Outcome report data extraction and risk of bias tools

Data extraction - outcome evaluations

Consider if study has reported following. Provide justifications for your response wherever possible.

Intervention Details

Intervention name

Intervention location
Details (country; area(s) of country)

Not stated

Description of intervention
Details (overall aim(s) of intervention; key components of intervention; content and activities; who delivered intervention; deviations from original protocol; etc.)

Not stated

Timing and duration of intervention
Details

Not stated

Target population
Details (specific characteristics; how y were recruited)

Not stated

Technology
Details

Not stated

Provider organisation description
Details

Not stated

How intervention was developed
Details

Not stated

Intervention theory of change
Details

Not stated

Intervention offered to control group
Not applicable
Details

Not stated

Evaluation Details

Research question(s) or hypotheses
Details

Not stated

Overall design within which evaluation is part
RCT
Quasi-experimental comparison
Other (please state):
Not stated

Timing of outcome evaluation
Details

Not stated

Sampling strategy
Details

Not stated

Sample size, overall response rates at baseline

Details (intervention group)

Details (control group)

Not stated

Sample size, overall response rates at follow-up

Details (intervention group)

Details (control group)

Not stated
Not appropriate

Sociodemographic characteristics at baseline/follow-up
Details

Not stated

Were baseline equivalence/differences between arms reported?
Details

Not stated

How were baseline differences between intervention and comparison groups controlled?
Details

Not stated

For *each* outcome measure, report following:

Description of outcome measure provided
Details

Not stated

Outcomes stated (in protocol, from outset)

Primary outcomes

Secondary outcomes

Not stated

Post-hoc outcomes reported
Description

Not stated

Evidence of reliability
Description

Not stated

Data collection methods
Description

Not stated

Baseline response rate

Details (intervention group)

Details (control group)

Not stated

Follow-up response rate

Details (intervention group)

Details (control group)

Not stated

Effect sizes

Overall

By sexuality

By gender identity

By socioeconomic status

By ethnic sub-group

No effect size

Study analysis was intention to treat?

Yes

No

Not stated

Study analysis appropriately accounted for any clustering?

Yes

No

Not stated

Study analysis adjusted for confounders?

Yes

No

Not stated

If study is randomised controlled trial:

Unit of allocation
Individual
Other (please indicate)

Not stated

Generation of allocation sequence (e.g. What was random component of sequence generation process—this may include minimisation?)
Details

Not stated

Concealment of allocation
Details

Not stated

Blinding
Details

Not stated

If study involved non-randomised control group(s):
Unit of comparison
Individual
Other (please indicate)

Not stated

How controls were identified (e.g. matching, restriction)
Details

Not stated

How confounding was minimised (e.g. matching, adjustment etc.)
Details

Not stated

If includes economic evaluations

Perspective taken for direct and indirect costs
Details

Not stated

Evaluation framework
Details

Not stated

Source of effectiveness estimates
Details

Not stated

Critical assumptions
Details

Not stated

Discount rates
Details

Not stated

Approach to cost-effectiveness in form of incremental cost-effectiveness ratios or net (health) benefits
Details

Not stated

Quality assessment tool – randomised controlled trials

Consider if study has reported following. Provide justifications for your response wherever possible.

Adequate generation (random) of allocation sequence
Yes
No
Not clear
Not stated

Concealed allocation
Yes
No
Not clear
Not stated

Blinding of participants/personnel
Yes
No
Not clear
Not stated

Blinding of outcome assessors

Yes
No
Not clear
Not stated

Complete outcome data (i.e. low attrition) *(were complete data for each outcome reported, and if not, were reasons given for incomplete reporting?)*
Yes
No
Not clear
Not stated

Reporting complete, not selective by measure
Yes
No
Not clear
Not stated

Controlled for confounding
Yes
No
Not clear
Not stated

Accounted for clustering
Yes
No
Not clear
Not stated

Aimed to reduce other forms of bias that may have entered study
Yes
No
Not clear
Not stated

