Come si Misura la Qualità di una Revisione Sistematica

MARIO NEGRI ISTITUTO DI RICERCHE FARMACOLOGICHE

Ivan Moschetti– Negrar – Marzo/2013

Lab. Metodologia per la Ricerca Biomedica;
Unità di metodologia delle revisioni sistematiche e produzione di linee guida

Esistono **molti** strumenti per misurare la qualità delle RS.

- La maggior parte **NON sono stati validati** e **spesso NON sono chiari** nella distinzione tra:

    1) **methodological quality**
    how well the SR was conducted

    2) **reporting quality**
    how well reviewers have reported
    their methodology and findings

Ivan Moschetti– Negrar – Marzo/2013

AMSTAR: 11 Items

A measurament tool to assess systematic reviews

PRISMA: 27 Items

Preferred Reporting Items for Systematic Reviews and Meta-Analyses

A systematic review is..

…a review of a **clearly formulated question** that uses **systematic** and **explicit** methods **to identify, select,** and **critically appraise** relevant **research,** and to **collect** and **analyze data** from the **studies** that are **included** in the review.

Statistical methods **(meta-analysis)** may or may not be used **to analyze** and **summarize** the results of the included studies.

AMSTAR is a measurement tool created **to assess the methodological quality** of systematic reviews.

The aim of the PRISMA is to help authors improve the **reporting** of SR and meta-analyses...

...**may also** be useful for **critical appraisal** of SR ± MA

...however, the **PRISMA checklist is not a quality assessment instrument** to gauge the quality of a systematic review.

AMSTAR

- **Richiede** una media di **15 minutes** a revisione
- **È stata prodotta** a partire da 37 Items testati su 151 revisioni
  - Factor analysis 37->11 buona face e content validity
  - Media
- **È stata validata** su 42 revisioni
  - "inter-observer agreement" moderato-elevato
  - "reliability of total AMSTAR score" eccellente

1. Was an 'a priori' design provided?

- The research question and inclusion criteria should be established before the conduct of the review.

- □ Yes
- □ No
- □ Can't answer
- □ Not applicable

2. Was there duplicate study selection and data extraction?

- There should be at least **two** independent data **extractors** and a consensus procedure for disagreements should be in place.

- □ Yes
- □ No
- □ Can't answer
- □ Not applicable

3. Was a comprehensive literature search performed?

- At least **two** electronic sources should be searched. The report must include years and **databases** used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. **All searches should be supplemented** by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.

- □ Yes □ No □ Can't answer □ Not applicable

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?

- The authors should state that they searched for reports **regardless of their publication type.** The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.

- ☐ Yes ☐ No ☐ Can't answer ☐ Not applicable

5. Was a list of studies (included and excluded) provided?

- **A list** of included and excluded studies should be provided.

- □ Yes
- □ No
- □ Can't answer
- □ Not applicable

6. Were the characteristics of the included studies provided?

- In an aggregated form such as a table, data from the original studies should be provided on the **participants**, **interventions** and **outcomes**. The ranges of characteristics in all the studies analyzed e.g. **age**, **race**, **sex**, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.

- □ Yes □ No □ Can't answer □ Not applicable

7. Was the scientific quality of the included studies assessed and documented?

- **A priori** methods of assessment should be provided **(**e.g., for effectiveness studies if the authors chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria**)**; for other types of studies alternative items will be relevant.

Risk of Bias Evaluation

- □ Yes □ No □ Can't answer □ Not applicable

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?

- The **results** of the **methodological rigor** and scientific quality should be considered in the **analysis** and the **conclusions** of the review, and explicitly stated in formulating recommendations.

  □ Yes □ No □ Can't answer □ Not applicable

9. Were the methods used to combine the findings of studies appropriate?

- For the pooled results, **a test** should be done **to ensure** the **studies** were **combinable**, to assess their homogeneity (i.e. Chi-squared test for homogeneity, I^2). If **heterogeneity** exists a **random effects model** should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?).

□ Yes □ No □ Can't answer □ Not applicable

10. Was the likelihood of publication bias assessed?

- An assessment of **publication bias** should include a combination of graphical aids (e.g., **funnel plot**, other available tests) and/or statistical tests (e.g., Egger regression test).

  □ Yes □ No □ Can't answer □ Not applicable

11. Was the conflict of interest included?

- Potential sources of support should be clearly acknowledged **in both** the systematic **review** and the included **studies**.

  □ Yes □ No □ Can't answer □ Not applicable

AMSTAR – External Validation

Table 1. Assessment of the inter-rater agreement for AMSTAR

Items	Kappa (95% CI)	PHI Φ
1. Was an 'a priori' design provided?	0.75 (0.55 to 0.96)	0.76
2. Was there duplicate study selection and data extraction?	0.81 (0.63 to 0.99)	0.83
3. Was a comprehensive literature search performed?	0.88 (0.73 to 1.00)	0.89
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	0.64 (0.40 to 0.88)	0.64
5. Was a list of studies (included and excluded) provided?	0.84 (0.67 to 1.00)	0.84
6. Were the characteristics of the included studies provided?	0.76 (0.55 to 0.96)	0.76
7. Was the scientific quality of the included studies assessed and documented?	0.90 (0.77 to 1.00)	0.91
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	0.51 (0.25 to 0.78)	0.56
9. Were the methods used to combine the findings of studies appropriate?	0.80 (0.63 to 0.99)	0.80
10. Was the likelihood of publication bias assessed?	0.85 (0.64 to 1.00)	0.85
11. Were potential conflicts of interest included?	1.00 (100% no)	1.00
Overall Score	0.84 (0.67 to 1.00)	0.85

doi:10.1371/journal.pone.0001350.t001

External Validation of AMSTAR; PLoS ONE | www.plosone.org 4 December 2007 | Issue 12 | e1350

Conclusioni

- Le scale per la valutazione della qualità sono strumenti vivi e in continua evoluzione

- La soggettività nel giudizio non è eliminabile

- Meta-analysis RCT ->metodologia abbastanza solida, altri disegni di studio-> Still in the infancy