Annex 3. Incorporating bias assessments into evidence synthesis

ANNEX 3.

Incorporating bias assessments into evidence synthesis

Alexander P. Keil

A3.1 Multiple-bias analysis: worked examples

The goal of bias analysis is to estimate the expected effect that would have been estimated in the study, had that study not been subject to the bias of concern. For the purposes of this book, estimates of the risk ratio are considered, contrasting the cumulative risk of an outcome at two levels of exposure. The risk ratio obtained in a given study is referred to as the apparent risk ratio, RR_{app}, and the risk ratio after performing bias analysis is referred to as the adjusted risk ratio, RR_{adj}. The true target of RR_{adj} is $RR_{unbiased}$, the risk ratio that would have been estimated in the absence of any systematic bias, but RR_{adj} is used to emphasize the necessary simplifying assumptions that feed into a bias analysis and the reality that the only bias parameters that may be available are typically, at best, approximations to the true bias parameters.

The following worked example offers a template for adapting multiple-bias analysis to new scenarios, but it also indicates a unique aspect of multiple-bias analysis that sets it apart from single-bias analyses: the approach to serial multiple-bias analysis ought to vary according to the order in which biases are thought to occur for the RR_{app} under consideration. Smith et al. (2021, p. 627) write, "In general, we can think of biases as layers that we must peel off sequentially and the order in which we do so is the reverse of the order in which they occurred in the data." Fox et al. (2021, pp. 416–417) state, "Bias-adjustment does not generally reduce to independent multiplicative bias factors […], so the order of bias-adjustments can affect the ultimate result."

Two primary approaches to bias analyses could be considered: (i) the approach of Smith et al. (2021), which uses (dependent) bias factors to estimate upper or lower bounds of bias for a range of bias parameters, and (ii) an approach given in Fox et al. (2021) that involves the calculation of pseudo-data. In the first approach, a bound of the value of RR_{adj} is established that is typically a direct answer to the question "What is the most extreme value of the true risk ratio that is still consistent with RR_{app} under the bias parameters?" This approach answers the useful question (for hazard identification) "Can we rule out bias as the sole explanation of a non-null effect estimate?" However, the approach described by Smith et al. (2021) has not been studied extensively, so it is not known how conservative the bound is (e.g. how likely it is that a bound will indicate consistency with

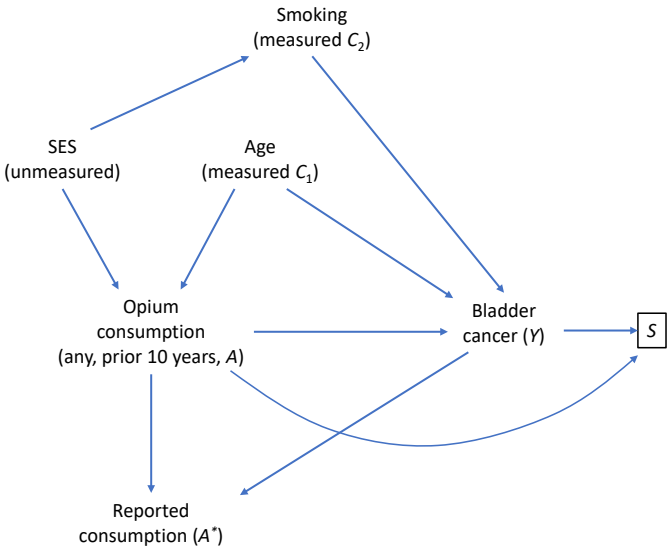
the null hypothesis). Furthermore, this bounding approach does not easily accommodate scenarios in which biases may offset each other. The second approach establishes a value for RR_{adj} that is derived by stacking the approaches used in other chapters of this book for individual biases. Rather than a bound, the second approach provides a single best estimate under a given set of bias parameters. This approach can be extended in a probabilistic bias analysis to accommodate uncertainty in the bias parameters.

A3.1.1 Worked example with a single study on opium consumption and bladder cancer

The primary examples are based on a study on opium consumption and bladder cancer (Aliramaji et al., 2015). Fig. A3.1 expresses three potential biases that might be considered. Specifically, these are related to issues of exposure misclassification, unmeasured confounding, and selection bias, the last of which arises from the method of selecting the study population. In the analysis of Aliramaji et al. (2015), which is a case–control study, odds ratios, rather than risk ratios, are reported. The bias correction methods used here rely variously on odds ratios and risk ratios, but note that, given a rare disease like bladder cancer, these can be considered nearly equivalent so that methods to adjust a risk ratio can be used with an odds ratio.