Quality assessment tool – quasi-experimental

ROBINS-I tool (Stage I): At protocol stage

<u>Specify question:-</u>

Participants=
Intervention=
Control=
Outcomes=
List confounding domains relevant to all/most studies=
List co-interventions which could be different between groups which could impact outcomes=

ROBINS-I tool (Stage II): For each study

Specify target RCT specific to study

Design= Individuall randomised / Cluster randomised / Matched
Participants=
Intervention =
Control=
Is study aim …?
-to assess effect of assignment to intervention □
-to assess effect of starting & adhering to intervention □

Specify the outcome

Specify which outcome being assessed for risk of bias (generally from among those earmarked for Summary of Findings table). Specify whether this is proposed intervention benefit or harm.
=

Specify numerical result being assessed

In case of multiple alternative analyses, specify numeric result (e.g. RR = 1.51 (95% CI 0.82 to 2.78) and/or reference (e.g. to table/figure/paragraph) which defines result being assessed.
=

Preliminary consideration of confounders

Complete row for each key confounding domain i) listed in review protocol; & ii) relevant to study setting or which study authors identify as possibly important.

"Important" domains are those for which, in this study context, adjustment is expected to produce important change in estimated effect of intervention. "Validity" refers to whether variable(s) fully measure confounding domain, while "reliability" refers to precision of measurement (more error means less reliability).

(i) Confounding domains listed in review protocol				
Confounding domain	Measure variable(s)	Evidence that controlling for this variable not necessary?*	Is confounding domain validly & reliably measured by this variable(s)?	OPTIONAL: Is failure to adjust for this variable alone expected to favour intervention or control?
			Yes / No / No info	Favour intervention / Favour control / No info

(ii) Additional confounding domains relevant to study setting or which study authors identified as important				
Confounding domain	Measured variable(s)	Evidence that controlling for variable not necessary?*	Is confounding domain validly & reliably measured by variable(s)?	OPTIONAL: Is failure to adjust for variable alone expected to favour intervention or control?
			Yes / No / No info	Favour intervention / Favour control / No info

* In context of particular study, variables may be demonstrated not to be confounders so not included in analysis: a) if they are not predictive of outcome; b) if they are not predictive of intervention; or c) because adjustment makes no/minimal difference to estimated effect of primary parameter. N.B. "no statistically significant association" is not same as "not predictive".

Preliminary consideration of co-interventions

Complete row for each important co-intervention i) listed in review protocol; & ii) relevant to study setting or which study authors identify as important.

"Important" co-interventions are those for which in study context adjustment is expected to lead to important change in estimated effect of intervention.

ii) Additional co-interventions relevant to study setting, or which study authors identified as important

i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was not necessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in intervention or control
		Favour intervention / Favour control / No info
		Favour intervention / Favour control / No info
		Favour intervention / Favour control / No info

(ii) Additional co-interventions relevant to study setting, or which study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was not necessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in intervention or control
		Favour intervention / Favour control / No info
		Favour intervention / Favour control / No info
		Favour intervention / Favour control / No info

Risk of bias assessment

Responses underlined in green are potential indicators for low risk of bias, & responses in red are potential indicators for risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Bias domain	Signalling questions	Elaboration	Response options
Bias due to confounding			
	1.1 Is there potential for confounding of effect of intervention in this study? If N/PN to 1.1: study can be considered to be at low risk of bias due to confounding & no further signalling questions need be considered	In rare situations, such as when studying harms which are very unlikely to be related to factors that influence treatment decisions, no confounding is expected & study can be considered to be at low risk of bias due to confounding, equivalent to fully randomised trial. There is no NI (No info) option for this signalling question.	Y / PY / PN / N
	If Y/PY to 1.1: determine whether there is need to assess time-varying confounding:		
	1.2. Was analysis based on splitting participants' follow up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3.	If participants could switch between intervention groups then associations between intervention & outcome may be biased by time-varying confounding. This occurs when prognostic factors influence switches between intended interventions.	NA / Y / PY / PN / N / NI
	1.3. Were intervention discontinuations or switches likely to be related to factors which are prognostic for outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline & time-varying confounding (1.7 & 1.8)	If intervention switches are unrelated to outcome, for example when outcome is unexpected harm, then time-varying confounding will not be present & only control for baseline confounding is required.	NA / Y / PY / PN / N / NI
	Questions relating to baseline confounding only		
	1.4. Did authors use appropriate	Appropriate methods to control for measured confounders include	NA / Y / PY / PN / N / NI

