BMJ Open

# Do social protection programmes have a causal effect on suicide mortality? A protocol for a systematic review and meta-analysis

Chungah Kim,[1] Karanpreet Azra,[1] Celine Teo,[1,2] Andrew Nielsen,[1,2] Zachary Bellows,[1] Thomas Young,[1] Antony Chum [1,2,3]

**To cite:** Kim C, Azra K, Teo C, *et al*. Do social protection programmes have a causal effect on suicide mortality? A protocol for a systematic review and meta-analysis. *BMJ Open* 2022;**12**:e054677. doi:10.1136/bmjopen-2021-054677

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-054677).

Received 24 June 2021
Accepted 15 December 2021



[1]Department of Applied Health Sciences, Brock University, St. Catharines, Ontario, Canada
[2]MAP Centre for Urban Health Solutions, St. Michael's Hospital Li Ka Shing Knowledge Institute, Toronto, Ontario, Canada
[3]Epidemiology Divison, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

**Correspondence to**
Dr Antony Chum;
antony.chum@utoronto.ca

## ABSTRACT

**Introduction** Recent international and national strategies to reduce suicide mortality have suggested that social protection programmes may be an effective multisectoral response given the link between material deprivation and suicides in observational studies. However, there is a lack of evidence on the causal relationship between social protection programmes and suicide, which may hinder substantial national budget reallocations necessary to implement these policies. Social protection programmes are government interventions that ensure adequate income now and in the future, through changes to earned income (eg, minimum wage increase) or social security (via cash transfers or cash equivalents). Our review aims to evaluate the existing evidence on a causal relationship between social protection programmes and suicide mortality by examining all relevant experimental and quasi-experimental studies between January 1980 and November 2021.

**Methods and analysis** The review will be conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guideline. We will search references published between 1 January 1980 and 30 November 2021 in 10 electronic databases, including MEDLINE (PubMed), PsycINFO, EMBASE and Applied Social Sciences Index Abstracts. Seven reviewers will independently participate in screening studies from titles, abstracts and full texts across all the stages. Experimental (ie, randomised controlled trials) and quasi-experimental studies (ie, non-randomised interventional studies) written in English, French, Spanish, German, Chinese, Korean and Japanese examining the impact of income security programmes on suicide mortality were included. Meta-analyses will be conducted if there are at least three studies with similar income security programmes.

**Ethics and dissemination** Our proposed review does not require ethical approval. In collaboration with our community partners, we will develop a policy brief for stakeholders to support efforts to implement social protection programmes to help prevent suicides. Our findings will be presented at conferences, published in a peer-reviewer journal and promoted on social media platforms.

**PROSPERO registration number** CRD42021252235.

## Strengths and limitations of this study

- By focusing on studies that use non-randomised/randomised experimental designs, our review is able to synthesise causal evidence of the effect of social protection programmes on suicide prevention.
- Our inclusion of a comprehensive set of social protection programmes will provide policy-makers novel insights on a range of diverse programmes for decision-making.
- Since the review will include a range of different social protection programmes, there is a greater chance that we will find heterogeneous effects.
- There is potential for reviews of secondary data to have publication bias, where published studies are more likely to report significant findings rather than null findings.

## INTRODUCTION

Suicide accounts for 1.4% of deaths worldwide,[1] and many more suicides are likely misclassified as unintentional or undeterminable injuries.[2] In 2014, the WHO formalised a global strategy to prevent suicides by calling on member states to implement multisectoral action, such as restricting common means (including pesticides, firearms and certain medications), reducing inappropriate media reporting, increasing access to services to manage mental illnesses, introducing appropriate alcohol policies and reducing stigma and increasing social support at the community level.[1] While poverty and material deprivation are well-established risk factors of suicides,[3] social protection programmes to reduce the risk of socioeconomic adversity on suicides have not been featured as a mainstream intervention in the global discourse. Social protection programmes are government interventions that ensure adequate income now

and in the future, through changes to earned income (eg, minimum wage increase) or social security (via cash transfers or cash equivalents).[4 5] Social protection programmes include a range of government programmes aimed at (partially) ameliorating the negative impact of predictable and unpredictable risks (eg, chronic poverty, dependency in childhood, frailty in old age, job loss, sickness/injuries and family breakdown). These programmes aim to compensate for income losses associated with these risks, and enable people to return to their everyday life. The impact of social protection programmes is not restricted to poverty alleviation but may include reducing income inequality and promoting the overall well-being of societies.

In 2017, the US Centers for Disease Control and Prevention developed a national suicide prevention strategy that included a focus on policies to strengthen economic support as part of the national multisectoral response to suicides.[6] This publication reflects a paradigm shift among suicide prevention strategies since no similar documents to date, at the global or national government levels, have recommended the promotion of social protection as part of comprehensive multisectoral action. Despite the new policy direction for suicide prevention, and the wider recognition that poverty, income loss and material deprivation are risk factors for suicide,[7] there are currently a lack of systematic reviews that evaluate the effectiveness of social protection programmes to reduce suicides. In order to provide strong evidence to justify the substantial national budget reallocations necessary to implement these policies, our study will systematically review evidence to evaluate the causal link between various social protection programmes and suicide mortality.

## Economic insecurity and suicides in observational studies

The association between material deprivation and suicide is well established in psychiatric epidemiology literature.[8–10] In a systematic review of psychiatric and socioeconomic risk factors for suicide in high-income countries, low income was associated with an increased relative risk of suicide by 2.18 in men and by 1.45 in women.[8] Similar associations have been identified in systematic reviews with evidence from low-income and middle-income countries. One review investigated suicide and poverty, and found that worse economic status and diminished wealth were positively associated with suicidal behaviour and ideation at the individual level, although these trends were not observed at the country level.[9] Across low-income and middle-income South and South-East Asian countries, another review found a consistent association between financial strain and suicide, where those in low socioeconomic positions had a threefold increased risk of suicide.[10]

Despite the consistent findings on the association between economic insecurity and suicide risk, observational studies have a limited ability to draw causal inference.[11] Potential shortcomings in these observational studies include: (1) the inability of case–control and cohort studies to effectively address potential endogeneity (eg, preexisting psychiatric disorder or genetic vulnerability as a common cause of material deprivation and suicide) and (2) suicide-related mortalities are rare outcomes in individual-level cohort studies and could result in an underpowered statistical analysis. Furthermore, observational studies cannot be used to infer the effectiveness of social protection programmes as part of suicide prevention strategy.

Randomised controlled trials (RCTs, ie, experimental studies) can resolve these limitations by ensuring that treatment assignment is exogenous (through random assignment). Exogeneity of exposure can help rule out selection bias and confounding since the exogenous exposure (eg, through random assignment) is not influenced by the outcome of interest or any variable associated with the outcome. Despite the high-quality standards of RCTs, they are difficult to conduct in non-clinical settings, since suicide events are extremely rare. Where manipulation to the exposure is not an option, quasi-experimental studies (ie, natural experiments) can be a viable alternative for causal inference since exogeneity can be established through other means such as through nature, policy and practice.[12 13] For example, the exogenous variation could be changes in levels of income driven by legislation and implementation of social protection programmes. Thus, recent studies have used exogenous variations in the time and the extent of the benefit level, naturally generated by the legislation of social protection programmes to identify the causal effects of increased income on suicide mortality.[14 15]

Although a growing body of literature examines the role of social and economic policy on suicide, there has been only one narrative review of the relationship between social protection programmes and suicide.[16] Social protection programmes include: however, the previous review (1) included studies that did not use quasi-experimental or RCT designs, and (2) did not evaluate quality of evidence; therefore, it had limited ability to provide evidence for causal inference. To address these limitations, our review will aim to identify all existing RCTs and quasi-experimental studies that examine social protection programmes conducted since 1980 on suicide mortality. We will only focus on mortality since individual-level socioeconomic positions may have a differential impact on non-fatal (eg, suicidal ideation and attempts) and fatal suicidal events.[17] Our systematic review of RCT and quasi-experimental studies on the impact of social protection programmes on suicides will have

the following objectives: (1) to provide evidence to support the decision-making process with regards to the implementation of social protection programmes as a core part of suicide prevention strategy and (2) to establish the broader effect of income on suicide by exploiting income security programmes as an exogenous shift. Our systematic review will answer the following research question: do social protection programmes have a causal effect on suicide mortality?

