Appendix 1. NIH Quality Appraisal [*, a, b, c]

Author	C1.	C2.	C3.	C4.	C5.	C6.	C7.	C8.	C9.	C10.	C11.	C12.	C13.	C14.	Quality
Giannopoulou et al. [34]	Yes	Yes	NR	Yes	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	No	Fair
Burdzovic Andreas and Brunborg [35]	Yes	Yes	CD	Yes	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	Yes	Good
Belanger et al. [36]	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	No	Good
Black et al. [37]	Yes	Yes	No	Yes	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	Yes	Good
Westrupp et al. [38]	Yes	Yes	NR	No	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	Yes	Fair
Mayne et al. [39]	Yes	Yes	NR	Yes	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	Yes	Good
Gladstone et al. [40]	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	Yes	Good
Jolliff et al. [41]	Yes	Yes	CD	No	Yes	N/A	Yes	N/A	Yes	Yes	Yes	No	N/A	Yes	Fair

[*] NR, not reported; CD, cannot be determined; N/A, not applicable.

[a] NIH criteria of 6 included studies.

[b] Quality was rated as poor (0 - 4 out of 14 questions), fair (5 - 10 out of 14 questions), or good (11 - 14 out of 14 questions).

[c] The criteria were: C1. Was the research question or objective in this paper clearly stated? C2. Was the study population clearly specified and defined? C3. Was the participation rate of eligible persons at least 50%? C4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants? C5. Was a sample size justification, power description, or variance and effect estimates provided? C6. For the analyses

in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured? C7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed? C8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)? C9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants? C10. Was the exposure(s) assessed more than once over time? C11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants? C12. Were the outcome assessors blinded to the exposure status of participants? C13. Was loss to follow-up after baseline 20% or less? C14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)