	analysis method which controlled for all important confounding domains?	stratification, regression, matching, standardization, & inverse probability weighting. They may control for individual variables or for estimated propensity score. Inverse probability weighting is based on function of propensity score. Each method depends on assumption that there is no unmeasured or residual confounding.	
	1.5. If Y/PY to 1.4: Were confounding domains which were controlled for measured validly & reliably by variables available in this study?	Appropriate control of confounding requires that variables adjusted for are valid & reliable measures of confounding domains. For some topics, list of valid & reliable measures of confounding domains will be specified in review protocol but for others such a list may not be available. Study authors may cite references to support use of a particular measure. If authors control for confounding variables with no indication of validity or reliability pay attention to subjectivity of measure. Subjective measures (e.g. based on self-report) may have lower validity & reliability than objective measures such as lab findings.	NA / Y / PY / PN / N / NI
	1.6. Did authors control for any post intervention variables which could have been affected by intervention?	Controlling for post-intervention variables which are affected by intervention is not appropriate. Controlling for mediating variables estimates direct effect of intervention & may introduce bias. Controlling for common effects of intervention & outcome introduces bias.	NA / Y / PY / PN / N / NI
	Questions relating to baseline & time-varying confounding		
	1.7. Did authors use appropriate analysis method which controlled for all important confounding domains & for time-varying confounding?	Adjustment for time-varying confounding is necessary to estimate effect of starting & adhering to intervention, in both randomised trials & NRSI. Appropriate methods include those based on inverse probability weighting. Standard regression models which include time-updated confounders may be problematic if time-varying confounding is present.	NA / Y / PY / PN / N / NI

	1.8. If Y/PY to 1.7: Were confounding domains which were controlled for measured validly & reliably by variables available in this study?	See 1.5 above.	NA / Y / PY / PN / N / NI
	Risk of bias judgement	See Table B	Low / Moderate / Serious / Critical / NI
	Optional: What is predicted direction of bias due to confounding?	Can true effect estimate be predicted to be greater or less than estimated effect in study because one or more of important confounding domains was not controlled for? Answering this question will be based on expert knowledge & results in other studies & therefore can only be completed after all of studies in body of evidence have been reviewed. Consider potential effect of each of unmeasured domains & whether all important confounding domains not controlled for in analysis would be likely to change estimate in same direction, or if one important confounding domain which was not controlled for in analysis is likely to have dominant impact.	Favours intervention / Favours control / Unpredictable

Bias in selection of participants into study			
	2.1. Was selection of participants into study (or into analysis) based on participant characteristics observed after start of intervention? If N/PN to 2.1: go to 2.4	This domain is concerned only with selection into study based on participant characteristics observed after start of intervention. Selection based on characteristics observed before start of intervention can be addressed by controlling for imbalances between intervention & control groups in baseline characteristics which are prognostic for outcome (baseline confounding).	Y / PY / PN / N / NI
	2.2. If Y/PY to 2.1: Were post intervention variables which influenced selection likely to be associated with intervention?	Selection bias occurs when selection is related to effect of either intervention or cause of intervention & effect of either outcome or cause of outcome. Therefore, result is at risk of selection bias if selection into study is related to both intervention & outcome.	NA / Y / PY / PN / N / NI
	2.3 If Y/PY to 2.2: Were post intervention variables which influenced selection likely to be influenced by outcome or cause of outcome?		NA / Y / PY / PN / N / NI
	2.4. Do start of follow-up & start of intervention coincide for most participants?	If participants are not followed from start of intervention then a period of follow up has been excluded, & individuals who experienced outcome soon after intervention will be missing from analyses. This problem may occur when prevalent, rather than new (incident), users of intervention are included in analyses.	Y / PY / PN / N / NI
	2.5. If Y/PY to 2.2 & 2.3, or N/PN to 2.4: Were adjustment techniques used which are likely to correct for presence of selection biases?	It is in principle possible to correct for selection biases, for example by using inverse probability weights to create pseudo-population in which selection bias has been removed, or by modelling distributions of missing participants or follow up times & outcome events & including them using missing data methodology. However such methods are rarely	NA / Y / PY / PN / N / NI