## METHODS

### Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

We conducted preliminary searches in May 2021 and registered the current protocol on the PROSPERO database on 4 May 2021. The current review protocol is written according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols guidelines. Revision history and any amendment to the protocol are available through PROSPERO. The review will start in December 2021.

### Definitions of key terms

#### Intervention: social protection programmes

Social protection programmes in the review are based on the widely recognised definition from Norton *et al*, which includes public actions that address 'the deprivation and vulnerabilities of the poor, and also with the needs of the non-poor for security in the face of shocks and the particular demands of different stages of the life cycle' (p22).[18] We also drew on a synthesised report (funded by the UK Department for International Development) aimed at summarising the evidence base on when and how social protection programmes can be used to minimise negative shocks in the global context.[19] Specifically, according to the report, social protection programmes consist of social assistance (ie, unremarkable tax-financed transfers in cash, vouchers or in-kind; fee waivers and subsidies), social insurance (ie, contributory schemes providing support in the event of contingencies, such as illness, injury, unemployment, old age and disability), social care services for individuals facing risks of social exclusion, and active (ie, strengthening skills and competencies to promote labour market participation) and passive (ie, ensuring minimum employment standards) labour market programmes. The specific programmes and policies with general terms and synonyms related to social protection programmes are presented in figure 1, and have been derived from a prior synthesis report.[20]

### Method: RCTs (ie, experimental study) and quasi-experimental studies

Our review will include both RCT and quasi-experimental studies. RCT refers to a form of intervention study in which participants are assigned to the intervention at random, assuming that all aspects other than assignment of the intervention are identical. The purpose of random assignment in an experimental study is to ensure both treatment and control groups are equivalent so that any preexisting attribute does not affect the outcome or any factor associated with the outcome (ie, to achieve exogeneity).[21] Although treatment is not randomly assigned, a well-defined quasi-experimental study can achieve exogeneity through a 'force of nature'[21] (ie, where the occurrence of an event with a natural cause) or a policy change

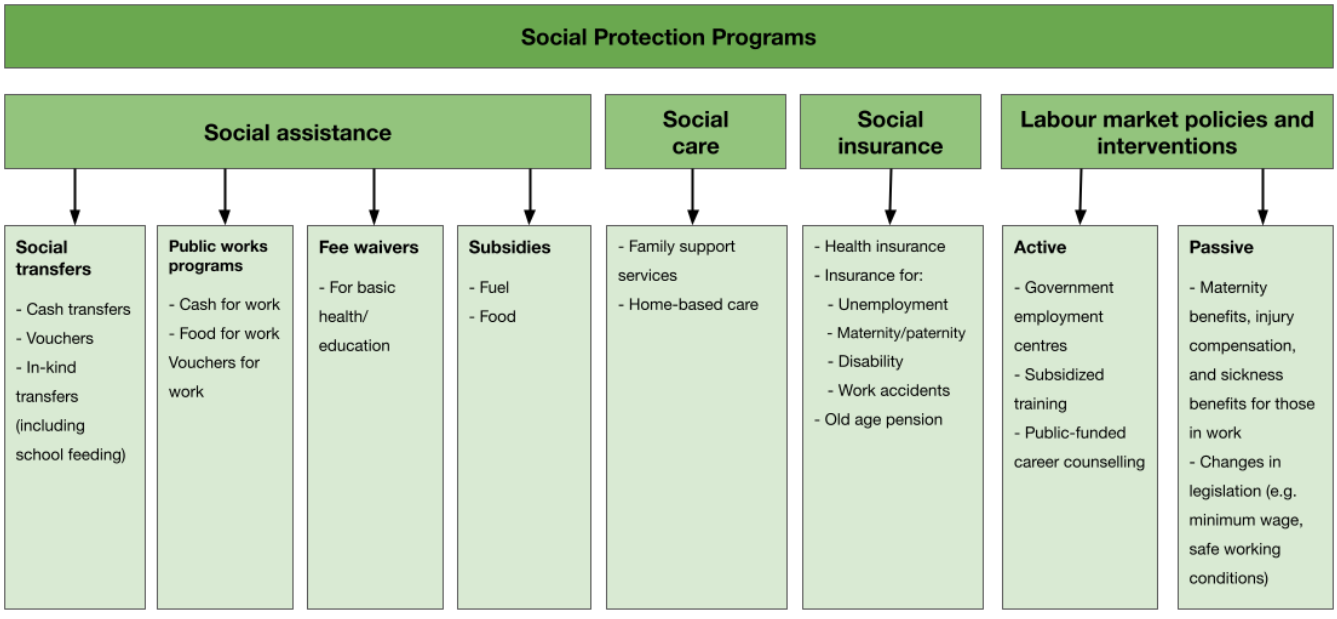


**Figure 1** Subtypes of social protection programmes, modified figure adapted from O'Brien *et al*.[20]

(ie, where exposure is allocated without the deliberate manipulation by researchers[21]).

### Suicide mortality

Suicide mortality refers to deaths from intentional self-harm, extracted using the International Classification of Diseases V.10 (ICD-10) is coded as X60–X84. We additionally include any (subset) of the following codes as potential suicide mortality: Y10–Y34 (undetermined deaths) and Y87.0 (sequelae of intentional self-harm, assault and events of undetermined intent). Many previous studies[22–24] have included undetermined deaths and sequelae of international self-harm as suicide mortality outcome because prior studies found that a large proportion of them are misclassified suicide cases. For instance, there is strong evidence that injury-related and poisoning-related undetermined deaths are likely to be suicides. Therefore, we included studies that used a broader definition of suicide outcomes, beyond X60–X84. For studies published before the release of the ICD-10, the above codes will be matched to the ICD-8 and ICD-9 equivalents. We will not exclude a study if ICD codes were not used. If a study does not use ICD or other standardised diagnostic codes at the full text review stage, we will try our best to match what is written in the paper to the above ICD definition (eg, contacting the author to confirm whether the deaths included in the study matches with the definitions we used above). Variability in the identification of suicides will be noted in the results of the review.

### Eligibility criteria

We will include all published studies, preprint studies and dissertations written in English. Studies in low-income, middle-income and high-income countries will be included. We will exclude studies that evaluated healthcare-related programmes or policy (eg, medical subsidy, Medicare and drug subsidy). While transfers and benefits directly related to healthcare utilisation are excluded, the use of eligibility for these subsidies as a criterion for other transfers and benefits is acceptable. For example, in South Korea, a medical aid programme, which provides medical service for the bottom 3%–4% of households of income, is often used as a means-testing criteria for social protection programmes.[25] Studies conducted prior to 1980 are excluded. Studies that do not have a specific government or non-government funded intervention or policy, such as those that investigated the impact of general macroeconomic changes (eg, economic boom or recession) will not be included since these changes are not considered exogenous that can be tested using causal inference (ie, quasi-experimental methods).

### Search strategy

#### Databases

Starting December 2021, the reviewers will use the following 10 databases to search for studies published between 1 January 1980 and 30 November 2021: MEDLINE (PubMed), PsycINFO, EMBASE, Applied Social Sciences Index, Grey Literature Report, Scopus (Elsevier), the Cochrane Central Register of Controlled Trials, ProQuest Dissertation Dissertation Database, EconLit and Research Papers in Economics. The electronic databases were selected for relevance to the research question as well as being frequently used in systematic literature searches. We will conduct additional hand-searching for references in relevant studies and key journals.

