

Table S1. Quality assessment of included studies.

Author, year	A Selection Bias	B Study Design	C Confounders	D Blinding	E Data Collection Method	F Withdrawals and Dropouts	Global Rating
Wang et al., 2022 [20]	M	M	S	N/A	S	N/A	Strong
Wang et al., 2022 [21]	M	M	S	N/A	S	N/A	Strong
Gupta et al., 2020 [22]	M	M	W	N/A	S	N/A	Moderate
Arzt et al., 2017 [23]	M	S	M	N/A	S	N/A	Strong
Oldenburg et al., 2007 [16]	M	M	M	N/A	S	N/A	Moderate
Bitter et al., 2009 [24]	M	M	M	N/A	M	N/A	Moderate
Chan et al., 1997 [25]	M	M	M	N/A	M	N/A	Moderate
Yumino et al., 2009 [26]	S	M	S	N/A	S	N/A	Strong
Herrscher et al., 2011 [27]	M	M	S	N/A	M	N/A	Moderate
Kalaydzhev et al., 2023 [28]	S	M	M	N/A	S	N/A	Strong
Arikawa et al., 2016 [29]	M	M	S	M	M	W	Moderate
Naito et al., 2022 [30]	M	M	M	W	M	M	Moderate
Kaneko et al., 2003 [31]	M	S	S	S	S	M	Strong
Mansfield et al., 2004 [32]	S	S	S	S	S	S	Strong
Fox et al., 2021 [33]	S	S	S	S	M	S	Strong
Kim et al., 2019 [34]	S	S	M	S	M	S	Strong
Gilman et al., 2008 [35]	S	S	S	S	S	S	Strong
Servantes et al., 2018 [36]	S	S	M	S	S	S	Strong
Egea et al., 2007 [37]	S	M	S	M	S	W	Moderate
Ryan et al., 2005 [38]	M	M	S	S	S	M	Strong
Cistulli et al., 2023 [39]	M	M	M	N/A	S	M	Moderate
Abdullah et al., 2018 [40]	M	M	M	N/A	M	N/A	Moderate
Malhotra et al., 2023 [41]	M	M	W	N/A	M	N/A	Moderate
Wojeck et al., 2023 [42]	M	M	S	S	M	W	Moderate
Neeland et al., 2020 [43]	M	M	S	S	S	W	Moderate
Furukawa et al., 2018 [45]	M	M	M	W	M	M	Moderate

Tang et al., 2019 [45]	S	S	M	S	M	W	Moderate
Sawada et al., 2018 [46]	M	M	M	W	M	W	Moderate
Owens et al., 2021 [48]	S	S	S	S	S	S	Strong
Pelaia et al., 2022 [49]	M	M	M	W	M	M	Moderate
Jaffuel et al., 2021 [50]	M	S	M	W	M	S	Moderate
Wang et al., 2023 [51]	M	M	M	W	M	M	Moderate
Passino et al., 2021 [52]	M	M	M	W	M	S	Moderate

Abbreviations; S, strong; M, moderate; W, weak; N/A, not applicable

Table S2. Effective Public Healthcare Panacea Project - Quality Assessment Tool for Quantitative Studies Dictionary.

Components of EPHPP	Questions	Instructions for completion
A. Selection Bias	<p>(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?</p> <p>1 Very likely 2 Somewhat likely 3 Not likely 4 Can't tell</p> <p>(Q2) What percentage of selected individuals agreed to participate?</p> <p>1 80 - 100% agreement 2 60 – 79% agreement 3 less than 60% agreement 4 Not applicable 5 Can't tell</p>	<p>Strong: The selected individuals are very likely to be representative of the target population (Q1 is 1) and there is greater than 80% participation (Q2 is 1).</p> <p>Moderate: The selected individuals are at least somewhat likely to be representative of the target population (Q1 is 1 or 2); and there is 60 - 79% participation (Q2 is 2). 'Moderate' may also be assigned if Q1 is 1 or 2 and Q2 is 5 (can't tell).</p> <p>Weak: The selected individuals are not likely to be representative of the target population (Q1 is 3); or there is less than 60% participation (Q2 is 3) or selection is not described (Q1 is 4); and the level of participation is not described (Q2 is 5)</p> <p>Strong - 1 Moderate - 2 Weak - 3</p>
B. Study Design	<p>(Q1) Indicate the study design</p> <p>1 Randomized controlled trial 2 Controlled clinical trial 3 Cohort analytic (two group pre + post) 4 Case-control</p>	<p>Strong: will be assigned to those articles that described RCTs and CCTs.</p> <p>Moderate: will be assigned to those that described a cohort analytic study, a case control study, a cohort design, or an</p>

