

## Supplementary material 5: Criterion-specific AMSTAR 2 credibility rating, over-all rating score, overall rating, for each included review.

Author, Year	PECO <sup>1</sup>	A-priori Methods <sup>2</sup>	Study Design Selection <sup>3</sup>	Search Strategy <sup>4</sup>	Study Selection <sup>5</sup>	Data Extraction <sup>6</sup>	Excluded Studies <sup>7</sup>	Included Studies <sup>8</sup>	RoB Assessment <sup>9</sup>	Funding Sources <sup>10</sup>	Statistical Methods <sup>11</sup>	Impact of RoB <sup>12</sup>	RoB Results <sup>13</sup>	Heterogeneity <sup>14</sup>	Publication Bias <sup>15</sup>	COI <sup>16</sup>	Rating score <sup>17</sup>	Overall Rating <sup>18</sup>
Samitz 2011	Yes	No	Yes	Yes	PY	PY	Yes	Yes	No	Yes	Yes	Yes	No	PY	Yes	Yes	<b>0.71</b>	<b>Low</b>
Coenen 2018	Yes	Yes	Yes	Yes	PY	PY	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	<b>0.78</b>	<b>Moderate</b>
Wendel Vos 2004	Yes	No	PY	No	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	<b>0.59</b>	<b>Low</b>
Jian Li 2013	Yes	No	Yes	PY	No	No	No	Yes	No	No	PY	No	No	No	PY	Yes	<b>0.34</b>	<b>Critically low</b>
Sattelmair 2011	Yes	No	Yes	PY	PY	PY	PY	Yes	No	Yes	Yes	No	No	no	Yes	Yes	<b>0.56</b>	<b>Low</b>
Wolin 2009	Yes	No	Yes	PY	No	No	No	No	No	No	PY	PY	No	PY	Yes	No	<b>0.31</b>	<b>Critically low</b>
Mahmood, 2017	Yes	Yes	Yes	PY	PY	PY	PY	Yes	No	No	Yes	No	No	PY	Yes	Yes	<b>0.59</b>	<b>Low</b>
Boyle 2012	Yes	No	Yes	PY	PY	PY	PY	Yes	Yes	Yes	Yes	Yes	Yes	PY	Yes	Yes	<b>0.78</b>	<b>Moderate</b>
Samad 2005	Yes	No	Yes	PY	No	No	No	Yes	No	No	Yes	No	No	No	No	No	<b>0.28</b>	<b>Critically low</b>
Robsahm 2013	Yes	No	Yes	PY	No	No	No	Yes	PY	No	Yes	Yes	Yes	Yes	No	Yes	<b>0.56</b>	<b>Low</b>
Wu Y, 2013	Yes	No	Yes	PY	PY	PY	No	Yes	No	No	Yes	No	No	Yes	Yes	PY	<b>0.5</b>	<b>Low</b>
Pizot 2016	Yes	No	Yes	PY	PY	No	No	Yes	No	Yes	No	no	No	Yes	Yes	Yes	<b>0.5</b>	<b>Low</b>
Chen X 2019	Yes	No	Yes	PY	PY	PY	No	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes	Yes	<b>0.75</b>	<b>Moderate</b>
Voskuil 2007	Yes	No	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes	No	No	PY	Yes	No	<b>0.53</b>	<b>Low</b>
Schmid 2015	Yes	No	Yes	PY	No	PY	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	<b>0.69</b>	<b>Low</b>
Vermaete 2013	Yes	No	Yes	PY	No	Yes	PY	Yes	Yes	Yes	Yes	No	No	PY	No	Yes	<b>0.56</b>	<b>Low</b>
Singh 2014	Yes	Yes	Yes	PY	Yes	PY	No	Yes	Yes	Yes	Yes	Yes	PY	PY	Yes	Yes	<b>0.75</b>	<b>Moderate</b>