**Table 1** Key terms for social protection interventions and policies

| Types of social protection programmes | Specific programmes/policies or synonyms |
|---|---|
| Social assistance | Social transfer, public works programme, fee waiver, housing support, housing benefit, housing subsidy, public housing, welfare, social policy, social assistance, social security, food stamp, food assistance, food aid, in-kind transfer, disability benefit, family allowance, child benefit, income benefit, income supplement, income support, income maintenance, cash transfer, income security, basic income, guaranteed income and cash-like transfers |
| Social care | Social care, family support, childcare, eldercare, residential care and home care |
| Social insurance | Unemployment insurance, employment insurance, pension, sickness benefit, income benefit and injury compensation |
| Labour market programmes | Minimum wage, (earned) income tax credit, maternity benefits, active labour market, employment service, wage subsidy, vocational training, job-search services and work sharing |
| Other related terms | Austerity, deaths of despair and poverty reduction |

#### Search terms

The two search terms for suicide-related studies include suici* and self-harm to ensure that studies examining suicide mortality are captured. The search terms for social protection were identified based on the goal of covering a range of specific programmes that fall under our definition of a social protection programme, and have been identified through previous literature.[4 26] For the purposes of presenting and organising the terms, social protection programmes are categorised into the following five groups based on a prior study (see table 1): labour market programmes, targeted social assistance, social insurance, other conditional/unconditional cash/cash-like transfers and general programmes. Related terms and specific modelling related to RCT and quasi-experimental studies are listed below (see table 2). See online supplemental file 1 for detailed instructions on how these terms are operationalised in each database.

Table 2 Search terms for RCT and quasi-experimental studies

| Study specifications | Related terms |
|---|---|
| Quasi-experimental study | Natural experiment, quasi-experiment, non-randomised, instrument, interrupted time series, propensity score, sharp design, fuzzy design, matched control, synthetic control, regression discontinuity and inverse probability weight |
| Randomised experimental study (RCT) | RCTs, randomised control trials, RCT, field experiment, experiment, social experiment and random |
| Terms for either RCT or quasi-experimental studies | Sibling, Mendelian randomisation, controlled before and after, difference in difference, difference study, exogenous variation, counterfactual, Rubin causal model and potential outcome |

RCT, randomised controlled trial.

### Study selection

We will import all the citations to a citation manager (ie, Zotero) for deduplication and then to an online software programme for systematic review (ie, Covidence) for screening. At stage 1, all authors (AC, CK, CT, KA, AN, ZB and TY) will screen all of the titles and abstracts to identify relevant studies by checking whether the target programme, outcome and methods were used. Each title and abstract are required to be screened by two authors, and any discrepancies that arise will be resolved through a discussion between all authors on its relevance based on the inclusion/exclusion criteria. At stage 2, another reviewer (CK) will screen a random sample of studies that were excluded at stage 1 with no discrepancies (ie, a 10% sample of the excluded studies). Any studies that are identified as inappropriately excluded at stage 1 will be discussed among CK, ZB, KA and AN, with another reviewer (AC) intervening to resolve any arising discrepancy. At stage 3, for the chosen studies screened through titles and abstracts, all team members will be working collaboratively to review the full texts (comparing results throughout the process), assess the eligibility of the texts and then appraise the quality of the included studies where results are determined by consensus. We will contact the authors if additional study information is required.

## Strategy for data synthesis

### Data extraction

We will create a table to provide a clear description of the data extracted from the selected studies, which will include the authors, years of publication, titles, populations, designs, data sources, data years, analytic approaches and results (see online supplemental file 2). The effect sizes and quality of the studies will be reviewed and critiqued. Data will be extracted by ZB, KA, AN and TY.

### Risk of bias (quality) assessment

All authors will use Cochrane Collaboration risk of bias (RoB) V.2.0 tool[27] for RCTs and the ROBINS-I ("Risk of Bias in Non-randomised Studies - of Interventions") for quasi-experiments and natural experiments[28] (see online supplemental file 3), for the final set of included studies after the full-text screening. Any disagreements will be discussed and resolved by another reviewer (AC). The RoB V.2.0 analyses six domains: random sequence generation, allocation concealment, blinding of patients and personnel, blinding of outcome assessor, incomplete outcome data and selective outcome reporting. The ROBINS-I consists of seven components assessing the following: bias due to confounding, selection of participants, classification of interventions, departure from intended interventions, missing data, measurement of outcomes and selection of reported results.

## Systematic narrative review and meta-analysis

We will provide a summary table of the included studies with effect sizes and details on programme specifications. We will consider each programme's economic contexts (eg, low-income or middle-income or high-income countries), study design (eg, use of individual-level or population-level data), types of programme (eg, universality, delivery and conditionality) and underlying mechanisms, and use this information to analytically categorise these programmes. The results will be summarised separately for each programme category. Based on these factors, if we have at least three studies of a similar programme, we will perform a meta-analysis. Otherwise, only a systematic narrative review will be performed. If we can conduct a meta-analysis, we will examine the heterogeneity of studies, and their sources, and conduct a fixed-effects or random-effects model based on the level of heterogeneity. We will also check for publication bias, and perform sensitivity analyses if necessary. All statistical analyses will be conducted using R. The strength of the body of evidence will be assessed using the Grading of Recommendations, Assessment, Development and Evaluations framework.

## Ethical considerations and disseminations

Ethical approval is not required for the present study since the review will be a synthesis of existing secondary data. In collaboration with our community partners, we will develop a policy brief for key stakeholders. Therefore, the study will provide policy-makers with evidence to modify or implement social protection programmes to prevent suicides. Findings from the review can be used to inform future research such as impact evaluation of social protection programmes. Our findings will be presented at international conferences and published in a peer-reviewed journal. The findings will also be promoted through social media platforms, such as Twitter and YouTube.

## DISCUSSION

The proposed systematic review will be the first to summarise the causal effects of social protection programmes on suicide mortality based on prior RCTs and quasi-experiments. Our review has the following policy and theoretical implications: first, evidence from our study could be used to support multisectoral suicide prevention strategies by clarifying the role of social protection programmes as a core component of these strategies in low-income–high-income countries. We recognise the numerous ways in which social protection programmes are implemented, and we include a wide range of these programmes to ensure a comprehensive review of relevant studies. Second, the review will contribute to a richer theoretical understanding of the causal impacts of income (ie, economic security) on suicide. By examining exogenous changes in income within RCTs and quasi-experimental studies, we can help identify possible causal links and mechanisms between income and suicide risk. In addition, to ensure that our findings reflect a valid representation of existing evidence, our study design is compliant with recommended and validated methods guidelines and will adhere to a systematic and transparent approach.

The proposed review has some limitations that we will take into consideration. First, since our review will include a range of different social protection programmes, there is a greater chance that we will find heterogeneous effects. Nevertheless, we believe the need to review the range of selected studies is significant to suicide prevention policy development. Second, reviews of secondary data may have publication bias, where published studies are more likely to report significant findings rather than null findings. We will minimise the publication bias risk by trying to find unpublished studies (eg, grey literature and dissertations) and conduct additional hand-searching in references. Funnel plots will be included to visually identify the presence of potential bias. Third, the review is limited to only include studies published in seven languages, which may exclude studies published in other languages.

## CONCLUSION

While traditional suicide prevention strategies have focused on individual-level and clinical inventions, social protection programmes may offer a unique solution to further reduce suicides. However, the current lack of evidence on their efficacy may be a barrier to their wider implementation. Our review will evaluate the evidence of a causal relationship between social protection programmes and suicide mortality, which may provide strong evidence for shaping the future of suicide prevention strategies.

**Contributors** CK conceived the idea. CK, CT, KA, AN, ZB and TY drafted the manuscript. AC provided feedback. CT, AN and TY contributed to the quality appraisal assessment strategy and data extraction criteria. CK, KA, AC and ZB developed the search strategy. All authors provided comments and amendments. All authors (CK, CT, KA, AC, AN, ZB and TY) contributed to the development of the selection criteria and approved the final manuscript.

**Funding** This study was supported by SSHRC insight grant (435-2020-1086, principal investigator (PI): AC) and CIHR project grant (421369, PI: AC). The funders have no role in the development or intellectual contribution to the protocol.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.



