**Supplementary Table 1.** Methodological quality assessment of the included studies

| Study | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Overall |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Aguilera (2010) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Low |
| Barrelet (1990) | Y | Y | Y | Y | N | Y | Y | Y | Y | N | Y | High |
| Bertrando (1992) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Brown (1962) | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Low |
| Brown (1972) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Ito (1995) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Ivanović (1994) | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Moderate |
| Jarbin (2000) | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | Y | Moderate |
| Karno (1987) | Y | Y | Y | Y | N | Y | Y | Y | Y | N | Y | High |
| King (1999) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Kopelowicz (2006) | Y | Y | Y | Y | N | Y | U | Y | Y | Y | Y | Moderate |
| Koutra (2015) | Y | Y | Y | Y | Y | Y | U | Y | Y | NA | Y | Low |
| Lee (2014) | Y | Y | Y | Y | Y | Y | Y | Y | Y | U | Y | Low |
| Leff (1987) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Leff (1990) |  Y |  Y | Y | Y | N | Y | Y | Y | Y | N | Y | High |
| Marom (2002) | Y | Y | Y | Y | Y | Y | U | Y | Y | NA | Y | Low |
| Marom (2005) |  Y |  Y | Y | Y | Y | Y | U | Y | Y | NA | Y | Low |
| Moline (1985) | Y | Y | Y | Y | N | Y | Y | Y | Y | N | Y | High |
| Montero (1992) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Mottaghipour (2001) | Y | Y | Y | Y | N | Y | Y | Y | Y | N | Y | High |
| Mozny (1992) | Y | Y | Y | Y | N | Y | Y | Y | Y | NA | Y | Moderate |
| Ng (2001) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Ng (2019) | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Low |
| Niedermeier (1992) | Y | Y | Y | Y | N | Y | Y | Y | Y | NA | Y | Moderate |
| Parker (1988) | Y | Y | Y | Y | N | Y | Y | Y | N | U | Y | Moderate |
| Roseliza-Murni (2014) | Y | Y | Y | Y | N | Y | U | U | Y | NA | Y | Low |
| Sadiq (2017) | Y | Y | Y | Y | Y | Y | Y | Y | Y |  U | Y | Moderate |
| Stirling (1991) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Stirling (1993) | Y | Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Tanaka (1995) | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | Y | Moderate |
| Mino (1997) |  Y |  Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Moderate |
| Uehara (1997) | Y | Y | Y | Y | N | Y | Y | Y | Y | NA | Y | Moderate |
| Vaughan (1992) | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Low |
| Vaughn (1976) | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y | Moderate |
| Leff (1981) |  Y |  Y | Y | Y | N | Y | Y | Y | Y | U | Y | Moderate |
| Vaughn (1984) | Y | Y | Y | Y | N | Y | Y | Y | Y | N | Y | High |
| Wang (2017) | Y | Y | Y | Y | N | Y | Y | Y | N | N | Y | High |
| Zanetti (2018) | Y | Y | Y | Y | N | Y | U | Y | Y | U | Y | Moderate |

\*JBI Critical Appraisal Checklist for Cohort Studies (Yes/No/Unclear/NA)

1. Were the two groups similar and recruited from the same population?
2. Were the exposures measured similarly to assign people to both exposed and unexposed groups?
3. Was the exposure measured in a valid and reliable way?
4. Were confounding factors identified?
5. Were strategies to deal with confounding factors stated?
6. Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)?
7. Were the outcomes measured in a valid and reliable way?
8. Was the follow up time reported and sufficient to be long enough for outcomes to occur?
9. Was follow up complete, and if not, were the reasons to loss to follow up described and explored?
10. Were strategies to address incomplete follow up utilized?
11. Was appropriate statistical analysis used?