHPs; they detailed how to report RCTs of NHPs; or they discussed the importance of quality assessment of NHP trials and provided suggested items of importance. Articles were not considered relevant if they met at least one of the following exclusion criteria: the instrument was meant for the assessment of meta-analyses, systematic reviews, or any trial design other than RCT; the instrument was not published in a journal, textbook or on a website; or the instrument was designed for the assessment of RCTs of therapies other than NHPs, e.g. surgery. If the same instrument was identified from several resources, it was included only once.

Data Extraction

A form was developed to aid data extraction. The first part of the form was a checklist comprised of items contained in the CONSORT Statements, including those items specific to herbs found in the CONSORT Statement for Herbal Intervention (14). Using the CONSORT Statements provided a good starting point for the data extraction form as they consist of lists of criteria deemed to be important in the reporting of an RCT such that a reader will understand how a trial was conducted and be able to assess the validity of results. Additionally, the CONSORT Statements provide detailed explanation of the meaning of the criteria used, which would aid the accurate extraction of data (12,14,15). Following the itemized extraction portion, the data extraction form contained an area for recording additional items found in instruments but not contained in the CONSORT Statements. These were termed miscellaneous items.

In the second section of the data extraction form, the articles were categorized by type of instrument (checklist, rating scale, numeric scale, combination of checklist and scale) or as a guidance document. The term guidance document was adapted from a previous publication and in this project was intended to capture articles that described how to report an RCT or discussed the importance of assessing publication quality and provided example items for consideration (7). The method used to develop the instrument, any process used for validation and provision of instructions for use, were also recorded. Two investigators (A.M.W. and L.L.) independently completed data extraction for all articles. Disagreements were resolved by the third investigator (T.J.).

Analysis

Data were analyzed using descriptive statistics. The identified instruments were separated into categories based on whether they were a type of instrument (checklist, numeric scale, rating scale, combination checklist and scale) or a guidance document. A list of items contained in instruments was compiled and the number of instruments that included each item was tabulated. Miscellaneous items were also tabulated. Methods of development and validation were examined to identify the most common processes used. References cited as a source of instrument development were reviewed to determine if the same pool of instruments were used as a basis for instrument development.

Results

Article Retrieval and Relevance Assessment

Online database searches and hand searching of investigators files, websites, textbooks and reference

lists identified 4442 citations (Fig. 1). Of these, 4032 citations were excluded after initial review indicated they were not pertinent to the project. Review of the remaining 410 abstracts identified a total of 29 potentially relevant articles, 16 of which met the inclusion criteria (14,16–30). Of the 16 instruments 6 were checklists, 3 combinations of a scale and a checklist, 3 rating scales and 4 numeric scales. No guidance documents specific to NHPs were identified.

Data Extraction

Items

In the 16 instruments developed specifically for use with NHPs, the most common items included randomization sequence generation (100%), blinding (100%), allocation concealment (75%), participant flow (75%) and baseline data (50%) (Fig. 2). Nine of these instruments included at least one item to appraise the identity or content of the NHP (14,17,18,20–22,25,29,30). Only 4 of the 16 instruments included items pertaining to the generalizability of the results (14,18,20,27). Twelve items not found in the CONSORT Statements were extracted from the instruments for NHPs and recorded as miscellaneous items (Table 1).

Instrument Development, Validation and Guidelines for Use

A description of the method used to develop the instrument was provided for a total of 12 of the 16 instruments (75%). Eleven of the instruments that reported method of development were adapted or modified from one or more existing instruments and one was developed as a result of a discussion of experts. The most common references cited as sources for modification/adaptation for new instrument development are listed in Table 2.

None of the 16 instruments stated that they had been validated. Instructions for using critical appraisal instruments were included with four of the identified instruments (14,18,21,25).