**ORCID iD**

Antony Chum http://orcid.org/0000-0002-6756-7031

## REFERENCES

1 Saxena S, Krug EG, Chestnov O, eds. *Preventing suicide: a global imperative*. Geneva: World Health Organization, 2014.
2 Mishara BL, Weisstub DN. The legal status of suicide: a global review. *Int J Law Psychiatry* 2016;44:54–74.
3 Cairns J-M, Graham E, Bambra C. Area-level socioeconomic disadvantage and suicidal behaviour in Europe: a systematic review. *Soc Sci Med* 2017;192:102–11.
4 Sun S, Huang J, Hudson DL, *et al*. Cash transfers and health. *Annu Rev Public Health* 2021;42:363–80.
5 ILO. *Economic security for a better world*. International Labour Organization, 2004.
6 Stone DM, Holland KM, Bartholow B. *Preventing suicide: a technical package of policies, programs, and practices*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 2017.
7 O'Connor RC, Pirkis J, eds. *The international handbook of suicide prevention*. Chichester, UK:: John Wiley & Sons, Ltd, 2016.
8 Li Z, Page A, Martin G, *et al*. Attributable risk of psychiatric and socio-economic factors for suicide from individual-level, population-based studies: a systematic review. *Soc Sci Med* 2011;72:608–16.
9 Iemmi V, Bantjes J, Coast E, *et al*. Suicide and poverty in low-income and middle-income countries: a systematic review. *Lancet Psychiatry* 2016;3:774–83.
10 Knipe DW, Carroll R, Thomas KH, *et al*. Association of socio-economic position and suicide/attempted suicide in low and middle income countries in South and south-east Asia – a systematic review. *BMC Public Health* 2015;15:1055.
11 Turecki G, Brent DA. Suicide and suicidal behaviour. *Lancet* 2016;387:1227–39.
12 Bärnighausen T, Tugwell P, Røttingen J-A, *et al*. Quasi-experimental study designs series—paper 4: uses and value. *J Clin Epidemiol* 2017;89:21–9.
13 Frenk J, Gómez-Dantés O. Quasi-experimental study designs series—paper 3: systematic generation of evidence through public policy evaluation. *J Clin Epidemiol* 2017;89:17–20.
14 Dow WH, Godøy A, Lowenstein C, *et al*. Can labor market policies reduce deaths of despair? *J Health Econ* 2020;74:102372.
15 Kaufman JA, Salas-Hernández LK, Komro KA, *et al*. Effects of increased minimum wages by unemployment rate on suicide in the USA. *J Epidemiol Community Health* 2020;74:219–24.
16 Kim C. The impacts of social protection policies and programs on suicide: a literature review. *Int J Health Serv* 2018;48:512–34.
17 Kim M-H, Jung-Choi K, Jun H-J, *et al*. Socioeconomic inequalities in suicidal ideation, parasuicides, and completed suicides in South Korea. *Soc Sci Med* 2010;70:1254–61.

18 Norton A, Conway T, Foster M. *Social protection concepts and approaches: implications for policy and practice in international development*. Citeseer, 2001.
19 Watson C, Lone T, Qazi U. *Shock-Responsive social protection systems research*. Oxford, 2017.
20 O'Brien C, Scott Z, Smith G. *Shock-Responsive social protection systems research: synthesis report*. Oxford, UK: Oxford Policy Management, 2018.
21 de Vocht F, Katikireddi SV, McQuire C, *et al*. Conceptualising natural and quasi experiments in public health. *BMC Med Res Methodol* 2021;21:32.
22 Björkenstam C, Johansson L-A, Nordström P, *et al*. Suicide or undetermined intent? A register-based study of signs of misclassification. *Popul Health Metr* 2014;12:1–11.
23 Mittendorfer-Rutz E, Rasmussen F, Wasserman D. Restricted fetal growth and adverse maternal psychosocial and socioeconomic conditions as risk factors for suicidal behaviour of offspring: a cohort study. *Lancet* 2004;364:1135–40.
24 Burrows S, Auger N, Gamache P, *et al*. Influence of social and material individual and area deprivation on suicide mortality among 2.7 million Canadians: a prospective study. *BMC Public Health* 2011;11:1–11.
25 Yoo K-B, Ahn H-U, Park E-C, *et al*. Impact of co-payment for outpatient utilization among medical aid beneficiaries in Korea: a 5-year time series study. *Health Policy* 2016;120:960–6.
26 Simpson J, Brown H, Bell Z, *et al*. Investigating the relationship between changes in social security benefits and mental health: a protocol for a systematic review. *BMJ Open* 2020;10:e035993.
27 Sterne JAC, Savović J, Page MJ, *et al*. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
28 Sterne JA, Hernán MA, Reeves BC, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.

**Supplementary File S1: Search strategy**

The research results were restricted by date to include results between January 1980 and November 30, 2021. No other restrictions were applied.

**Pubmed**
((suici* OR self-harm) AND (minimum wage OR income tax?credit OR maternity benefit* OR active lab* OR employment service OR wage subsidy OR vocational training OR job?search service* OR work sharing OR housing support OR housing benefit* OR housing subsidy OR public housing OR welfare OR social policy OR social assistance OR social security OR food stamp OR food assistance OR food aid OR in?kind transfer OR social?transfer OR public works program OR fee waiver* OR family allowance OR child benefit* OR social care OR family support OR childcare OR eldercare OR residential care OR home care OR disability benefit* OR unemployment insurance OR employment insurance OR pension OR sickness benefit* OR income benefit* OR injury compensation OR income supplement OR income support OR income maintenance OR cash?transfer* OR income security OR basic income OR guaranteed income OR austerity OR deaths of despair OR poverty reduction)) AND (experiment* OR randomi?ed control* OR RCT OR randomi?ed OR non?randomi?ed OR interrupted time?series OR propensity?score OR sharp?design OR fuzzy?design OR matched?control OR synthetic control OR regression?discontinuity OR inverse?probability weight OR mendelian?randomi?ation OR controlled before and after OR difference?in?difference* OR difference?stud* OR exogenous varia* OR counterfactual OR rubin causal model OR potential outcome)

**PsycInfo**
Any Field: “suici*” OR “self-harm” *AND* Any Field: “minimum wage” OR “income tax?credit” OR “maternity benefit*” OR “active lab*” OR “employment service” OR “wage subsidy” OR “vocational training” OR “job?search service*” OR “work sharing” OR “housing support” OR “housing benefit*” OR “housing subsidy” OR “public housing” OR “welfare” OR “social policy” OR “social assistance” OR “social security” OR “food stamp” OR “food assistance” OR “food aid” OR “in?kind transfer” OR “social?transfer” OR “public works program” OR “fee waver*” OR “family allowance” OR “child benefit*” OR “social care” OR “family support” OR “childcare” OR “eldercare” OR “residential care” OR “home care” OR “disability benefit*” OR “unemployment insurance” OR “employment insurance” OR “pension” OR “sickness benefit*” OR “income benefit*” OR “injury compensation” OR “income supplement” OR “income support” OR “income maintenance” OR “cash?transfer*” OR “income security” OR “basic income” OR “guaranteed income” OR “austerity” OR “deaths of despair” OR “poverty reduction” *AND* Any Field: “experiment*” OR “randomi?ed control*” OR “RCT” OR “randomi?ed” OR “non?randomi?ed” OR “interrupted time?series” OR “propensity?score” OR “sharp?design” OR “fuzzy?design” OR “matched?control” OR “synthetic control” OR “regression?discontinuity” OR “inverse?probability weight” OR “mendelian?randomi?ation” OR “controlled before and after” OR “difference?in?difference*” OR “difference?stud*” OR “exogenous varia*” OR “counterfactual” OR “rubin causal model” OR “potential outcome”

**Embase**
((suici* or self-harm) and (minimum wage or income tax?credit or maternity benefit* or active lab* or employment service or wage subsidy or vocational training or job?search service* or work sharing OR housing support or housing benefit* or housing subsidy or public housing or welfare or social policy or social assistance or social security or food stamp or food assistance or food aid or in?kind transfer or social?transfer or public works program or fee waiver* or family

allowance or child benefit* or social care or family support or childcare or eldercare or residential care or home care or disability benefit* or unemployment insurance or employment insurance or pension or sickness benefit* or income benefit* or injury compensation or income supplement or income support or income maintenance or cash?transfer* or income security or basic income or guaranteed income or austerity or deaths of despair or poverty reduction) and (((experiment* or randomi?ed control* or RCT or randomi?ed or non?randomi?ed or interrupted time?series or propensity?score or sharp?design or fuzzy?design or matched?control or synthetic control or regression?discontinuity or inverse?probability weight or mendelian?randomi?ation or controlled before) and after) or difference?in?difference* or difference?stud* or exogenous varia* or counterfactual or rubin causal model or potential outcome)).af.