	<p>5 Cohort (one group pre + post (before and after)) 6 Interrupted time series 7 Other specify 8 Can't tell</p> <p>Was the study described as randomized? If NO, go to Component C. No/Yes If Yes, was the method of randomization described? No/Yes If Yes, was the method appropriate? No/Yes</p>	<p>interrupted time series. Weak: will be assigned to those that used any other method or did not state the method used.</p> <p>Strong - 1 Moderate - 2 Weak - 3</p>
C. Confounders	<p>(Q1) Were there important differences between groups prior to the intervention? 1 Yes 2 No 3 Can't tell The following are examples of confounders: 1 Race 2 Sex 3 Marital status/family 4 Age 5 SES (income or class) 6 Education 7 Health status 8 Pre-intervention score on outcome measure</p>	<p>Strong: will be assigned to those articles that controlled for at least 80% of relevant confounders (Q1 is 2); or (Q2 is 1). Moderate: will be given to those studies that controlled for 60 – 79% of relevant confounders (Q1 is 1) and (Q2 is 2). Weak: will be assigned when less than 60% of relevant confounders were controlled (Q1 is 1) and (Q2 is 3) or control of confounders was not described (Q1 is 3) and (Q2 is 4).</p>

	<p>(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?</p> <p>1 80 – 100% (most) 2 60 – 79% (some) 3 Less than 60% (few or none) 4 Can't Tell</p>	<p>Strong - 1 Moderate - 2 Weak - 3</p>
D. Blinding	<p>(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?</p> <p>1 Yes 2 No 3 Can't tell</p> <p>(Q2) Were the study participants aware of the research question?</p> <p>1 Yes 2 No 3 Can't tell</p>	<p>Strong: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); and the study participants are not aware of the research question (Q2 is 2).</p> <p>Moderate: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); or the study participants are not aware of the research question (Q2 is 2); or blinding is not described (Q1 is 3 and Q2 is 3).</p> <p>Weak: The outcome assessor is aware of the intervention status of participants (Q1 is 1); and the study participants are aware of the research question (Q2 is 1).</p>

		Strong - 1 Moderate - 2 Weak - 3
E. Data Collection Method	<p>(Q1) Were data collection tools shown to be valid? 1 Yes 2 No 3 Can't tell</p> <p>(Q2) Were data collection tools shown to be reliable? 1 Yes 2 No 3 Can't tell</p>	<p>Strong: The data collection tools have been shown to be valid (Q1 is 1); and the data collection tools have been shown to be reliable (Q2 is 1). Moderate: The data collection tools have been shown to be valid (Q1 is 1); and the data collection tools have not been shown to be reliable (Q2 is 2) or reliability is not described (Q2 is 3). Weak: The data collection tools have not been shown to be valid (Q1 is 2) or both reliability and validity are not described (Q1 is 3 and Q2 is 3).</p> <p>Strong - 1 Moderate - 2 Weak - 3</p>
F. Withdrawals and Dropouts	<p>(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group? 1 Yes 2 No 3 Can't tell 4 Not Applicable (i.e. one time surveys or interviews)</p>	<p>Strong: will be assigned when the follow-up rate is 80% or greater (Q2 is 1). Moderate: will be assigned when the follow-up rate is 60 – 79% (Q2 is 2) OR Q2 is 5 (N/A). Weak: will be assigned when a follow-up rate is less than 60% (Q2 is 3) or if the withdrawals and drop-outs were not described (Q2 is 4).</p>

	<p>(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).</p> <p>1 80 -100%</p> <p>2 60 - 79%</p> <p>3 less than 60%</p> <p>4 Can't tell</p> <p>5 Not Applicable (i.e. Retrospective case-control)</p>	<p>Strong - 1</p> <p>Moderate - 2</p> <p>Weak - 3</p>
Global Rating	<p>GLOBAL RATING FOR THIS PAPER (circle one):</p> <p>1 STRONG (no WEAK ratings)</p> <p>2 MODERATE (one WEAK rating)</p> <p>3 WEAK (two or more WEAK ratings)</p> <p>With both reviewers discussing the ratings: Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?</p> <p>No/Yes</p> <p>Final decision of both reviewers (circle one):</p> <p>1 STRONG</p> <p>2 MODERATE</p> <p>3 WEAK</p>	

Abbreviations: EPHPP, Effective Public Healthcare Panacea Project; Q1, question 1; Q2, question 2.