Discussion

As anticipated, there was a paucity of instruments for the critical appraisal of RCTs evaluating the efficacy of NHPs. The CONSORT Statement for Herbal Interventions (14) most closely addressed the four components (quality of methods, details of the identity and content of NHP tested, generalizability of results and instructions for use) felt to be important in such an instrument. However, the CONSORT Statement for Herbal Interventions had two limitations. It was designed for use exclusively with herbal medicines, and

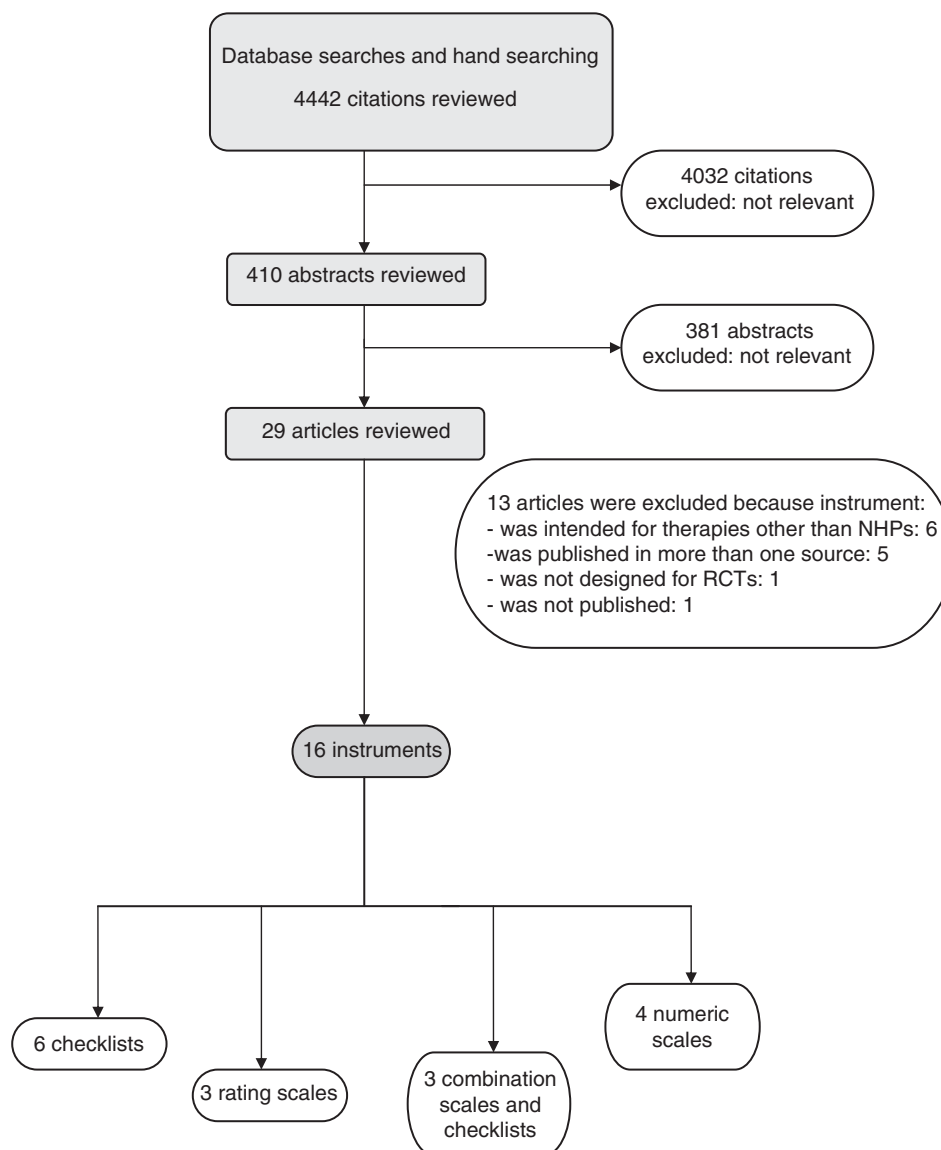


Figure 1. Results of literature search and relevance assessment.

therefore was not useful for the other non-herbal types of NHPs and there was no indication that the CONSORT Statement for Herbal Interventions had been validated for use as an instrument for critically appraising an RCT.

There was a wide variety in the number and wording of the items included in the identified appraisal instruments. As well as the items contained in the CONSORT Statement for Herbal Interventions, an additional 12 unique miscellaneous items were extracted from the instruments. This illustrated the lack of apparent consensus on which items are essential for inclusion in a critical appraisal instrument. Encouragingly, there was evidence that the most common items found (e.g. randomization sequence generation (31) blinding (31–34), allocation concealment (31–33) and participant

flow (35–37) have been shown to be important in minimizing bias in trials and/or in improving the quality of reporting. Additionally, each of these items was part of validated instruments used with conventional medications (38–41). This provides good rationale for inclusion of items in critical appraisal instruments.

