**Table, Supplemental Digital Content 3. Risk of Bias Assessment.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | Selection | Comparability | Outcome | Total Stars (quality\*) |
| Allegretti et al. | \*\* | - | \* | 3 (poor) |
| Amiot et al. | \*\*\*\* | - | \*\*\* | 7 (poor) |
| Baert et al. | \*\*\* | - | \*\* | 5 (poor) |
| Bossuyt et al. | \*\*\*\* | - | \*\* | 6 (poor) |
| Buda et al. | \*\* | - | \*\* | 4 (poor) |
| Buisson et al. | \*\* | - | \*\*\* | 5 (poor) |
| Chan et al. | \*\*\* | - | \*\* | 5 (poor) |
| Chaparro et al. | \*\*\* | - | \*\* | 5 (poor) |
| Chu et al. | \*\* | - | \*\*\* | 5 (poor) |
| Dufour et al. | \*\*\* | - | \*\* | 5 (poor) |
| Lucidarme et al. | \*\*\* | - | \*\*\* | 6 (poor) |
| Mantzaris et al. | \*\*\*\* | \*\* | \*\*\* | 9 (good) |
| Olivares et al. | \*\*\* | - | \*\* | 5 (poor) |
| Paul et al. | \*\*\* | - | \*\* | 5 (poor) |
| Petitcollin et al. | \*\*\* | - | \*\*\* | 6 (poor) |
| Pouillon et al. | \*\*\* | - | \*\*\* | 6 (poor) |
| Van de Vondel et al. | \*\*\* | - | \*\* | 5 (poor) |
| Van Steenbergen et al. | \*\*\*\* | \*\* | \*\*\* | 9 (good) |
| Vázquez-Morón et al. | \*\* | - | \*\*\* | 5 (poor) |
| Viazis et al. | \*\*\* | - | \*\*\* | 6 (poor) |

\* Thresholds for converting the Newcastle-Ottawa scales to Agency for Healthcare Research and Quality standards (good, fair, and poor): good quality: 3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain, fair quality: 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain, poor quality: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in outcome/exposure domain