Reviewer's report

**Title:** Development and evaluation of an instrument for the critical appraisal of randomized controlled trials of natural products.

**Version:** 1  **Date:** 30 January 2009

**Reviewer:** Benjamin Djulbegovic

**Reviewer's report:**

This is an ambitious project attempting to develop a new instrument for the critical appraisal of RCT on natural products (NPs). The authors first performed systematic review of all existing instruments for critical appraisal of NP RCTs. (This in itself is the most important contribution). The authors then used Delphi method to identify items deemed essential in describing NPs. They then field tested their instrument against existing ones.

While the overall approach is solid, I believe that the authors need to make some clarification in the paper (perhaps, under Limitation Section);

- The most instruments they identified pertain to the quality of REPORTING standards (e.g. CONSORT authors themselves warned that the CONSORT checklist should not be used to assess the quality of research)

- The authors should discuss/acknowledge the difference between the quality of reporting and risk for bias (the latest Cochrane Handbook provides the useful discussion on this issue + identify the items that are empirically linked to bias)

- Related to the same issue, the statement that “4 of 99 instruments have been validated’ (page 13) is actually not correct (e.g. Jadad’s score has been abandoned as its scale did not show usefulness in several studies as it was case with general scale-based instruments etc)

- I found description of the field experiment lacking and ironically the authors did not follow their own advice in designing/reporting this trial (e.g. the primary outcome was not well defined, sample size was not provided etc )

- The term “double-blind” is now discouraged (instead one should specify who was “blinded” to what and why; Gordon Guyatt’s group discussed this in JAMA article several years ago)

- The authors should distinguish evidence from decision-making ; in other words, the quality of appraisal does and should not be in a straightforward fashion translated into application of evidence (items #28). Many factors play role in making recommendation (e.g. benefits/harms ratio, other relevant evidence, preferences etc). I think the item #28 should be dropped from their checklist.

- I am surprised that pharmacological/biological plausibility with respect to NP mechanism of action has not been included in the checklist. For example, most
people found homeopathy trials not credible because of the lack of solid scientific theory.

- Finally, please include definition of CAM and NP medicine.

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

'I declare that I have no competing interests’