		used & answer to this question will usually be “No”.	
	Risk of bias judgement	See Table B	Low / Moderate / Serious / Critical / NI
	Optional: What is predicted direction of bias due to selection of participants into study?	If likely direction of bias can be predicted, it is helpful to state this. Direction might be characterized either as being towards (or away from) null, or as being in favour of one of interventions.	Favours intervention / Favours control / Towards null /Away from null / Unpredictable

Bias in classification of interventions			
	3.1 Were intervention groups clearly defined?	A pre-requisite for appropriate comparison of interventions is that interventions are well defined. Ambiguity in definition may lead to bias in classification of participants. For individual-level interventions, criteria for considering individuals to have received each intervention should be clear & explicit, covering issues such as type, setting, dose, frequency, intensity and/or timing of intervention. For population-level interventions (e.g. measures to control air pollution), question relates to whether population is clearly defined, & answer is likely to be 'Yes'.	Y / PY / PN / N / NI
	3.2 Was information used to define intervention groups recorded at start of intervention?	In general, if info about interventions received is available from sources which could not have been affected by subsequent outcomes, then differential misclassification of intervention status is unlikely. Collection of info at time of intervention makes it easier to avoid such misclassification. For population-level interventions (e.g. measures to control air pollution), answer to this question is likely to be 'Yes'.	Y / PY / PN / N / NI
	3.3 Could classification of intervention status have been affected by knowledge of outcome or risk of outcome?	Collection of info at time of intervention may not be sufficient to avoid bias. Way in which data are collected for purposes of NRSI should also avoid misclassification.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table B	Low / Moderate / Serious / Critical / NI
	Optional: What is predicted direction of bias due to measurement of outcomes or interventions?	If likely direction of bias can be predicted, it is helpful to state this. Direction might be characterized either as being towards (or away from) null, or as being in favour of one of interventions.	Favours intervention / Favours control / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions			
	If your aim for this study is to assess effect of assignment to intervention, answer questions 4.1 & 4.2		
	4.1. Were there deviations from intended intervention beyond what would be expected in usual practice?	Deviations which happen in usual practice following intervention (for example, cessation of drug intervention because of acute toxicity) are part of intended intervention & therefore do not lead to bias in ffect of assignment to intervention. Deviations may arise due to expectations of difference between intervention & control (for example because participants feel unlucky to have been assigned to control group & therefore seek active intervention, or components of it, or other interventions). Such deviations are not part of usual practice, so may lead to biased effect estimates. However these are not expected in observational studies of individuals in routine care.	Y / PY / PN / N / NI
	4.2. If Y/PY to 4.1: Were these deviations from intended intervention unbalanced between groups & likely to have affected outcome?	Deviations from intended interventions which do not reflect usual practice will be important if they affect outcome, but not otherwise. Furthermore, bias will arise only if there is imbalance in deviations across two groups.	NA / Y / PY / PN / N / NI
	If your aim for this study is to assess the effect of starting & adhering to intervention, answer questions 4.3 to 4.6		
	4.3. Were important co-interventions balanced across intervention groups?	Risk of bias will be higher if unplanned co-interventions were implemented in way that would bias estimated effect of intervention. Co-interventions will be important if they affect outcome, but not otherwise. Bias will arise only if there is imbalance in such co-interventions between intervention groups. Consider co-interventions, including any pre-specified co-interventions, that are likely to affect outcome & to have been administered in this study. Consider whether these co-interventions are balanced between intervention groups.	Y / PY / PN / N / NI

