#### Item 12: RISK OF BIAS IN INDIVIDUAL STUDIES.

Describe methods used for assessing risk of bias in individual studies (including specification of whether this was done at the study or outcome level, or both), and how this information is to be used in any data synthesis.

**Example.** “To ascertain the validity of eligible randomized trials, pairs of reviewers working independently and with adequate reliability determined the adequacy of randomization and concealment of allocation, blinding of patients, health care providers, data collectors, and outcome assessors; and extent of loss to follow-up (i.e. proportion of patients in whom the investigators were not able to ascertain outcomes).”

“To explore variability in study results (heterogeneity) we specified the following hypotheses before conducting the analysis. We hypothesised that effect size may differ according to the methodological quality of the studies.”

#### Explanation.

The likelihood that the treatment effect reported in a systematic review approximates the truth depends on the validity of the included studies, as certain methodological characteristics may be associated with effect sizes . For example, trials without reported adequate allocation concealment exaggerate treatment effects on average compared to those with adequate concealment. Therefore, it is important for authors to describe any methods that they used to gauge the risk of bias in the included studies and how that information was used . Additionally, authors should provide a rationale if no assessment of risk of bias was undertaken. The most popular term to describe the issues relevant to this item is “quality,” but for the reasons that are elaborated in [Box 4](http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000100#pmed-1000100-box004) we prefer to name this item as “assessment of risk of bias.”

Many methods exist to assess the overall risk of bias in included studies, including scales, checklists, and individual components . As discussed in [Box 4](http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000100#pmed-1000100-box004), scales that numerically summarize multiple components into a single number are misleading and unhelpful . Rather, authors should specify the methodological components that they assessed. Common markers of validity for randomized trials include the following: appropriate generation of random allocation sequence ; concealment of the allocation sequence ; blinding of participants, health care providers, data collectors, and outcome adjudicators ; proportion of patients lost to follow-up ; stopping of trials early for benefit ; and whether the analysis followed the intention-to-treat principle . The ultimate decision regarding which methodological features to evaluate requires consideration of the strength of the empiric data, theoretical rationale, and the unique circumstances of the included studies.

Authors should report how they assessed risk of bias; whether it was in a blind manner; and if assessments were completed by more than one person, and if so, whether they were completed independently . Similarly, we encourage authors to report any calibration exercises among review team members that were done. Finally, authors need to report how their assessments of risk of bias are used subsequently in the data synthesis (see Item 16). Despite the often difficult task of assessing the risk of bias in included studies, authors are sometimes silent on what they did with the resultant assessments . If authors exclude studies from the review or any subsequent analyses on the basis of the risk of bias, they should tell readers which studies they excluded and explain the reasons for those exclusions (see Item 6). Authors should also describe any planned sensitivity or subgroup analyses related to bias assessments (see Item 16).