**Applied Social Sciences Index and Abstracts (ASSIA)**
noft("suici*" OR "self-harm") AND noft("minimum wage" OR "income tax?credit" OR "maternity benefit*" OR "active lab*" OR "employment service" OR "wage subsidy" OR "vocational training" OR "job?search service*" OR "work sharing" OR "housing support" OR "housing benefit*" OR "housing subsidy" OR "public housing" OR "welfare" OR "social policy" OR "social assistance" OR "social security" OR "food stamp" OR "food assistance" OR "food aid" OR "in?kind transfer" OR "social?transfer" OR "public works program" OR "fee waiver*" OR "family allowance" OR "child benefit*" OR "social care" OR "family support" OR "childcare" OR "eldercare" OR "residential care" OR "home care" OR "disability benefit*" OR "unemployment insurance" OR "employment insurance" OR "pension" OR "sickness benefit*" OR "income benefit*" OR "injury compensation" OR "income supplement" OR "income support" OR "income maintenance" OR "cash?transfer*" OR "income security" OR "basic income" OR "guaranteed income" OR "austerity" OR "deaths of despair" OR "poverty reduction") AND noft("experiment*" OR "randomi?ed control*" OR "RCT" OR "randomi?ed" OR "non?randomi?ed" OR "interrupted time?series" OR "propensity?score" OR "sharp?design" OR "fuzzy?design" OR "matched?control" OR "synthetic control" OR "regression?discontinuity" OR "inverse?probability weight" OR "mendelian?randomi?ation" OR "controlled before and after" OR "difference?in?difference*" OR "difference?stud*" OR "exogenous varia*" OR "counterfactual" OR "rubin causal model" OR "potential outcome")

**Google Scholar**
(minimum wage OR income OR econ* OR benefit) AND (suicid*)

**Cochrane Central Register Of Controlled Trials (CENTRAL)**
suici* OR self-harm in Title Abstract Keyword AND minimum wage OR income tax?credit OR maternity benefit* OR active lab* OR employment service OR wage subsidy OR vocational training OR job?search service* OR work sharing OR housing support OR housing benefit* OR housing subsidy OR public housing OR welfare OR social policy OR social assistance OR social security OR food stamp OR food assistance OR food aid OR in?kind transfer OR social?transfer OR public works program OR fee waiver* OR family allowance OR child benefit* OR social care OR family support OR childcare OR eldercare OR residential care OR home care OR disability benefit* OR unemployment insurance OR employment insurance OR pension OR sickness benefit* OR income benefit* OR injury compensation OR income supplement OR income support OR income maintenance OR cash?transfer* OR income security OR basic income OR guaranteed income OR austerity OR deaths of despair OR poverty reduction in Title Abstract Keyword AND experiment* OR randomi?ed control* OR RCT OR randomi?ed OR non?randomi?ed OR interrupted time?series OR propensity?score OR sharp?design OR fuzzy?design OR matched?control OR synthetic control OR regression?discontinuity OR

inverse?probability weight OR mendelian?randomi?ation OR controlled before and after OR difference?in?difference* OR difference?stud* OR exogenous varia* OR counterfactual OR rubin causal model OR potential outcome in Title Abstract Keyword

**Proquest Dissertation Dissertation Database**
noft("suici*" OR "self-harm") AND noft("minimum wage" OR "income tax?credit" OR "maternity benefit*" OR "active lab*" OR "employment service" OR "wage subsidy" OR "vocational training" OR "job?search service*" OR "work sharing" OR "housing support" OR "housing benefit*" OR "housing subsidy" OR "public housing" OR "welfare" OR "social policy" OR "social assistance" OR "social security" OR "food stamp" OR "food assistance" OR "food aid" OR "in?kind transfer" OR "social?transfer" OR "public works program" OR "fee waiver*" OR "family allowance" OR "child benefit*" OR "social care" OR "family support" OR "childcare" OR "eldercare" OR "residential care" OR "home care" OR "disability benefit*" OR "unemployment insurance" OR "employment insurance" OR "pension" OR "sickness benefit*" OR "income benefit*" OR "injury compensation" OR "income supplement" OR "income support" OR "income maintenance" OR "cash?transfer*" OR "income security" OR "basic income" OR "guaranteed income" OR "austerity" OR "deaths of despair" OR "poverty reduction") AND noft("experiment*" OR "randomi?ed control*" OR "RCT" OR "randomi?ed" OR "non?randomi?ed" OR "interrupted time?series" OR "propensity?score" OR "sharp?design" OR "fuzzy?design" OR "matched?control" OR "synthetic control" OR "regression?discontinuity" OR "inverse?probability weight" OR "mendelian?randomi?ation" OR "controlled before and after" OR "difference?in?difference*" OR "difference?stud*" OR "exogenous varia*" OR "counterfactual" OR "rubin causal model" OR "potential outcome")

**Econlit**
( suici* OR self-harm ) AND ( minimum wage OR income tax?credit OR maternity benefit* OR active lab* OR employment service OR wage subsidy OR vocational training OR job?search service* OR work sharing OR housing support OR housing benefit* OR housing subsidy OR public housing OR welfare OR social policy OR social assistance OR social security OR food stamp OR food assistance OR food aid OR in?kind transfer OR social?transfer OR public works program OR fee waiver* OR family allowance OR child benefit* OR social care OR family support OR childcare OR eldercare OR residential care OR home care OR disability benefit* OR unemployment insurance OR employment insurance OR pension OR sickness benefit* OR income benefit* OR injury compensation OR income supplement OR income support OR income maintenance OR cash?transfer* OR income security OR basic income OR guaranteed income OR austerity OR deaths of despair OR poverty reduction ) AND ( experiment* OR randomi?ed control* OR RCT OR randomi?ed OR non?randomi?ed OR interrupted time?series OR propensity?score OR sharp?design OR fuzzy?design OR matched?control OR synthetic control OR regression?discontinuity OR inverse?probability weight OR mendelian?randomi?ation OR controlled before and after OR difference?in?difference* OR difference?stud* OR exogenous varia* OR counterfactual OR rubin causal model OR potential outcome )

**RePEc (Research Papers in Economics)**
(minimum wage OR income OR econ* OR benefit) AND (suicid*)

**Scopus (Elsevier)**
( TITLE-ABS-KEY ( "suici*" OR "self-harm" ) AND TITLE-ABS-KEY ( "minimum wage" OR "income tax?credit" OR "maternity benefit*" OR "active lab*" OR "employment service" OR "wage subsidy" OR "vocational training" OR "job?search service*" OR "work sharing" OR "housing support" OR "housing benefit*" OR "housing subsidy" OR "public housing" OR

“welfare” OR “social policy” OR “social assistance” OR “social security” OR “food stamp” OR “food assistance” OR “food aid” OR “in?kind transfer” OR “social?transfer” OR “public works program” OR “fee waiver*” OR “family allowance” OR “child benefit*” OR “social care” OR “family support” OR “childcare” OR “eldercare” OR “residential care” OR “home care” OR “disability benefit*” OR “unemployment insurance” OR “employment insurance” OR “pension” OR “sickness benefit*” OR “income benefit*” OR “injury compensation” OR “income supplement” OR “income support” OR “income maintenance” OR “cash?transfer*” OR “income security” OR “basic income” OR “guaranteed income” OR “austerity” OR “deaths of despair” OR “poverty reduction” ) AND TITLE-ABS-KEY ( “experiment*” OR “randomi?ed control*” OR “RCT” OR “randomi?ed” OR “non?randomi?ed” OR “interrupted time?series” OR “propensity?score” OR “sharp?design” OR “fuzzy?design” OR “matched?control” OR “synthetic control” OR “regression?discontinuity” OR “inverse?probability weight” OR “mendelian?randomi?ation” OR “controlled before and after” OR “difference?in?difference*” OR “difference?stud*” OR “exogenous varia*” OR “counterfactual” OR “rubin causal model” OR “potential outcome” ) )

