**Supplementary material 1:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **RCT study design** | **Steadman (2001)** | **Swanson (2000/****2001)** | **Cohort study design** | **Gilbert (2010)** | **Hiday (1978)** | **Ogilvie (2022)** | **Phelan (2010)** | **Pollack (2005)** | **Power (1992)** | **Segal (2019)** | **Segal (2023)** | **Link (2011)** |
| 1) Was true randomisation used for assignment of participants to treatment groups? (Yes, No, Unclear, N/a) | Yes | Unclear | 1) Were the two groups similar and recruited from the same population? | Yes | Yes | Yes | Yes | Yes | Unclear | Yes | Yes | Yes |
| 2) Was allocation to treatment groups concealed? (Yes, No, Unclear, N/a) | Yes | Unclear | 2) Were the exposures measured similarly to assign people to both exposed and unexposed groups? | Yes | Yes | Yes | Yes | Yes | Unclear | Yes | Yes | Yes |
| 3) Were treatment groups similar at the baseline? (Yes, No, Unclear, N/a) | Yes | Unclear | 3) Was the exposure measured in a valid and reliable way? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 4) Were participants blind to treatment assignment? (Yes, No, Unclear, N/a) | No | No | 4) Were confounding factors identified? | Yes | No | Yes | Yes | Yes | Yes | Yes | No | Yes |
| 5) Were those delivering the treatment blind to treatment assignment? (Yes, No, Unclear, N/a) | No | No | 5) Were strategies to deal with confounding factors stated? | Yes | No | Yes | Yes | Unclear | Yes | Yes | No | Yes |
| 6) Were treatment groups treated identically other than the intervention of interest? (Yes, No, Unclear, N/a) | Yes | Yes | 6) Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | Yes | Unclear | Yes | Unclear | Yes | Unclear | Yes | Yes | Unclear |
| 7) Were outcome assessors blind to treatment assignment? (Yes, No, Unclear, N/a) | Yes | Unclear\* | 7) Were the outcomes measured in a valid and reliable way? | Yes | Yes | Yes | Yes | Yes | Unclear | Yes | Yes | Yes |
| 8) Were outcomes measured in the same way for treatment groups? (Yes, No, Unclear, N/a) | Yes | Yes | 8) Was the follow up time reported and sufficient to be long enough for outcomes to occur? | Yes | Unclear | Yes | Yes | Yes | Yes | Unclear | Unclear | Yes |
| 9) Were outcomes measured in a reliable way? (Yes, No, Unclear, N/a) | Yes | Unclear | 9) Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | Yes | Yes | Yes | Yes | Yes | Unclear | Yes | Yes | Yes |
| 10) Was follow-up complete and if not, were differences between groups in terms of their follow-up adequately described and analysed? (Yes, No, Unclear, N/a) | Yes | Yes | 10) Were strategies to address incomplete follow up utilized? | N/a\*\* | N/a\*\* | N/a\*\* | N/a\*\* | N/a\*\* | N/a\*\* | N/a\*\* | N/a\*\* | N/a\*\* |
| 11) Were participants analysed in the groups to which they were randomised? (Yes, No, Unclear, N/a) | Yes | Yes | 11) Was appropriate statistical analysis used? | Yes | No | Yes | Yes | Unclear | Unclear | Yes | No | Yes |
| 12) Was appropriate statistical analysis used? (Yes, No, Unclear, N/a) | Yes | Yes | **Overall appraisal (Include or Exclude)** | Include | Include | Include | Include | Include | Include | Include | Include | Include |
| 13) Was the trial design appropriate and any deviations from the standard RCT design (individual randomization, parallel groups) accounted for in the conduct and analysis of the trial? (Yes, No, Unclear, N/a) | Yes | Yes | **Overall score** **(Out of 11)** | 10 | 5 | 10 | 9 | 8 | 4 | 9 | 6 | 9 |
| **Overall appraisal (Include or Exclude)** | Include | Include |  |  |  |  |  |  |  |  |  |  |
| **Overall score** **(Out of 13)** | 11 | 6 |  |  |  |  |  |  |  |  |  |  |

\*Unclear because the two studies reporting on the same cohort used different approaches.
\*\*Incomplete follow-up strategies were unnecessary, so these studies should not be penalised for a N/a response to this JBI criterion.