

Supplementary Material 2

	Bias due to confounding	Bias in selection of participants	Bias in classifying interventions	Bias due to deviation from intended intervention	Bias due to missing data	Bias in outcome measurement	Bias in selection of reported results	Overall risk of bias
Arnold, 2021	2	1	1	1	2	2	1	1
Chambers, 2022	2	1	1	1	3	2	2	2
Doernberg, 2015	2	1	1	1	2	2	1	1
Durham, 2020	2	1	1	1	3	2	2	2
Hanlon, 2021	2	1	1	1	1	2	1	1
Harbin, 2023	3	1	1	2	3	2	1	2
Hartman, 2023	2	1	1	1	2	2	1	1
Kassett, 2016	2	1	1	2	1	2	1	1
Konig, 2024	3	1	1	1	3	2	1	2
Loeb, 2005	2	1	1	1	3	2	1	2
Liao, 2020	2	1	1	1	2	2	1	1
McMaughan, 2016	2	1	1	2	3	2	2	2
Mitchell, 2021	2	1	1	4	2	2	1	2
Nace, 2020	2	1	1	2	3	2	1	2
Pasay, 2019	2	1	1	2	3	2	1	2
Rutten, 2022	2	1	1	1	2	2	1	1
Salem-Schatz, 2020	2	1	1	2	3	2	2	2
Sloane, 2020	2	1	1	2	3	2	1	2
Zabarsky, 2008	3	1	1	2	3	2	1	2
Zimmerman, 2014	2	1	1	2	3	2	1	2

Supplementary material 2. Summary of study risk of bias with regard to each domain and overall risk of bias.

	Is there no potential for confounders affecting the intervention?	Was analysis based on splitting participant follow-up time?	Did authors use appropriate analysis to control confounding domains?	Were confounding domains measured validly and reliably?	Were study outcomes not affected by confounders?	Did authors use an appropriate analysis method for time varying confounding?	Were time varying confounders measured validly and reliably?	Risk of bias due to confounding
Arnold, 2021	4	4	1	2	1	2	2	2
Chambers, 2022	4	4	2	1	1	1	1	2
Doernberg, 2015	4	4	1	3	1	2	2	2
Durham, 2020	4	4	1	2	2	2	2	2
Hanlon, 2021	4	4	2	1	1	1	1	2
Harbin, 2023	4	4	1	4	1	2	2	3
Hartman, 2023	4	4	2	1	1	1	1	2
Kassett, 2016	4	4	2	2	1	2	2	2
Konig, 2024	4	4	2	3	1	2	2	3
Loeb, 2005	4	4	1	3	1	2	2	2
Liao, 2020	4	4	1	2	1	2	2	2
McMaughan, 2016	4	4	1	3	1	2	2	2
Mitchell, 2021	4	4	1	1	1	1	1	2
Nace, 2020	4	4	1	1	1	1	1	2
Pasay, 2019	4	4	1	1	1	2	2	2
Rutten, 2022	4	4	1	3	1	2	2	2
Salem-Schatz, 2020	4	4	1	2	1	1	1	2
Sloane, 2020	4	4	1	2	1	1	1	2
Zabarsky, 2008	4	4	N/A	2	N/A	N/A	N/A	3
Zimmerman, 2014	4	4	1	2	1	2	2	2

Risk of bias assessment to each signalling question within the risk of bias due to confounding domain.

	Was selection of participants based on characteristics observed BEFORE the intervention?	Start of intervention and start of follow-up did not coincide.	If no to Q1: Post-intervention factors leading to selection were not due to intervention?	If no to Q1: Post-intervention factors leading to selection were not due to the outcome?	If no to Q1 or Q2: Were adjustment techniques used?	Bias in selection of participants
Arnold, 2021	1	1	N/A	N/A	N/A	1
Chambers, 2022	1	1	N/A	N/A	N/A	1
Doernberg, 2015	1	1	N/A	N/A	N/A	1
Durham, 2020	1	1	N/A	N/A	N/A	1
Hanlon, 2021	1	1	N/A	N/A	N/A	1
Harbin, 2023	1	1	N/A	N/A	N/A	1
Hartman, 2023	1	1	N/A	N/A	N/A	1
Kassett, 2016	1	1	N/A	N/A	N/A	1
Konig, 2024	1	1	N/A	N/A	N/A	1
Loeb, 2005	1	1	N/A	N/A	N/A	1
Liao, 2020	1	1	N/A	N/A	N/A	1
McMaughan, 2016	1	1	N/A	N/A	N/A	1
Mitchell, 2021	1	1	N/A	N/A	N/A	1
Nace, 2020	1	1	N/A	N/A	N/A	1
Pasay, 2019	1	1	N/A	N/A	N/A	1
Rutten, 2022	1	1	N/A	N/A	N/A	1
Salem-Schatz, 2020	1	1	N/A	N/A	N/A	1
Sloane, 2020	1	1	N/A	N/A	N/A	1
Zabarsky, 2008	1	1	N/A	N/A	N/A	1
Zimmerman, 2014	1	1	N/A	N/A	N/A	1

Risk of bias assessment to each signalling question within the risk of bias in the selection of participants domain.