	4.4. Was intervention implemented successfully for most participants?	Risk of bias will be higher if intervention was not implemented as intended by, for example, health care professionals delivering care during trial. Consider whether implementation of intervention was successful for most participants.	Y / PY / PN / N / NI
	4.5. Did study participants adhere to the assigned intervention regimen?	Risk of bias will be higher if participants did not adhere to intervention as intended. Lack of adherence includes imperfect compliance, cessation of intervention, crossovers to control intervention and switches to another active intervention. Consider available info on proportion of study participants who continued with their assigned intervention throughout follow up, & answer 'No' or 'Probably No' if this proportion is high enough to raise concerns. Answer 'Yes' for studies of interventions which are administered once, so that imperfect adherence is not possible. We distinguish between analyses where follow-up time after interventions switches (including cessation of intervention) is assigned to 1) new intervention or 2) original intervention. 1) is addressed under time varying confounding, & should not be considered further here.	Y / PY / PN / N / NI
	4.6. If N/PN to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate effect of starting & adhering to intervention?	It is possible to conduct analysis which corrects for some types of deviation from intended intervention. Examples of appropriate analysis strategies include inverse probability weighting or instrumental variable estimation. It is possible that paper reports such analysis without reporting info on deviations from intended intervention, but it would be hard to judge such analysis to be appropriate in absence of such info. Specialist advice may be needed to assess studies which used these approaches.	NA / Y / PY / PN / N / NI

		If everyone in one group received co-intervention, adjustments cannot be made to overcome this.	
	Risk of bias judgement	See Table C	Low / Moderate / Serious / Critical / NI
	Optional: What is predicted direction of bias due to deviations from intended interventions?	If likely direction of bias can be predicted, it is helpful to state this. Direction might be characterized either as being towards (or away from) null, or as being in favour of one of interventions.	
Bias due to missing data			
	5.1 Were outcome data available for all, or nearly all, participants?	"Nearly all" should be interpreted as "enough to be confident of findings", & suitable proportion depends on context. In some situations, availability of data from 95% (or possibly 90%) of participants may be sufficient, providing that events of interest are reasonably common in both intervention groups. One aspect of this is that review authors would ideally try & locate analysis plan for study.	Y / PY / PN / N / NI
	5.2 Were participants excluded due to missing data on intervention status?	Missing intervention status may be problem. This requires that intended study sample is clear, which it may not be in practice.	Y / PY / PN / N / NI
	5.3 Were participants excluded due to missing data on other variables needed for analysis?	This question relates particularly to participants excluded from analysis because of missing info on confounders which were controlled for in analysis.	Y / PY / PN / N / NI
	5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are proportion of participants & reasons for missing data similar across interventions?	This aims to elicit whether either i) differential proportion of missing observations or ii) differences in reasons for missing observations could substantially impact on our ability to answer question being addressed. "Similar" includes some minor degree of discrepancy across intervention groups as expected by chance.	NA / Y / PY / PN / N / NI
	5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to presence of missing data?	Evidence for robustness may come from how missing data were handled in analysis & whether sensitivity analyses were performed by investigators, or occasionally from additional analyses performed by systematic reviewers. It is important to assess whether assumptions	NA / Y / PY / PN / N / NI