Aliramaji et al. (2015) do not report measures of association, but the apparent odds ratio is calculable as the crude odds ratio (Table A3.1). Note that the crude odds ratio of 2.72 is different from the odds ratio given in Table 6.4, which was calculated as the crude odds ratio among tobacco non-smokers. The odds of having bladder cancer (case odds) are 2.15 among those who reported opium use and 0.79 among those who did not, leading to an apparent odds ratio of 2.72. As noted, for this odds ratio there is a concern over unmeasured confounding (because an adjusted estimate of association for key confounders of age and smoking was not reported). Selection bias concerns arise because the control participants were selected from among hospitalized patients who were being treated surgically for gall bladder stones; in addition, as noted in Fig. A3.2, selection bias may arise because there was frequency matching on sex (Mansournia et al., 2018). Opium exposure misclassification concerns arise because these were identified in prior validation studies on self-reported opium consumption. This bias analysis is focused on the scenario in Fig. A3.2, in which selection bias due to matching is of greater concern than Berkson bias (which was discussed in Chapter 5), because of the recruitment of hospitalized control participants.

Fig. A3.1. Simplified directed acyclic graph (DAG) showing potential areas where bias correction may be used in an analysis of a study on opium consumption and bladder cancer (Aliramaji et al., 2015). This DAG illustrates three biases: (i) differential exposure measurement error: measured exposure (A^*) is a mismeasured proxy of true exposure (A), in which measurement error depends on the outcome (Y); (ii) selection bias: the recruitment of hospitalized control participants raises concerns that selection in the study may be affected by opium use, because opium use can cause other hospitalizable outcomes; and (iii) confounding by age and smoking (C_1, C_2). SES, socioeconomic status.



A3.1.2 Order of bias corrections

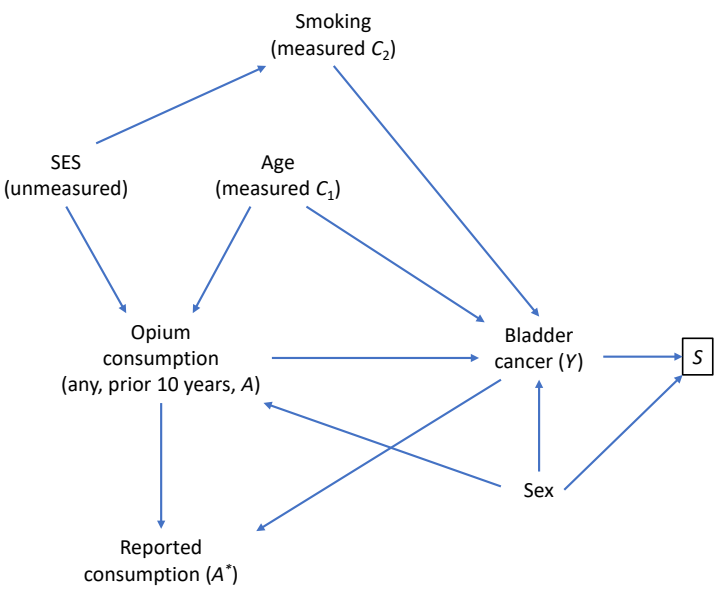
This example, drawn from Aliramaji et al. (2015), is an interesting case study in multiple-bias analysis because it demonstrates issues of measured

Table A3.1. Raw exposure and case status data, and calculated crude odds ratio

		Case participants	Control participants	Case odds
Opium use	Yes	58	27	2.15
	No	117	148	0.79
Crude/apparent odds ratio				2.72

Source: Aliramaji et al. (2015).

Fig. A3.2. Simplified directed acyclic graph (DAG) showing alternative specification in which bias correction may be used in an analysis of a study on opium consumption and bladder cancer (Aliramaji et al., 2015). This DAG illustrates three biases: (i) differential exposure measurement error: measured exposure (A^*) is a mismeasured proxy of true exposure (A), in which measurement error depends on the outcome (Y); (ii) selection bias: hospitalized control participants were being treated surgically for gall bladder stones, for which there is no known association with bladder cancer; however, control participants were matched with case participants on sex; this latter factor was left uncontrolled in the analysis, leading to selection bias; (iii) confounding by age and smoking (C_1, C_2). SES, socioeconomic status.