Psaltopoulou 2015	Yes	No	Yes	PY	PY	PY	PY	Yes	Yes	No	Yes	No	No	PY	Yes	Yes	0.81	Moderate	
Chen Y 2014	Yes	No	Yes	PY	PY	PY	PY	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	0.69	Low	
Behrens 2014	Yes	No	Yes	Yes	No	No	No	Yes	Yes	No	Yes	Yes	PY	PY	No	No	0.5	Low	
Behrens, 2013	Yes	no	Yes	Yes	No	No	PY	Yes	Yes	No	Yes	Yes	No	Yes	Yes	no	0.59	Low	
Shephard 2016	Yes	No	Yes	PY	No	No	No	Yes	No	Yes	N/A	N/A	N/A	N/A	N/A	Yes	0.5	Low	
Krstev 2019	Yes	No	Yes	No	No	no	PY	Yes	no	Yes	NO	No	No	PY	no	Yes	0.38	Critically low	
Benke, 2018	Yes	No	Yes	PY	No	PY	PY	Yes	Yes	Yes	Yes	Yes	No	PY	Yes	Yes	0.69	Low	
Shephard, 2017	Yes	No	Yes	PY	No	No	No	Yes	No	Yes	N/A	N/A	N/A	N/A	No	Yes	0.5	Low	
Liu 2011	Yes	No	Yes	Yes	No	Yes	PY	Yes	Yes	Yes	Yes	Yes	Yes	PY	Yes	Yes	0.81	Moderate	
O Rorke, 2010	Yes	No	Yes	Yes	PY	PY	No	Yes	Yes	Yes	Yes	Yes	No	No	PY	Yes	No	0.59	Low
Bao 2008	Yes	No	Yes	PY	No	No	PY	Yes	No	Yes	Yes	Yes	No	PY	No	Yes	Yes	0.53	Low
Keimling 2014	Yes	No	Yes	PY	PY	No	PY	Yes	No	No	Yes	No	No	Yes	Yes	No	0.47	Critically low	
Aune 2015	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	PY	PY	Yes	Yes	Yes	0.75	Moderate
McWilliams 2011	Yes	No	Yes	Yes	No	PY	PY	Yes	Yes	Yes	No	No	No	No	No	Yes	0.5	Low	
Gignac 2019	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	N/A	N/A	N/A	N/A	No	no	0.58	Low	
Palmer 2012	Yes	No	No	PY	No	No	No	Yes	Yes	Yes	N/A	N/A	N/A	N/A	No	no	0.38	Critically low	
White 2017	Yes	Yes	Yes	Yes	Yes	Yes	PY	Yes	Yes	Yes	Yes	Yes	PY	Yes	PY	PY	Yes	0.88	Moderate
Yang B, 2018	Yes	No	Yes	PY	PY	No	No	Yes	No	PY	Yes	No	No	Yes	Yes	No	0.5	Low	
Huai 2013	Yes	No	Yes	PY	No	PY	PY	Yes	Yes	Yes	Yes	No	No	Yes	Yes	no	0.59	Low	
Total amount of Yes	36	4	34	10	3	5	3	35	21	23	26	11	7	11	24	23			

Abbreviations: COI = conflict of interest; PECO = population, exposure, comparator, outcome; PY = partial yes; RoB = risk of bias

<sup>1</sup> Did the research questions and inclusion criteria for the review include the components of PICO/PECO?

<sup>2</sup> Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

<sup>3</sup> Did the review authors explain their selection of the study designs for inclusion in the review?

<sup>4</sup> Did the review authors use a comprehensive literature search strategy?

<sup>5</sup> Did the review authors perform study selection in duplicate?

<sup>6</sup> Did the review authors perform data extraction in duplicate?

<sup>7</sup> Did the review authors provide a list of excluded studies and justify the exclusions?

<sup>8</sup> Did the review authors describe the included studies in adequate detail?

- <sup>9</sup> Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?
- <sup>10</sup> Did the review authors report on the sources of funding for the studies included in the review?
- <sup>11</sup> If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?
- <sup>12</sup> If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?
- <sup>13</sup> Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?
- <sup>14</sup> Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?
- <sup>15</sup> If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?
- <sup>16</sup> Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?
- <sup>17</sup> Shea et al. 2017. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. (7)