Surprisingly, seven of the 16 instruments developed for use with NHPs did not contain any items to address the identity and content of the NHPs being tested. An adequate appraisal of the design and results of an RCT of an NHP must include an assessment of the information provided in the report of the RCT detailing the identity and chemical content of the NHP tested in the trial. Documenting the name of the NHP that was evaluated is not always sufficient to ensure unambiguous identity of the NHP, as it is well known that the chemical

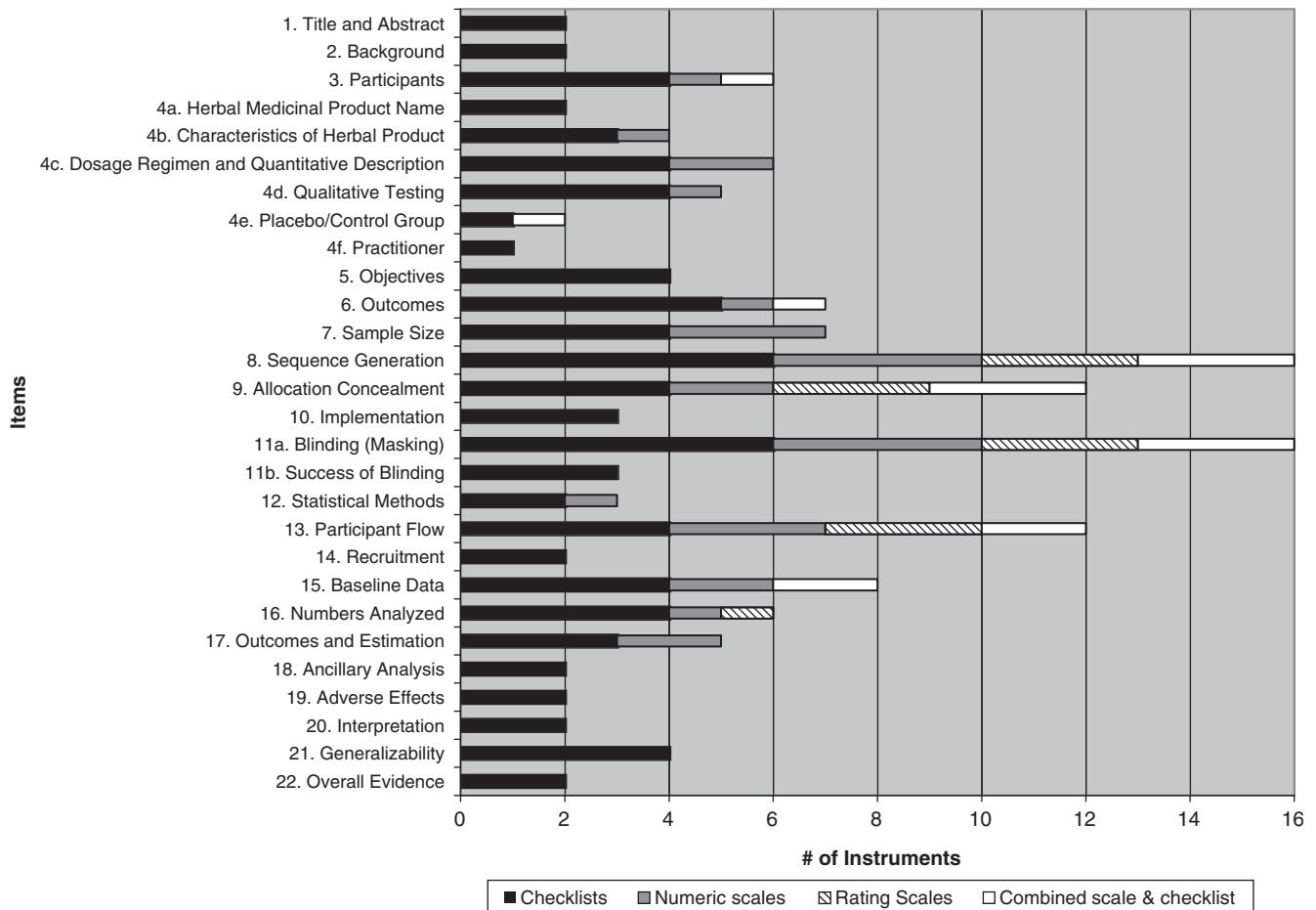


Figure 2. Frequency of items included in instruments for assessing quality of RCTs of NHPs.

content of some NHPs can vary significantly between preparations, thus affecting the pharmacological effects of the product (9,10). Most importantly, documenting the detailed identity of the NHP used in the RCT allows an interested person to select a product of similar content and be able to expect similar outcomes.