		employed in analyses are clear & plausible. Both content knowledge & statistical expertise will often be required for this. For instance, use of a statistical method such as multiple imputation does not guarantee an appropriate answer. Review authors should seek naïve (complete-case) analyses for comparison, & clear differences between complete-case & multiple imputation-based findings should lead to careful assessment of validity of methods used.	
	Risk of bias judgement	See Table C	Low / Moderate / Serious / Critical / NI
	Optional: What is predicted direction of bias due to missing data?	If the likely direction of bias can be predicted, it is helpful to state this. Direction might be characterized either as being towards (or away from) null, or as being in favour of one of interventions.	Favours intervention / Favours control / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes			
	6.1 Could outcome measure have been influenced by knowledge of intervention received?	Some outcome measures involve negligible assessor judgment, e.g. all-cause mortality or non-repeatable automated laboratory assessments. Risk of bias due to measurement of these outcomes would be expected to be low.	Y / PY / PN / N / NI
	6.2 Were outcome assessors aware of intervention received by study participants?	If outcome assessors were blinded to intervention status, answer to this question would be 'No'. In other situations, outcome assessors may be unaware of interventions being received by participants despite there being no active blinding by study investigators; answer this question would then also be 'No'. In studies where participants report their outcomes themselves, for example in questionnaire, outcome assessor is study participant. In observational study, answer to this question will usually be 'Yes' when participants report their outcomes themselves.	Y / PY / PN / N / NI
	6.3 Were methods of outcome assessment comparable across intervention groups?	Comparable assessment methods (i.e. data collection) would involve same outcome detection methods & thresholds, same time point, same definition, & same measurements.	Y / PY / PN / N / NI
	6.4 Were any systematic errors in measurement of outcome related to intervention received?	This question refers to differential misclassification of outcomes. Systematic errors in measuring outcome, if present, could cause bias if they are related to intervention or to confounder of intervention-outcome relationship. This will usually be due either to outcome assessors being aware of intervention received or to non-comparability of outcome assessment methods, but there are examples of differential misclassification arising despite these controls being in place.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table C	Low / Moderate / Serious / Critical / NI
	Optional: What is predicted direction of bias due to measurement of	If likely direction of bias can be predicted, it is helpful to state this. Direction might be characterized either as being towards (or away	Favours intervention / Favours control / Towards null /Away

	outcomes?	from) null, or as being in favour of one of interventions.	from null / Unpredictable
Bias in selection of reported result			
	Is reported effect estimate likely to be selected, on basis of results, from...		
	7.1. ... multiple outcome measurements within outcome domain?	For specified outcome domain, it is possible to generate multiple effect estimates for different measurements. If multiple measurements were made, but only one or subset is reported, there is risk of selective reporting on basis of results.	Y / PY / PN / N / NI
	7.2 ... multiple analyses of intervention-outcome relationship?	Because of limitations of using data from non-randomised studies for analyses of effectiveness (need to control confounding, substantial missing data, etc), analysts may implement different analytic methods to address these limitations. Examples include unadjusted & adjusted models; use of final value vs change from baseline vs analysis of covariance; different transformations of variables; continuously scaled outcome converted to categorical data with different cut-points; different sets of covariates used for adjustment; & different analytic strategies for dealing with missing data. Application of such methods generates multiple estimates of effect of intervention versus control on the outcome. If analyst does not pre-specify methods to be applied, & multiple estimates are generated but only one or subset is reported, there is risk of selective reporting on basis of results.	Y / PY / PN / N / NI
	7.3 ... different subgroups?	Particularly with large cohorts often available from routine data sources, it is possible to generate multiple effect estimates for different subgroups or simply to omit varying proportions of original cohort. If multiple estimates are generated but only one or subset is reported, there is risk of selective reporting on basis of results.	Y / PY / PN / N / NI
	Risk of bias judgement	See Table C	Low / Moderate / Serious / Critical / NI

	Optional: What is predicted direction of bias due to selection of reported result?	If likely direction of bias can be predicted, it is helpful to state this. Direction might be characterized either as being towards (or away from) null, or as being in favour of one of interventions.	Favours intervention / Favours control / Towards null /Away from null / Unpredictable

Overall bias			
	Risk of bias judgement	See Table D	Low / Moderate / Serious / Critical / NI
	Optional: What is overall predicted direction of bias for this outcome?		Favours intervention / Favours control / Towards null /Away from null / Unpredictable

Table B. Reaching risk of bias judgements in ROBINS-I: pre-intervention & at-intervention domains

