

Supplement 2. Methodological quality appraisal criteria and question weight

| Systematic review methodological quality criteria (question weight) |
|--|
| Q1. Is the research question/objective clearly stated? (1) |
| Q2. Was a primary hypothesis clearly stated? (1) |
| Q3. Were comparison group(s) used? (1) |
| Q4. Was an intervention allocation randomized? (3) |
| Q5. Was allocation concealment adequate? (1) |
| Q6. Were sample inclusion/exclusion criteria adequately described? (2) |
| Q7. Was recruitment (or participation) rate reported and adequate? (1) |
| Q8. Was the sampling frame representative of the target population? (1) |
| Q9. Did the author(s) examine whether important differences existed between those who participated and those who did not? (1) |
| Q10. Were pre-intervention (baseline) characteristics described and equivalent between groups? (2) |
| Q11. Was loss to follow up (attrition) significant and problematic? (1) |
| Q12. Did attrition affect groups equally? (1) |
| Q13. Did the author(s) compare post intervention those who dropped out to those who remained in the study? (1) |
| Q14. Was the intervention process adequately described to allow for replication? (3) |
| Q15. Was there any potential for contamination or co-intervention? (1) |
| Q16. Were the effects of the intervention on key intervention mechanisms described? (3) |
| Q17. Was compliance with the intervention in all groups described and adequate? (1) |
| Q18. Was blinding of participants and personnel adequate? (1) |
| Q19. Were the instruments used to assess the outcomes valid and reliable? (3) |
| Q20. Were all important outcomes described at baseline and follow-up? (3) |
| Q21. Was the proportion of missing outcome data greater than 40% or substantially different between groups? (2) |
| Q22. Was there adjustment for important covariates, confounders or pre-intervention differences (if necessary)? (2) |
| Q23. Were the statistical analyses optimized for best results? (2) |
| Q24. Were all participants' outcomes analyzed by the groups to which they were originally allocated (intention-to-treat analysis)? (2) |
| Q25. Was there a direct between group comparison? (3) |