Since there did not appear to be a gold standard appraisal instrument for RCTs of NHPs, it was important to determine how instruments were developed and how items were chosen for inclusion in instruments. Approximately 75% of the articles that were identified provided information about how the instrument was developed. Of those who reported their development techniques, almost 70% were adapted or modified from existing instruments. The instrument that was published by Jadad *et al.* (40). was most commonly used as a template for the development of other instruments. Some of the references cited were used to develop only one other instrument. This may suggest that authors are selecting their preferred existing instrument and then modifying it to meet their needs. However, in the new instruments developed from previous instruments, there was little specific information provided as to why an item was

included/excluded or how/why the item was modified. Thus there appeared to be minimal empirical basis for the development of many of the instruments. Although none of the instruments stated that they were validated, six of the instruments were adapted from the validated instrument published by Jadad *et al.* (40). However, in most cases, it was the non-validated instruments that were used as a starting point for the development of new instruments, meaning that potential problems with existing instruments are being perpetuated in the newly created instruments. It was discouraging to see that only 25% of the identified instruments included some instructions for use. Guidance in using the instrument as the developers intended should provide some consistency in the interpretation of items and therefore less variability in results obtained from the use of the instrument.

Strengths and Limitations

To our knowledge this is the first article to systematically review the literature to identify instruments for the critical appraisal of RCTs evaluating efficacy of NHPs.

Table 1. Miscellaneous items that appeared in instruments for critical appraisal of RCTs of NHPs

Item	Number of times item appeared
Enough detail for reanalysis by reader/study replicable	3
Intervention described in detail	2
Groups treated equally	1
Compliance	1
Indication for further intervention	1
Components known to be predictive of clinical effect?	1
Possibility of publication bias	1
Is journal peer-reviewed?	1
Informed consent/ethics approval	1
Syndrome of disease based on Chinese herbal medicine	1
Rationale of Chinese herbal medication composition	1
Selection bias after allocation	1

Table 2. References^a cited most frequently as source for adapting/modifying for instrument development

Reference	Number of times cited
Jadad <i>et al.</i> (40)	6 (38%)
Moher <i>et al.</i> (revised CONSORT) (15)	3 (19%)
JAMA Users' guides 1993–94 (11,42,43)	3 (19%)
Schulz <i>et al.</i> (44)	2 (13%)
Cochrane handbook (45)	2 (13%)
Juni <i>et al.</i> (46)	2 (13%)
Meinert <i>et al.</i> (47)	1 (6%)
Linde <i>et al.</i> (48)	1 (6%)

^aSome articles reported more than one reference used as a source.

Although appraisal instruments for other types of study design exist, we limited our review to only instruments designed to assess quality of RCTs as the efficacy of NHPs are being increasingly evaluated using this methodology.

Although we followed systematic review methodology as rigorously as possible when conducting this research we did encounter difficulties particularly with regards to literature searching. Selecting search terms to use to identify relevant articles in various databases proved problematic, as indexing terms varied considerably between databases. It is possible that some instruments may not have been identified as search terms were not exhaustive. It also appeared as though MeSH headings were not very effective at identifying desired articles. As a result, text word searching was used, which resulted in many hits that were deemed to be irrelevant to the research. For example, using the term 'instrument'

resulted in articles that pertained to implements/tools used for surgery or other procedures and were not at all relevant to our study. Similar difficulties were also encountered by the Agency for Healthcare Research and Quality (7). We followed their suggestions of not relying only on literature searches, and used published reports, such as theirs, to identify existing instruments. In the end, we did find that many articles, not identified through database searching, were located by hand searching the reference lists of other articles.

Due to resource constraints we were not able to search non-English literature or contact other researchers/experts in this field for identification of other potential instruments. It is possible that instruments were missed by not searching in languages other than English. However, we believe we were able to identify many of the commonly used instruments, as the same instruments began to appear repeatedly in the databases we searched and in the hand searching of reference lists. These are likely the same instruments that would also be found by any health care provider trying to find a critical appraisal instrument for use in practice or research. Thus, despite the limitations we have described, we believe that we were able to identify and review the most readily available assessment instruments for RCTs of NHPs.

