

Supplementary file 5. ROBINS-I assessments for studies providing sufficient data to calculate the difference in change between quitters and continuing smokers from baseline to longest follow-up: Mixed anxiety and depression

Study	Bias due to confounding	Bias in selection of participants into the study (pre-filled)	Bias in classification of interventions (Bias in classification of quitting smoking)	Bias due to deviations from the intended intervention (Bias due to deviations from quitting/continuing smoking (i.e., relapsing) or access to psychoactive treatments)	Bias due to missing data	Bias in measurement of outcomes (pre-filled)	Bias in selection of the reported result	Overall risk of bias
Blalock 2008	Serious	Low	Low	No information	No information	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: Smoking cessation was clearly defined. Continuous abstinence - defined as self-reported sustained abstinence beginning 2 weeks after quit date, through the 3-month follow-up assessment.	Justification: Not enough information was provided to answer the signalling questions in this domain.	Justification: The final sample included 72% (21/29) participants who entered the trial. Not enough information was provided to answer the remaining signalling questions in this domain.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis. There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure group.	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple analyses of the association in question, or different subgroups.	Justification: One domain judged as being at serious risk of bias.
Carroll 2019	Serious	Low	Low	Low	Low	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: Smoking cessation was clearly defined as 24-hour point-prevalence abstinence (i.e. not a cigarette, even a puff in past 24 hours).	Justification: Psychoactive treatments were not offered as part of the study. Participants who are likely to need psychoactive treatment were excluded from the	Justification: Retention was 84.1% at 12 weeks.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis.	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple analyses of the	Justification: One domain judged as being at serious risk of bias.

				study "Exclusion criteria included use of other tobacco products or smoking cessation treatments, current diagnosis of substance abuse disorder, having a current medical problem for which varenicline use is contraindicated (e.g., allergy), a lifetime DSM-IV diagnosis of psychotic or bipolar disorder or current unstable or untreated major depression, current suicidality or a past attempt as identified by the Mini International Neuropsychiatric Interview".		There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.	association in question, or different subgroups.	
Chassin 2002	Serious	Low	Low	No information	Moderate	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: Smoking cessation was clearly defined.	Justification: Not enough information was provided to answer all signalling questions in this domain. No psychoactive treatments were offered as part of the study.	Justification: Only 57% of the sample provided smoking status data at both waves and were included in the analysis. For our meta-analysis we used unadjusted data, so exclusion based on other variables is not relevant here.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis. There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple analyses of the association in question, or different subgroups.	Justification: Two domains judged as being at serious risk of bias.

Cinciripini 2013	Serious	Low	Low	Moderate	Moderate	Low	Serious	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: Smoking cessation was clearly defined. Prolonged abstinence (primary smoking outcome) defined as 7 or more consecutive days of smoking or smoking at least 1 cigarette over two consecutive weeks from the end of the grace period to a selected future time point (e.g. EOT, 3- and 6-months post-quit date); seven-day point-prevalence abstinence defined as self-report of no smoking (not even a puff), in 7 days prior to time point of interest; continuous abstinence defined as self-report of no smoking (not even a puff) from 2-weeks post quit date to future time point; or beginning with last 4 weeks of treatment or week 8 of medication. Bio-verified by expired CO < 10 ppm; salivary cotinine values < 15 ng/mL.	Justification: Bupropion treatment rates were higher in quitters. No appropriate analysis was used to address this.	Justification: Outcome data were missing for 62.9% (180/286) of participants.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis. There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.	Justification: CESD outcome data is not reported. Because no differences were found by including covariates the results are reported for the unadjusted but not the adjusted models.	Justification: Three domains judged as being at serious risk of bias.
Leventhal 2014	Serious	Low	Low	Serious	Moderate	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: Smoking cessation was clearly defined. Quitters were defined as having carbon monoxide-confirmed 7-day point prevalence abstinence.	Justification: Bupropion provided to one trial arm but was not adjusted for in the analysis.	Justification: Outcome data were available for 66% (999/1504) of participants. Participants who did not provide outcome data were coded as non-abstinent.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis. There is no reason to believe that participants would assess their	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple analyses of the association in question, or different subgroups.	Justification: Two domains judged as being at serious risk of bias.

						mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.		
Mino 2000	Serious	Low	Moderate	No information	No information	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: The definition used was not reported.	Justification: Not enough information was provided to answer all signalling questions in this domain. No psychoactive treatments were offered as part of the study.	Justification: Outcome data were available for 91.6% (175/191) at 6 months and 84.8% (162/191) at 12 months. Definition of smoking is not reported so we cannot tell if participants were excluded based on smoking status. No information about if participants were excluded due to missing data for other variables.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis. There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple analyses of the association in question, or different subgroups.	Justification: One domain judged as being at serious risk of bias.
Robinson 2019	Serious	Low	Low	No information	Serious	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed	Justification: Smoking cessation was clearly defined. 7-day point prevalence abstinence at week 24 and week 52 post-quit, bio-verified by expired CO (≤ 8 ppm).	Justification: Not enough information was provided to answer all signalling questions in this domain. Psychoactive treatment did not differ across	Justification: Outcome data were available for 66% (822/1245) of participants. More smoking participants lost to follow-up (40% versus 14%). Our	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple	Justification: Two domains judged as being at serious risk of bias.

		after the start of quitting smoking.		randomised groups. Some psychoactive treatments were exclusions, but not all.	analysis does not account for missing data. No evidence of any other analyses accounting for missing data.	review/meta-analysis. There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.	analyses of the association in question, or different subgroups.	
Steinberg 2011	Serious	Low	Low	No information	Low	Low	Low	Serious
	Justification: Authors did not use an appropriate analysis method that controlled for all the important time-varying confounding.	Justification: Selection of participants into the study (or into the analysis) was not based on participant characteristics observed after the start of quitting smoking.	Justification: Smoking cessation was clearly defined as self-reported 7-day point-prevalence abstinence bio-verified by exhaled CO (< 10 ppm).	Justification: Not enough information was provided to answer all signalling questions in this domain. Patients may have received psychoactive medications based on clinical criteria and were not randomly allocated to treatment. Therefore, some of the observed effects could have been due to differences in medication treatment groups that were not controlled for.	Justification: Outcome data were available for > 70% of participants. Those lost to follow-up were classified as smokers.	Justification: Participants would need to remember what they scored at baseline, and participants will be blinded to the hypothesis of our review/meta-analysis. There is no reason to believe that participants would assess their mental health differently due to quitting status. The same questionnaires would have been administered at the same time and in the same manner for all participants. It is unlikely that measurement error in mental health outcomes would be split by exposure.	Justification: No evidence that the effect estimate was likely to be selected on the basis of results from multiple outcome measurements, multiple analyses of the association in question, or different subgroups.	Justification: One domain judged as being at serious risk of bias.