confounding, effect measure modification, and selection bias arising from matching on confounders. The order of bias correction is guided by the order in which biases may appear in the data. In this study, frequency matching of the study design on confounding factors (age and sex) can introduce selection bias in analyses that are unadjusted for these factors. Thus, selection bias occurs because of an open path from the outcome to the (correctly classified) exposure through the selection node and the frequency-matched factors. Adjusting for this bias requires that pathway to be closed; this can be done by adjusting for the matched factors. Selection bias can thus be considered as the first bias to address, given that true exposure need not be measured. Had exposure directly influenced selection (as might occur if opium were a cause of a condition that resulted in a person's being selected as a control participant), misclassification bias would necessarily first have been considered for adjustment. Once study selection is adjusted for, exposure misclassification can be adjusted for. Finally, consideration is given to confounding, which is considered to be a function of reality, rather than study design or measurement issues. This bias would therefore be considered to happen first (in temporal order), and correction for it would come last. Smith et al. (2021) and Fox et al. (2021) both consider alternative orderings in multiple-bias analysis.

A notable issue when selecting the order of bias correction for this example is in regard to the available data. Fox et al. (2021, p. 417) perfectly encapsulate this scenario: "Classification parameters might be measured in a population-based setting (i.e. negligible selection bias), but be applied to a data set where selection bias is a concern. In this setting, the analyst should bias-adjust for selection bias before bias-adjusting for misclassification, even if the selection bias preceded the misclassification in the data generation process." Thus, even if true exposure plays a role in selection bias, the role of opium in inducing selection bias is less pertinent to the biases at issue than the role of misclassification in effect estimation. Thus, there is an additional reason in this example to apply misclassification parameters in data that are already adjusted for selection bias, rather than the other way around.

A3.2 Overview of multiple-bias analysis of the data of Aliramaji et al. (2015) using pseudo-data and bias-factor approaches

The pseudo-data approach (using bias parameters to calculate bias-adjusted data, which can then be analysed as though they were real data) is used for plausible (under reasonable assumptions) ranges of bias arising from exposure misclassification and selection bias. The pseudo-data approach begins with an apparent odds ratio that is unadjusted for confounders. Consideration is also given to measured confounding, similar to the selection bias adjustment example of Chapter 4. The bias-factor approach, as described by Smith et al. (2021), is used to place a lower bound on the multiple-bias-adjusted odds ratio, given a set of bias parameters. The bias-factor approach begins with an apparent odds ratio that is adjusted for confounders. The parameters needed in the bias-factor approach are summarized in Table A3.2. Both approaches are demonstrated in the R code provided in Annex 2 (online only; available from: https://publications.iarc.who.int/634#supmat).

A3.2.1 The apparent odds ratio as the basis for subsequent calculations

The crude odds ratio of Table A3.1 can be used as the apparent odds ratio for subsequent analysis in the pseudo-data approach. Alternatively, Aliramaji et al. (2015) give enough information to infer smoking-stratified results (in the results, sample sizes were reported for case and control participants who had both consumed opium and smoked cigarettes for longer than 1 year); these figures are given in Table A3.3 and yield a smoking-adjusted odds ratio of 1.25 from unconditional logistic regression and stratum-specific odds ratios of 4.1 among non-smokers and 0.5 among smokers. The stratum-specific results indicate substantial odds ratio modification; this is a key consideration, as discussed for unmeasured confounding in Section 3.3. This adjusted odds ratio can, nonetheless, be selected as the apparent odds ratio for the bias-factor approach because it does not rely on tabulated data.

A3.2.2 Selection bias adjustment using bias factors

As shown in Table A3.2, adjustment for selection bias using the bias-factor approach involves consideration of a factor, U_S, that influences selection into the study and is also conditionally associated with the outcome. In this analysis, U_S is considered as sex only, given that opium use varies strongly with sex, and sex was used as a matching factor for the study but was not subsequently adjusted for. Matching without adjustment for sex created a backdoor biasing pathway, because of conditioning on the collider S (opium consumption ← sex → S ← bladder cancer). For a causal interpretation of the selection-bias adjustment, adjustment for U_S should be sufficient to render the study outcome and study selection (the node S in Figs. A3.1 and A3.2) independent, given other factors that are included in the analysis. Crucially, the selection bias under consideration here affects the meaning of S. By definition, $S = 1$ is the value of S for members in the study population; $S = 0$ is given for individuals who would have been part of the study data, had they not been selected out of the study. Generally, to adjust for selection bias, one must know or assume something about those for whom $S = 0$. Here, those for whom $S = 0$ are a (potentially hypothetical) group of women who were at the hospital used in the study for surgical treatment of gall bladder stones but were not included in the study. To simplify further calculations, it is assumed that selection into this eligible population is not related to sex; thus, a similar sex ratio among potential control participants to that in the underlying source population (i.e. 1:1) is expected.