Conclusion

Although this systematic review was successful in identifying appraisal instruments for evaluating the quality of RCTs of NHPs, none of those identified were widely accepted as the gold standard instrument. The CONSORT Statement for Herbal Interventions comes closest to fulfilling requirements for a quality instrument to critically appraise RCTs of NHPs, however it's limitation is that it is specific for herbal medicines. There is therefore a need for the development and validation of an instrument that can be used to reliably assess the quality of results of RCTs of any NHPs, whether herbs or single, natural sourced chemicals. Empirical evidence and results of this review suggest that randomization sequence generation, allocation concealment, blinding and participant flow are four elements that improve the quality of an RCT and should be included as items in a critical appraisal instrument. Items pertaining specifically to the identity and content of NHPs as well as applicability of results to practice must be included. It remains to be determined if a group of items specific to NHPs could be developed, validated and added to an existing gold standard instrument used for assessing quality of trials of conventional medication or whether an instrument specific for assessing trials of NHPs will need to be developed.

Acknowledgements

Funding for the project was provided by the Dalhousie Pharmacy Endowment Fund and Studentship funding was provided by the Canadian Institutes of Health Research. Many thanks are extended to Ms Elizabeth Foy for her assistance in developing search strategies and to Dr David Gardner and Melissa MacDonald for helping with early project development.