**Supplementary File 2**

## Data Extraction Form

**Inclusion/Exclusion form:**

| Reference details | |
|---|---|
| Title of paper | |
| Journal | |
| Year of publication | |
| Authors | |
| Publication type | |
| Assessor's name | |
| Date | |

Study included in the review:

| Yes | | No | |
|---|---|---|---|

| If excluded, reason(s) for exclusion | |
|---|---|
| Other types of suicidal behaviour (e.g. ideation, attempt, etc.) | |
| Examining macroeconomic change (e.g. recession, COVID restrictions) | |
| Non-interventional study (e.g. no pre-defined control groups) | |

**Data extraction form:**

| Reference details | |
|---|---|
| Title of paper | |
| Journal | |
| Year of publication | |
| Authors | |
| Publication type | |
| Assessor's name | |
| Date | |

| **Study details** | |
|---|---|
| Start date | |
| End date | |
| Aim of study | |
| Study design | |
| Ethical approval needed/obtained for study | |
| Setting | |
| Population description | |
| Age | |
| Sex | |
| Race/ethnicity | |
| Inclusion criteria | |
| Exclusion criteria | |
| Method of recruitment | |
| Total # randomized/total pop at start | |
| Baseline imbalances | |
| Intervention(s) | |
| Theoretical basis of intervention | |
| Outcome(s) | |
| Quality of vital statistics | |
| Imputation of missing data | |
| Assumed risk estimate | |
| Study findings | |
| Data analysis | |

| Notes | |
|---|---|

# Supplementary File 3

## The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool

(version for cohort-type studies)

Developed by: Jonathan AC Sterne, Miguel A Hernán, Barnaby C Reeves, Jelena Savović, Nancy D Berkman, Meera Viswanathan, David Henry, Douglas G Altman, Mohammed T Ansari, Isabelle Boutron, James Carpenter, An-Wen Chan, Rachel Churchill, Asbjørn Hróbjartsson, Jamie Kirkham, Peter Jüni, Yoon Loke, Terri Pigott, Craig Ramsay, Deborah Regidor, Hannah Rothstein, Lakhbir Sandhu, Pasqualina Santaguida, Holger J Schünemann, Beverly Shea, Ian Shrier, Peter Tugwell, Lucy Turner, Jeffrey C Valentine, Hugh Waddington, Elizabeth Waters, Penny Whiting and Julian PT Higgins

**Version 1 August 2016**



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

### Specify the review question

| | |
|---|---|
| Participants | |
| Experimental intervention | |
| Comparator | |
| Outcomes | |

### List the confounding domains relevant to all or most studies

| |
|---|
| |

### List co-interventions that could be different between intervention groups and that could impact on outcomes

| |
|---|
| |

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

### Specify a target randomized trial specific to the study

| | |
|---|---|
| Design | Individually randomized / Cluster randomized / Matched (e.g. cross-over) |
| Participants | |
| Experimental intervention | |
| Comparator | |

### Is your aim for this study…?

☐ to assess the effect of *assignment to* intervention

☐ to assess the effect of *starting and adhering to* intervention

### Specify the outcome

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

| |
|---|
| |

### Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

| |
|---|
| |



Kim C, *et al. BMJ Open* 2022; 12:e054677. doi: 10.1136/bmjopen-2021-054677

## Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

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

| **(i) Confounding domains listed in the review protocol** | | | | |
|---|---|---|---|---|
| Confounding domain | Measured variable(s) | Is there evidence that controlling for this variable was unnecessary?* | Is the confounding domain measured validly and reliably by this variable (or these variables)? | OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator? |
| | | | Yes / No / No information | Favour experimental / Favour comparator / No information |
| | | | | |
| | | | | |
| | | | | |

| **(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important** | | | | |
|---|---|---|---|---|
| Confounding domain | Measured variable(s) | Is there evidence that controlling for this variable was unnecessary?* | Is the confounding domain measured validly and reliably by this variable (or these variables)? | OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator? |
| | | | Yes / No / No information | Favour experimental / Favour comparator / No information |
| | | | | |
| | | | | |
| | | | | |

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

## Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

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

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

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

## Risk of bias assessment (cohort-type studies)

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a 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 the effect of intervention in this study?<br>**If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered | In rare situations, such as when studying harms that are very unlikely to be related to factors that influence treatment decisions, no confounding is expected and the study can be considered to be at low risk of bias due to confounding, equivalent to a fully randomized trial. There is no NI (No information) option for this signalling question. | Y / PY / PN / N |
| | **If Y/PY to 1.1**: determine whether there is a need to assess time-varying confounding: | | |
| | 1.2. Was the analysis based on splitting participants' follow up time according to intervention received?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, proceed to question 1.3. | If participants could switch between intervention groups then associations between intervention and 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 that are prognostic for the outcome?<br>**If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6)<br>**If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) | If intervention switches are unrelated to the outcome, for example when the outcome is an unexpected harm, then time-varying confounding will not be present and only control for baseline confounding is required. | NA / Y / PY / PN / N / NI |
| | **Questions relating to baseline confounding only** | | |
| | 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? | Appropriate methods to control for measured confounders include stratification, regression, matching, standardization, and inverse probability weighting. They may control for individual variables or for the estimated propensity score. Inverse probability weighting is based on a function of the propensity score. Each method depends on the assumption that there is no unmeasured or residual confounding. | NA / Y / PY / PN / N / NI |

| | | | |
|---|---|---|---|
| | 1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? | Appropriate control of confounding requires that the variables adjusted for are valid and reliable measures of the confounding domains. For some topics, a list of valid and reliable measures of confounding domains will be specified in the review protocol but for others such a list may not be available. Study authors may cite references to support the use of a particular measure. If authors control for confounding variables with no indication of their validity or reliability pay attention to the subjectivity of the measure. Subjective measures (e.g. based on self-report) may have lower validity and reliability than objective measures such as lab findings. | NA / Y / PY / PN / N / NI |
| | 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? | Controlling for post-intervention variables that are affected by intervention is not appropriate. Controlling for mediating variables estimates the direct effect of intervention and may introduce bias. Controlling for common effects of intervention and outcome introduces bias. | NA / Y / PY / PN / N / NI |
| | **Questions relating to baseline and time-varying confounding** | | |
| | 1.7. Did the authors use an appropriate analysis method that adjusted for all the important confounding domains and for time-varying confounding? | Adjustment for time-varying confounding is necessary to estimate the effect of starting and adhering to intervention, in both randomized trials and NRSI. Appropriate methods include those based on inverse probability weighting. Standard regression models that 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 that were adjusted for measured validly and reliably by the variables available in this study? | See 1.5 above. | NA / Y / PY / PN / N / NI |
| | **Risk of bias judgement** | See Table 1. | Low / Moderate / Serious / Critical / NI |
| | Optional: What is the predicted direction of bias due to confounding? | Can the true effect estimate be predicted to be greater or less than the estimated effect in the study because one or more of the important confounding domains was not controlled for? Answering this question will be based on expert knowledge and results in other studies and therefore can only be completed after all of the studies in the body of evidence have been reviewed. Consider the potential effect of each of the unmeasured domains and whether all important confounding domains not controlled for in the analysis would be likely to change the estimate in the same direction, or if one important confounding domain that was not controlled for in the analysis is likely to have a dominant impact. | Favours experimental / Favours comparator / Unpredictable |