	Were intervention groups clearly defined?	Was the information used to define intervention groups recorded at the start?	Classification of intervention status was not affected by knowledge of the outcome?	Bias in classifying interventions
Arnold, 2021	1	1	1	1
Chambers, 2022	1	1	1	1
Doernberg, 2015	1	1	1	1
Durham, 2020	1	2	1	1
Hanlon, 2021	1	1	1	1
Harbin, 2023	1	1	1	1
Hartman, 2023	1	1	1	1
Kassett, 2016	1	1	1	1
Konig, 2024	1	1	1	1
Loeb, 2005	1	1	1	1
Liao, 2020	1	1	1	1
McMaughan, 2016	1	1	1	1
Mitchell, 2021	1	1	1	1
Nace, 2020	1	1	1	1
Pasay, 2019	1	1	1	1
Rutten, 2022	1	1	1	1
Salem-Schatz, 2020	1	1	1	1
Sloane, 2020	1	1	1	1
Zabarsky, 2008	1	1	1	1
Zimmerman, 2014	1	1	1	1

Risk of bias assessment to each signalling question within the risk of bias in the classifying interventions domain.

	There were no deviations in intervention beyond what could be expected?	If there were deviations, they were balanced between groups and did not affect outcome?	Risk of bias due to deviation from intended intervention
Arnold, 2021	1	N/A	1
Chambers, 2022	1	N/A	1
Doernberg, 2015	1	N/A	1
Durham, 2020	1	N/A	1
Hanlon, 2021	1	N/A	1
Harbin, 2023	2	2	2
Hartman, 2023	1	N/A	1
Kassett, 2016	2	2	2
Konig, 2024	1	N/A	1
Loeb, 2005	1	N/A	1
Liao, 2020	1	N/A	1
McMaughan, 2016	2	2	2
Mitchell, 2021	3	4	4
Nace, 2020	2	2	2
Pasay, 2019	2	2	2
Rutten, 2022	1	N/A	1
Salem-Schatz, 2020	2	2	2
Sloane, 2020	2	2	2
Zabarsky, 2008	2	2	2
Zimmerman, 2014	2	2	2

Risk of bias assessment to each signalling question within the risk of bias due to deviation from intended intervention domain.

	Did all participants have outcome data (i.e. no exclusion or drop-outs). If yes, skip Q2 and Q3.	Are proportions and reasons for missing data similar across groups?	Is there evidence that the results were robust to the presence of missing data?	Risk of bias due to missing data
Arnold, 2021	4	2	1	2
Chambers, 2022	4	4	2	3
Doernberg, 2015	4	1	1	2
Durham, 2020	4	4	1	3
Hanlon, 2021	1	N/A	N/A	1
Harbin, 2023	4	4	1	3
Hartman, 2023	4	1	2	2
Kassett, 2016	1	N/A	N/A	1
Konig, 2024	4	4	2	3
Loeb, 2005	4	4	2	3
Liao, 2020	4	1	1	2
McMaughan, 2016	4	4	1	3
Mitchell, 2021	4	1	1	2
Nace, 2020	4	4	2	3
Pasay, 2019	4	3	3	3
Rutten, 2022	4	1	1	2
Salem-Schatz, 2020	4	4	2	3
Sloane, 2020	4	4	2	3
Zabarsky, 2008	4	4	2	3
Zimmerman, 2014	4	4	2	3

Risk of bias assessment to each signalling question within the risk of bias due to missing data domain.

	Outcome measure was not influenced by knowledge of receiving the intervention	Assessors were blinded to intervention received by participants	Method of outcome assessment was comparable across groups	There were no systematic errors in measurement of outcome	Risk of bias in outcome measurement
Arnold, 2021	3	1	1	1	2
Chambers, 2022	3	3	1	1	2
Doernberg, 2015	3	3	1	1	2
Durham, 2020	3	3	1	1	2
Hanlon, 2021	3	1	1	1	2
Harbin, 2023	3	3	1	2	2
Hartman, 2023	3	1	1	1	2
Kassett, 2016	3	3	1	2	2
Konig, 2024	3	3	1	1	2
Loeb, 2005	3	3	1	1	2
Liao, 2020	3	1	1	1	2
McMaughan, 2016	3	3	1	2	2
Mitchell, 2021	3	1	1	4	2
Nace, 2020	3	1	1	2	2
Pasay, 2019	3	1	1	2	2
Rutten, 2022	3	1	1	1	2
Salem-Schatz, 2020	3	3	1	2	2
Sloane, 2020	3	3	1	2	2
Zabarsky, 2008	3	3	1	2	2
Zimmerman, 2014	3	3	1	2	2

Risk of bias assessment to each signalling question within the risk of bias in outcome measurement.

	Reported effect was not selected based on multiple analyses of the outcome	Reported effect was not selected based on multiple analyses of the intervention- outcome relationship	Reported effect was not selected based on differing subgroups	Bias in selection of reported results
Arnold, 2021	1	1	1	1
Chambers, 2022	3	1	1	2
Doernberg, 2015	1	1	1	1
Durham, 2020	3	1	1	2
Hanlon, 2021	1	1	1	1
Harbin, 2023	1	1	1	1
Hartman, 2023	1	1	1	1
Kassett, 2016	1	1	1	1
Konig, 2024	1	1	1	1
Loeb, 2005	1	1	1	1
Liao, 2020	1	1	1	1
McMaughan, 2016	2	1	3	2
Mitchell, 2021	1	1	1	1
Nace, 2020	1	1	1	1
Pasay, 2019	1	1	1	1
Rutten, 2022	1	1	1	1
Salem-Schatz, 2020	2	3	2	2
Sloane, 2020	1	1	1	1
Zabarsky, 2008	1	1	1	1
Zimmerman, 2014	1	1	1	1

Risk of bias assessment to each signalling question within the risk of bias in the selection of reported results.