Judgement	Bias due to confounding	Bias in selection of participants into study	Bias in classification of interventions
Low risk of bias (study is comparable to a well-performed randomised trial with regard to this domain)	No confounding expected.	i) All participants who would have been eligible for target trial were included in study; & ii) For each participant, start of follow up & start of intervention coincided.	(i) Intervention status is well defined; & ii) Intervention definition is based solely on info collected at time of intervention.
Moderate risk of bias (study is sound for non randomised study with regard to this domain but cannot be considered comparable to well-performed randomised trial)	i) Confounding expected, all known important confounding domains appropriately measured & controlled for; & ii) Reliability & validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.	i) Selection into study may have been related to intervention & outcome; & Authors used appropriate methods to adjust for selection bias; or ii) Start of follow-up & start of intervention do not coincide for all participants; & a) proportion of participants for which this was case was too low to induce important bias; or b) authors used appropriate methods to adjust for selection bias; or c) review authors are confident that rate (hazard) ratio for effect of intervention remains constant over time.	i) Intervention status is well defined; &ii) Some aspects of assignments of intervention status were determined retrospectively.
Serious risk of bias (study has some important problems)	i) At least one known important domain was not appropriately measured, or not controlled for; or ii) Reliability or validity of measurement of important domain was low enough that we expect serious residual confounding.	i) Selection into study was related (but not very strongly) to intervention & outcome; & This could not be adjusted for in analyses; or ii) Start of follow up & start of intervention do not coincide; & A potentially important amount of follow-up time is missing from analyses;	i) Intervention status is not well defined; or ii) Major aspects of assignments of intervention status were determined in way that could have been affected by knowledge of outcome.

		& Rate ratio is not constant over time.	
Critical risk of bias (study is too problematic to provide any useful evidence on effects of intervention)	i) Confounding inherently not controllable or ii) Use of negative controls strongly suggests unmeasured confounding.	i) Selection into study was very strongly related to intervention & outcome; & This could not be adjusted for in analyses; or ii) Substantial amount of follow-up time is likely to be missing from analyses; & Rate ratio is not constant over time.	(Unusual) Extremely high amount of misclassification of intervention status, e.g. because of unusually strong recall biases.
No information on which to base a judgement about risk of bias for this domain	No information on whether confounding might be present.	No info is reported about selection of participants into study or whether start of follow up & start of intervention coincide.	No definition of intervention or no explanation of source of info about intervention status is reported.

Table C. Reaching risk of bias judgements in ROBINS-I: post-intervention domains

Judgement	**Bias due to deviations from intended intervention**	**Bias due to missing data**	**Bias in measurement of outcomes**	**Bias in selection of reported result**
Low risk of bias (study is comparable to a well-performed randomised trial with regard to this domain)	Effect of assignment to intervention: i) Any deviations from intended intervention reflected usual practice; or ii) Any deviations from usual practice were unlikely to impact on outcome.	i) Data were reasonably complete; or ii) Proportions of & reasons for missing participants were similar across intervention groups; or iii) Analysis addressed missing data & is likely to have removed any risk of bias.	i) Methods of outcome assessment were comparable across intervention groups; & ii) Outcome measure was unlikely to be influenced by knowledge of intervention received by study participants (i.e. is objective) or outcome assessors were unaware of intervention received by study participants; & iii) Any error in measuring outcome is unrelated to intervention status.	There is clear evidence (usually through examination of pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses & subcohorts.
Moderate risk of bias (study is sound for non randomised study with regard to this domain but cannot be considered comparable to a well-performed randomised trial)	Effect of assignment to intervention: There were deviations from usual practice, but their impact on outcome is expected to be slight. Effect of starting & adhering to intervention: i) There were deviations from	i) Proportions of & reasons for missing participants differ slightly across intervention groups; & ii) Analysis is unlikely to have removed risk of bias arising from missing data.	i) Methods of outcome assessment were comparable across intervention groups; & ii) Outcome measure is only minimally influenced by knowledge of intervention received by study participants; &	i) Outcome measurements & analyses are consistent with a priori plan; or are clearly defined & both internally & externally consistent; & ii) There is no indication of selection of reported analysis from among multiple analyses;