References

- Natural health products directorate [homepage on the Internet]. Available from: http://www.hc-sc.gc.ca/dhp-mps/prodnatur/index_e.html (last accessed date December 13, 2007).
- Chiappelli F, Prolo P, Rosenblum M, Edgerton M, Cajulis OS. Evidence-based research in complementary and alternative medicine II: the process of evidence-based research. *Evid Based Complement Alternat Med* 2006;3:3–12.
- Sackett DL, Richardson WS, Roseberg W, Haynes RB. *Evidence-based Medicine: How to Practice and Teach*. New York: Churchill Livingstone, 1997.
- Gagnier JJ, DeMelo J, Boon H, Rochon P, Bombardier C. Quality of reporting of randomized controlled trials of herbal medicine interventions. *Am J Med* 2006;119:e800–11.
- Glenny AM. No “gold standard” critical appraisal tool for allied health research. *Evid Based Dent* 2005;6:100–1.
- Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol* 2004;4:22.
- Agency for healthcare research and quality: Systems to rate the strength of scientific evidence. Rockville: 2002. Report No.: Evidence Report/Technology Assessment No.47, Publication No.02-E016.
- Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials* 1995;16:62–73.
- Busse W. The significance of quality for efficacy and safety of herbal medicinal products. *Drug Inf J* 2000;34:15–23.
- Bauer R. Quality criteria and standardization of phytopharmaceuticals: can acceptable drug standards be achieved? *Drug Inf J* 1998;32:101–10.
- Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-based medicine working group. *JAMA* 1994;271:59–63.
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–94.
- Drug information resources: a guide for pharmacists [homepage on the Internet]. Available from: <http://dir.pharmacy.dal.ca> (last accessed date September, 2006).
- Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C, et al. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med* 2006;144:364–7.
- Moher D, Schulz KF, Altman DG CONSORT GROUP (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001;134:657–62.
- Bian Z, Wu T, Liu L, Miao J, Wong H, Song L, et al. Effectiveness of the chinese herbal formula TongXieYaoFang for irritable bowel syndrome: a systematic review. *J Altern Complement Med* 2006;12:401–7.
- Bian ZX, Li YP, Moher D, Dagenais S, Liu L, Wu TX, et al. Improving the quality of randomized controlled trials in chinese herbal medicine, part I: clinical trial design and methodology. *Zhong Xi Yi Jie He Xue Bao* 2006;4:120–9.
- Bian ZX, Moher D, Dagenais S, Li YP, Wu TX, Liu L, et al. Improving the quality of randomized controlled trials in chinese herbal medicine, part IV: applying a revised CONSORT checklist to measure reporting quality. *Zhong Xi Yi Jie He Xue Bao* 2006;4:233–42.
- Chen XY, Wu TX, Liu GJ, Wang Q, Zheng J, Wei J, et al. Chinese medicinal herbs for influenza. *Cochrane Database of Systematic Reviews* 2005, Issue 1. Art No.: CD004559. DOI: 10.1002/14651858.CD004559. pub 2.
- Gardner DM. Evidence-based decisions about herbal products for treating mental disorders. *J Psychiatry Neurosci* 2002;27:324–33.
- Kleijnen J, Mackerras D. Vitamin E for intermittent claudication. *Cochrane Database of Systematic Reviews* 1998, Issue 1. Art. No.: CD000987. DOI: 10.1002/14651858.CD 000987.
- Kleijnen J, Knipschild P, ter Riet G. Clinical trials of homeopathy. *Br Med J* 1991;302:316–23.
- Linde K, Ramirez G, Mulrow CD, Pauls A, Weidenhammer W, Melchart D. St john's wort for depression: an overview and meta-analysis of randomised clinical trials. *Br Med J* 1996;313:253–8.
- Linde K, Berner M, Egger M, Mulrow C. St john's wort for depression: Meta-analysis of randomised controlled trials. *Br J Psychiatry* 2005;186:99–107.
- Low Dog T. Clinical trial reviewer's guidance and checklist. In: Barrett M (ed). *The Handbook of Clinically Tested Herbal Remedies*, Vol. 1. Birmingham, New York: The Haworth Press, Inc, 2004, 141–8.
- Qiong W, Yiping W, Jinlin Y, Tao G, Zhen G, Pengcheng Z. Chinese medicinal herbs for acute pancreatitis. *Cochrane Database of Systematic Reviews* 2005, Issue 1. Art. No.: CD003631. DOI: 10.1002/14651858.CD 003631. pub 2.
- Sarubin A. *The Health Professional's Guide to Popular Dietary Supplements*. United States: The American Dietetic Association, 2000.
- Taixiang W, Munro AJ, Guanjian L. Chinese medical herbs for chemotherapy side effects in colorectal cancer patients. *Cochrane Database of Systematic Reviews* 2005, Issue 1. Art. No.: CD004540. DOI: 10.1002/14651858.CD004540. pub 2.
- Zhang W, Leonard T, Bath-Hextall F, Chambers CA, Lee C, Humphreys R, et al. Chinese herbal medicine for atopic eczema. *Cochrane Database of Systematic Reviews* 2004, Issue 4. Art. No.: CD002291. DOI: 10.1002/14651858.CD002291.pub 3.
- Hemilä H, Chalker E, D'Souza RRD, Douglas RM, Treacy B. Vitamin C for preventing and treating the common cold. *Cochrane Database of Systematic Reviews* 2004, Issue 2. Art. No.: CD004782. DOI: 10.1002/14651858.CD004782.
- Kjaergard LL, Villumsen J, Gluud C. Quality of randomised clinical trials affects estimates of intervention efficacy. *Proceedings of the 7th cochrane colloquium*. 1999. Rome: Universita S. Tommaso D'Aquino, 1999, 57.
- Schulz KF. Subverting randomization in controlled trials. *JAMA* 1995;274:1456–8.
- Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *Br Med J* 2001;323:42–6.
- Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994;44:16–20.
- Egger M, Juni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996–9.
- May GS, Chir B, Demets DL, Friedman LM, Furberg C, Passamani E. The randomized clinical trial: bias in analysis. *Circulation* 1981;64:669–73.
- Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *New Engl J Med* 1979;301:1410–2.
- Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377–84.
- Sindhu F, Carpenter L, Seers K. Development of a tool to rate the quality assessment of randomized controlled trials using a delphi technique. *J Adv Nurs* 1997;25:1262–8.

40. 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;17:1–12.
41. Cho MK, Bero LA. Instruments for assessing the quality of drug studies published in the medical literature. *JAMA* 1994;272:101–4.
42. Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. how to get started. the evidence-based medicine working group. *JAMA* 1993;270:2093–5.
43. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. how to use an article about therapy or prevention. A. are the results of the study valid? evidence-based medicine working group. *JAMA* 1993;270:2598–601.
44. Schulz KF, Chalmers I, Hayes JA, Altman DG. Empirical evidence of bias. dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–12.
45. Alderson P, Green S, Higgins JPT (eds) *Cochrane reviewers' handbook 4.2.2*, 2004.
46. Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *Br Med J* 2001;323:42–6.
47. Meinert CL. Questions, & factors to consider when reading a report from a clinical trial. In: Meinert C (ed). *Clinical Trials: Design, Conduct, and Analysis*. New York: Oxford University Press, 1986, 272–6.
48. Linde K, Jonas WB, Melchart D, Willich S. The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture. *Int J Epidemiol* 2001;30:526–31.

Received May 23, 2007; accepted December 21, 2007