| | | | |
|---|---|---|---|
| Bias in selection of participants into the study | 2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?<br>**If N/PN to 2.1:** go to 2.4 | This domain is concerned only with selection into the study based on participant characteristics observed *after* the start of intervention. Selection based on characteristics observed *before* the start of intervention can be addressed by controlling for imbalances between experimental intervention and comparator groups in baseline characteristics that are prognostic for the outcome (baseline confounding). | Y / PY / PN / N / NI |
| | 2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention? | Selection bias occurs when selection is related to an effect of either intervention or a cause of intervention **and** an effect of either the outcome or a cause of the outcome. Therefore, the result is at risk of selection bias if selection into the study is related to both the intervention and the outcome. | NA / Y / PY / PN / N / NI |
| | 2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? | | NA / Y / PY / PN / N / NI |
| | 2.4. Do start of follow-up and start of intervention coincide for most participants? | If participants are not followed from the start of the intervention then a period of follow up has been excluded, and individuals who experienced the outcome soon after intervention will be missing from analyses. This problem may occur when prevalent, rather than new (incident), users of the intervention are included in analyses. | Y / PY / PN / N / NI |
| | 2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases? | It is in principle possible to correct for selection biases, for example by using inverse probability weights to create a pseudo-population in which the selection bias has been removed, or by modelling the distributions of the missing participants or follow up times and outcome events and including them using missing data methodology. However such methods are rarely used and the answer to this question will usually be "No". | NA / Y / PY / PN / N / NI |
| | **Risk of bias judgement** | See Table 1. | Low / Moderate / Serious / Critical / NI |
| | Optional: What is the predicted direction of bias due to selection of participants into the study? | If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions. | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| | | | |
|---|---|---|---|
| Bias in classification of interventions | 3.1 Were intervention groups clearly defined? | A pre-requisite for an appropriate comparison of interventions is that the interventions are well defined. Ambiguity in the definition may lead to bias in the classification of participants. For individual-level interventions, criteria for considering individuals to have received each intervention should be clear and 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), the question relates to whether the population is clearly defined, and the answer is likely to be 'Yes'. | Y / PY / PN / N / NI |
| | 3.2 Was the information used to define intervention groups recorded at the start of the intervention? | In general, if information about interventions received is available from sources that could not have been affected by subsequent outcomes, then differential misclassification of intervention status is unlikely. Collection of the information at the time of the intervention makes it easier to avoid such misclassification. For population-level interventions (e.g. measures to control air pollution), the 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 the outcome or risk of the outcome? | Collection of the information at the time of the intervention may not be sufficient to avoid bias. The way in which the data are collected for the purposes of the NRSI should also avoid misclassification. | Y / PY / PN / N / NI |
| | **Risk of bias judgement** | See Table 1. | Low / Moderate / Serious / Critical / NI |
| | Optional: What is the predicted direction of bias due to measurement of outcomes or interventions? | If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions. | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| | | | |
|---|---|---|---|
| Bias due to deviations from intended interventions | **If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2** | | |
| | 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? | Deviations that happen in usual practice following the intervention (for example, cessation of a drug intervention because of acute toxicity) are part of the intended intervention and therefore do not lead to bias in the effect of assignment to intervention.<br><br>Deviations may arise due to expectations of a difference between intervention and comparator (for example because participants feel unlucky to have been assigned to the comparator group and therefore seek the 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 *and* likely to have affected the outcome? | Deviations from intended interventions that do not reflect usual practice will be important if they affect the outcome, but not otherwise. Furthermore, bias will arise only if there is imbalance in the deviations across the two groups. | NA / Y / PY / PN / N / NI |
| | **If your aim for this study is to assess the effect of starting and 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 a way that would bias the estimated effect of intervention. Co-interventions will be important if they affect the outcome, but not otherwise. Bias will arise only if there is imbalance in such co-interventions between the intervention groups. Consider the co-interventions, including any pre-specified co-interventions, that are likely to affect the outcome and 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 the intervention implemented successfully for most participants? | Risk of bias will be higher if the intervention was not implemented as intended by, for example, the health care professionals delivering care during the trial. Consider whether implementation of the 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 the intervention as intended. Lack of adherence includes imperfect compliance, cessation of intervention, crossovers to the comparator intervention and switches to another active intervention. Consider available information on the proportion of study participants who continued with their assigned | Y / PY / PN / N / NI |

| | | | |
|---|---|---|---|
| | | intervention throughout follow up, and answer 'No' or 'Probably No' if this proportion is high enough to raise concerns. Answer 'Yes' for studies of interventions that are administered once, so that imperfect adherence is not possible.<br>We distinguish between analyses where follow-up time after interventions switches (including cessation of intervention) is assigned to (1) the new intervention or (2) the original intervention. (1) is addressed under time-varying confounding, and should not be considered further here. | |
| | 4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? | It is possible to conduct an analysis that corrects for some types of deviation from the intended intervention. Examples of appropriate analysis strategies include inverse probability weighting or instrumental variable estimation. It is possible that a paper reports such an analysis without reporting information on the deviations from intended intervention, but it would be hard to judge such an analysis to be appropriate in the absence of such information. Specialist advice may be needed to assess studies that used these approaches.<br><br>If everyone in one group received a co-intervention, adjustments cannot be made to overcome this. | NA / Y / PY / PN / N / NI |
| | **Risk of bias judgement** | See Table 2 | |
| | Optional: What is the predicted direction of bias due to deviations from the intended interventions? | If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the 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 the findings", and a suitable proportion depends on the context. In some situations, availability of data from 95% (or possibly 90%) of the 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 and locate an analysis plan for the study. | Y / PY / PN / N / NI |
| | 5.2 Were participants excluded due to missing data on intervention status? | Missing intervention status may be a problem. This requires that the *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 the analysis? | This question relates particularly to participants excluded from the analysis because of missing information on confounders that were controlled for in the analysis. | Y / PY / PN / N / NI |
| | 5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and 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 the 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 the presence of missing data? | Evidence for robustness may come from how missing data were handled in the analysis and whether sensitivity analyses were performed by the investigators, or occasionally from additional analyses performed by the systematic reviewers. It is important to assess whether assumptions employed in analyses are clear and plausible. Both content knowledge and 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, and clear differences between complete-case and multiple imputation-based findings should lead to careful assessment of the validity of the methods used. | NA / Y / PY / PN / N / NI |
| | **Risk of bias judgement** | See Table 2 | Low / Moderate / Serious / Critical / NI |
| | Optional: What is the predicted direction of bias due to missing data? | If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions. | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| | | | |
|---|---|---|---|
| Bias in measurement of outcomes | 6.1 Could the outcome measure have been influenced by knowledge of the 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 the intervention received by study participants? | If outcome assessors were blinded to intervention status, the answer to this question would be 'No'. In other situations, outcome assessors may be unaware of the interventions being received by participants despite there being no active blinding by the study investigators; the answer this question would then also be 'No'. In studies where participants report their outcomes themselves, for example in a questionnaire, the outcome assessor is the study participant. In an observational study, the answer to this question will usually be 'Yes' when the participants report their outcomes themselves. | Y / PY / PN / N / NI |
| | 6.3 Were the methods of outcome assessment comparable across intervention groups? | Comparable assessment methods (i.e. data collection) would involve the same outcome detection methods and thresholds, same time point, same definition, and same measurements. | Y / PY / PN / N / NI |
| | 6.4 Were any systematic errors in measurement of the outcome related to intervention received? | This question refers to differential misclassification of outcomes. Systematic errors in measuring the outcome, if present, could cause bias if they are related to intervention or to a confounder of the intervention-outcome relationship. This will usually be due either to outcome assessors being aware of the 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 2 | Low / Moderate / Serious / Critical / NI |
| | Optional: What is the predicted direction of bias due to measurement of outcomes? | If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions. | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| | | | |
|---|---|---|---|
| Bias in selection of the reported result | Is the reported effect estimate likely to be selected, on the basis of the results, from...<br>7.1. ... multiple outcome *measurements* within the outcome domain? | For a specified outcome domain, it is possible to generate multiple effect estimates for different measurements. If multiple measurements were made, but only one or a subset is reported, there is a risk of selective reporting on the basis of results. | Y / PY / PN / N / NI |
| | 7.2 ... multiple *analyses* of the intervention-outcome relationship? | Because of the limitations of using data from non-randomized 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 and adjusted models; use of final value vs change from baseline vs analysis of covariance; different transformations of variables; a continuously scaled outcome converted to categorical data with different cut-points; different sets of covariates used for adjustment; and different analytic strategies for dealing with missing data. Application of such methods generates multiple estimates of the effect of the intervention versus the comparator on the outcome. If the analyst does not pre-specify the methods to be applied, and multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the 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 the original cohort. If multiple estimates are generated but only one or a subset is reported, there is a risk of selective reporting on the basis of results. | Y / PY / PN / N / NI |
| | **Risk of bias judgement** | See Table 2 | Low / Moderate / Serious / Critical / NI |
| | Optional: What is the predicted direction of bias due to selection of the reported result? | If the likely direction of bias can be predicted, it is helpful to state this. The direction might be characterized either as being towards (or away from) the null, or as being in favour of one of the interventions. | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| | | | |
|---|---|---|---|
| Overall bias | **Risk of bias judgement** | See Table 3. | Low / Moderate / Serious / Critical / NI |
| | Optional:<br>What is the overall predicted direction of bias for this outcome? | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