	intended intervention, but their impact on outcome is expected to be slight. or ii) Important co-interventions were not balanced across intervention groups, or there were deviations from intended interventions (in terms of implementation and/or adherence) which were likely to impact on outcome; & Analysis was appropriate to estimate effect of starting & adhering to intervention, allowing for deviations (in terms of implementation, adherence & co-intervention) which were likely to impact on outcome.		iii) Any error in measuring outcome is only minimally related to intervention status.	& iii) There is no indication of selection of cohort or subgroups for analysis & reporting on basis of results.
Serious risk of bias (study has some important problems)	Effect of assignment to intervention: There were deviations from usual practice which were unbalanced between intervention groups & likely to have affected outcome. Effect of starting & adhering to intervention:	i) Proportions of missing participants differ substantially across interventions; or Reasons for missingness differ substantially across interventions; & ii) Analysis is unlikely to have removed risk of	i) Methods of outcome assessment were not comparable across intervention groups; or ii) Outcome measure was subjective (i.e. vulnerable to influence by knowledge of intervention received by study	i) Outcomes are defined in different ways in methods & results sections, or in different publications of study; or ii) There is igh risk of selective reporting from among multiple analyses; or

	i) Important co-interventions were not balanced across intervention groups, or there were deviations from intended interventions (in terms of implementation and/or adherence) which were likely to impact on outcome; & (ii) Analysis was not appropriate to estimate effect of starting & adhering to intervention, allowing for deviations (in terms of implementation, adherence & co-intervention) which were likely to impact on outcome.	bias arising from missing data; or Missing data were addressed inappropriately in analysis; or nature of missing data means that risk of bias cannot be removed through appropriate analysis.	participants); & Outcome was assessed by assessors aware of intervention received by study participants; or (iii) Error in measuring outcome was related to intervention status.	iii) Cohort or subgroup is selected from larger study for analysis & appears to be reported on basis of results.
Critical risk of bias (study is too problematic to provide any useful evidence on effects of intervention)	Effect of assignment to intervention: There were substantial deviations from usual practice that were unbalanced between intervention groups & likely to have affected outcome. Effect of starting & adhering to intervention: (i) There were substantial imbalances in important co-interventions	i) (Unusual) There were critical differences between interventions in participants with missing data; & ii) Missing data were not, or could not, be addressed through appropriate analysis.	Methods of outcome assessment were so different that they cannot reasonably be compared across intervention groups.	i) There is evidence or strong suspicion of selective reporting of results; & ii) Unreported results are likely to be substantially different from reported results.

	across intervention groups, or there were substantial deviations from intended interventions (in terms of implementation and/or adherence) which were likely to impact on outcome; & ii) Analysis was not appropriate to estimate effect of starting & adhering to intervention, allowing for deviations (in terms of implementation, adherence & co-intervention) which were likely to impact on outcome.			
No info on which to base a judgement about risk of bias for this domain	No info is reported on whether there is deviation from intended intervention.	No info is reported about missing data or potential for data to be missing.	No info is reported about methods of outcome assessment.	There is too little info to make judgement (for example, if only abstract is available for study).

Table D. Interpretation of domain-level & overall risk of bias judgements in ROBINS-I*

Judgement	Within each domain	Across domains	Criterion
Low risk of bias	Study is comparable to well-performed randomised trial with regard to this domain.	Study is comparable to well performed randomised trial.	Study is judged to be at low risk of bias for all domains.
Moderate risk of bias	Study is sound for non-randomised study with regard to this domain but cannot be considered comparable to a well-performed randomised trial.	Study provides sound evidence for a non-randomised study but cannot be considered comparable to well-performed randomised trial.	Study is judged to be at low or moderate risk of bias for all domains.
Serious risk of bias	Study has some important problems in this domain.	Study has some important problems.	Study is judged to be at serious risk of bias in at least one domain, but not at critical risk of bias in any domain.
Critical risk of bias	Study is too problematic in this domain to provide any useful evidence on effects of intervention.	Study is too problematic to provide any useful evidence & should not be included in any synthesis.	Study is judged to be at critical risk of bias in at least one domain.
No info No info on which to	No info on which to base judgement about risk of bias for this domain.	No infon on which to base judgement about risk of bias.	There is no clear indication that study is at serious or critical risk of bias & there is lack of information in one or more key domains of bias (a judgement is required for this).

*Also saved as table 2 in main article.