Supplementary material 2. Risk of bias scale obtained from Kmet et al. (2004)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Criteria | Yes (2) | Partial (1) | No (0) | N/A |
| 1. | Question / objective sufficiently described? |  |  |  |  |
| 2. | Study design evident and appropriate? |  |  |  |  |
| 3. | Method of subject/comparison group selection *or* source of information/input variables described and appropriate? |  |  |  |  |
| 4. | Subject (and comparison group, if applicable) characteristics sufficiently described? |  |  |  |  |
| 5a. | Exposure measure(s) well defined and robust to measurement / misclassification bias? Means of assessment reported? |  |  |  |  |
| 5b. | Outcome measure(s) well defined and robust to measurement / misclassification bias? Means of assessment reported? |  |  |  |  |
| 6. | Sample size appropriate? |  |  |  |  |
| 7. | Analytic methods described/justified and appropriate? |  |  |  |  |
| 8. | Some estimate of variance is reported for the main results? |  |  |  |  |
| 9. | Sufficiently1 controlled for relevant other factors? |  |  |  |  |
| 10. | Results reported in sufficient detail? |  |  |  |  |
| 11. | Conclusions supported by the results? |  |  |  |  |
|  | Summary score2 |  |  |  |  |

1 ‘Sufficiently’ was defined as adjustment for age and gender, and at least one other relevant factor, including socio-economic (e.g., education or income), lifestyle (e.g., smoking, alcohol use or leisure-time physical activity) or health-related factors (e.g., body mass index or blood pressure).

2 Summary scores are calculated as: total sum[(number of ‘yes’ × 2) + (number of ‘partial’ × 1)]/total possible sum[24 − (number of ‘N/A’ × 2)]\*100%, with a maximum possible total score of 100%.