Table 1. Reaching risk of bias judgements in ROBINS-I: pre-intervention and at-intervention domains

| **Judgement** | **Bias due to confounding** | **Bias in selection of participants into the study** | **Bias in classification of interventions** |
|---|---|---|---|
| Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain) | No confounding expected. | (i) All participants who would have been eligible for the target trial were included in the study;<br>*and*<br>(ii) For each participant, start of follow up and start of intervention coincided. | (i) Intervention status is well defined;<br>*and*<br>(ii) Intervention definition is based solely on information collected at the time of intervention. |
| Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial): | (i) Confounding expected, all known important confounding domains appropriately measured and controlled for;<br>*and*<br>(ii) Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding. | (i) Selection into the study may have been related to intervention and outcome;<br>*and*<br>The authors used appropriate methods to adjust for the selection bias;<br>*or*<br>(ii) Start of follow up and start of intervention do not coincide for all participants;<br>*and*<br>(a) the proportion of participants for which this was the case was too low to induce important bias;<br>*or*<br>(b) the authors used appropriate methods to adjust for the selection bias;<br>*or*<br>(c) the review authors are confident that the rate (hazard) ratio for the effect of intervention remains constant over time. | (i) Intervention status is well defined;<br>*and*<br>(ii) Some aspects of the assignments of intervention status were determined retrospectively. |

| | | | |
|---|---|---|---|
| Serious risk of bias (the study has some important problems); | (i) At least one known important domain was not appropriately measured, or not controlled for;<br>*or*<br>(ii) Reliability or validity of measurement of an important domain was low enough that we expect serious residual confounding. | (i) Selection into the study was related (but not very strongly) to intervention and outcome;<br>*and*<br>This could not be adjusted for in analyses;<br>*or*<br>(ii) Start of follow up and start of intervention do not coincide;<br>*and*<br>A potentially important amount of follow-up time is missing from analyses;<br>*and*<br>The rate ratio is not constant over time. | (i) Intervention status is not well defined;<br>*or*<br>(ii) Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. |
| Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention); | (i) Confounding inherently not controllable<br>*or*<br>(ii) The use of negative controls strongly suggests unmeasured confounding. | (i) Selection into the study was very strongly related to intervention and outcome;<br>*and*<br>This could not be adjusted for in analyses;<br>*or*<br>(ii) A substantial amount of follow-up time is likely to be missing from analyses;<br>*and*<br>The rate ratio is not constant over time. | (Unusual) An 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 information is reported about selection of participants into the study or whether start of follow up and start of intervention coincide. | No definition of the intervention or no explanation of the source of information about intervention status is reported. |

Table 2. 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 the reported result** |
|---|---|---|---|---|
| Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain) | **Effect of assignment to intervention:**<br>(i) Any deviations from intended intervention reflected usual practice;<br>*or*<br>(ii) Any deviations from usual practice were unlikely to impact on the outcome.<br><br>**Effect of starting and adhering to intervention:**<br>The important **co-interventions** were balanced across intervention groups, and there were no deviations from the intended interventions (in terms of **implementation or adherence**) that were likely to impact on the outcome. | (i) Data were reasonably complete;<br>*or*<br>(ii) Proportions of and reasons for missing participants were similar across intervention groups;<br>*or*<br>(iii) The analysis addressed missing data and is likely to have removed any risk of bias. | (i) The methods of outcome assessment were comparable across intervention groups;<br>*and*<br>(ii) The outcome measure was unlikely to be influenced by knowledge of the intervention received by study participants (i.e. is objective) or the outcome assessors were unaware of the intervention received by study participants;<br>*and*<br>(iii) Any error in measuring the outcome is unrelated to intervention status. | There is clear evidence (usually through examination of a pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses and sub-cohorts. |

| | | | | |
|---|---|---|---|---|
| Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial): | **Effect of assignment to intervention:**<br>There were deviations from usual practice, but their impact on the outcome is expected to be slight.<br><br>**Effect of starting and adhering to intervention:**<br>(i) There were deviations from intended intervention, but their impact on the outcome is expected to be slight.<br>*or*<br>(ii) The important co-interventions were not balanced across intervention groups, or there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome;<br>*and*<br>The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome. | (i) Proportions of and reasons for missing participants differ slightly across intervention groups;<br>*and*<br>(ii) The analysis is unlikely to have removed the risk of bias arising from the missing data. | (i) The methods of outcome assessment were comparable across intervention groups;<br>*and*<br>(ii) The outcome measure is only minimally influenced by knowledge of the intervention received by study participants;<br>*and*<br>(iii) Any error in measuring the outcome is only minimally related to intervention status. | (i) The outcome measurements and analyses are consistent with an *a priori* plan; or are clearly defined and both internally and externally consistent;<br>*and*<br>(ii) There is no indication of selection of the reported analysis from among multiple analyses;<br>*and*<br>(iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. |

| | | | | |
|---|---|---|---|---|
| Serious risk of bias (the study has some important problems); | **Effect of assignment to intervention:**<br>There were deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome.<br><br>**Effect of starting and adhering to intervention:**<br>(i) The important co-interventions were not balanced across intervention groups, or there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome;<br>*and*<br>(ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome. | (i) Proportions of missing participants differ substantially across interventions;<br>*or*<br>Reasons for missingness differ substantially across interventions;<br>*and*<br>(ii) The analysis is unlikely to have removed the risk of bias arising from the missing data;<br>*or*<br>Missing data were addressed inappropriately in the analysis;<br>*or*<br>The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis. | (i) The methods of outcome assessment were not comparable across intervention groups;<br>*or*<br>(ii) The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);<br>*and*<br>The outcome was assessed by assessors aware of the intervention received by study participants;<br>*or*<br>(iii) Error in measuring the outcome was related to intervention status. | (i) Outcomes are defined in different ways in the methods and results sections, or in different publications of the study;<br>*or*<br>(ii) There is a high risk of selective reporting from among multiple analyses;<br>*or*<br>(iii) The cohort or subgroup is selected from a larger study for analysis and appears to be reported on the basis of the results. |

| | | | | |
|---|---|---|---|---|
| Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention); | **Effect of assignment to intervention:**<br>There were substantial deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome.<br><br>**Effect of starting and adhering to intervention:**<br>(i) There were substantial imbalances in important co-interventions across intervention groups, or there were substantial deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome;<br>*and*<br>(ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome. | (i) (Unusual) There were critical differences between interventions in participants with missing data;<br>*and*<br>(ii) Missing data were not, or could not, be addressed through appropriate analysis. | The 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;<br>*and*<br>(ii) The unreported results are likely to be substantially different from the reported results. |

| | | | | |
|---|---|---|---|---|
| No information on which to base a judgement about risk of bias for this domain. | No information is reported on whether there is deviation from the intended intervention. | No information is reported about missing data or the potential for data to be missing. | No information is reported about the methods of outcome assessment. | There is too little information to make a judgement (for example, if only an abstract is available for the study). |

Table 3. Interpretation of domain-level and overall risk of bias judgements in ROBINS-I

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