Calculating the bias factor for this example involves specifying (for binary misclassified exposure A^*) $\Pr(U_S = u \mid A^* = a, S = a, C = c)$, $\Pr(U_S = u \mid A^* = a, S = 1 - a, C = c)$ (the prevalence parameters), and $\Pr(Y = 1 \mid A^* = a, C = c, U_S = u)$ (the risk parameters), which are the prevalence of the unmeasured factor at some level u (i.e. male or female for the binary in this example) and the risk of the outcome at specified values of A^*, C, and U_S.

To inform the prevalence parameters, data were included from a national survey of residents of the Islamic Republic of Iran conducted by Moradinazar et al. (2020), who estimated the prevalence of drug use, stratified by several demographic variables, using the survey question "Have you used illicit drug more than one time during a lifetime?" The average prevalence was estimated as 24.1% among men and 2.2% among women. These survey data correspond to the sex-specific prevalences, $\Pr(A^* = a \mid U_S = u)$; it is assumed that this does not vary meaningfully across levels of covariates C

Table A3.2. Summary of parameters that determine bias factors according to the multiple-bias bounding method of Smith et al. (2021)

Type of bias[a]	Identifier	Parameter	Definition and notes	Investigator-specified values	Calculated value
Unmeasured confounding	D	$\Pr(U_c = u \mid A = 1, C = c)$	Prevalence of $U_c = u$, among exposed, given observed confounders.	✓	
	E	$\Pr(U_c = u \mid A = 0, C = c)$	Prevalence of $U_c = u$, among unexposed, given observed confounders.	✓	
	RR_{AU_C}	$\max_u(D/E)$	The maximum factor (over levels of the unmeasured confounder) by which exposure is conditionally associated with a given value of the unmeasured confounder in an analysis free of selection bias and misclassification bias. For binary U_c, this is the maximum prevalence ratio contrasting levels of exposure, given measured confounders.		✓
	F	$\Pr(Y = 1 \mid A = a, C = c, U_c = u)$	The risk of Y, given exposure, observed confounders, and unmeasured confounders.	✓	
	RR_{U_CY}	$\max_a[\max_u(F)/\min_u(F)]$	The maximum value (across levels of exposure) of the ratio of the maximum risk (across different levels of U_c) and minimum risk (across different levels of U_c). This is the maximum possible risk ratio contrasting the outcome risk across levels of U_c and describes confounding bias above and beyond measured confounding in an analysis free of selection bias and misclassification bias.		✓
	BF_C	$g(RR_{U_CY}, RR_{AU_C})$	Multiplicative bias factor, confounding. This is interpreted as the confounding-bias risk ratio or the ratio of the risk ratio adjustment for U_c to the risk ratio with unmeasured confounding by U_c.		
Differential or non-differential exposure misclassification	G	$\Pr(A^* = 1 \mid Y = 1, A = 0, S = 1, C = c)$	False-positive probability among case participants (1 − specificity)	✓	
	H	$\Pr(A^* = 1 \mid Y = 0, A = 0, S = 1, C = c)$	False-positive probability among non-case participants (1 − specificity)	✓	
	I	$\Pr(A^* = 1 \mid Y = 1, A = 1, S = 1, C = c)$	True-positive probability among case participants (sensitivity)	✓	
	J	$\Pr(A^* = 1 \mid Y = 0, A = 1, S = 1, C = c)$	True-positive probability among non-case participants (sensitivity)	✓	
	FPOR	$(G/H)/[(1 - G)/(1 - H)]$	False-positive odds ratio		✓
	SEOR	$(I/J)/[(1 - I)/(1 - J)]$	Sensitivity odds ratio		✓
	CCR	$(I/J)/[(1 - G)/(1 - H)]$	Correct classification ratio		✓
	ICR	$(G/H)/[(1 - I)/(1 - J)]$	Incorrect classification ratio		✓
	OR_{A^*Y}	max(FPOR, SEOR, CCR, ICR)	Maximum selection odds ratio		✓
	BF_M	OR_{A^*Y}	Multiplicative bias factor, differential exposure misclassification. Note that this is a bias odds ratio and applies when the effect estimate is an odds ratio. In rare disease settings, this approximates the risk ratio and can be used for risk ratios.		✓

Table A3.2. Summary of parameters that determine bias factors according to the multiple-bias bounding method of Smith et al. (2021) (continued)

Type of bias[a]	Identifier	Parameter	Definition and notes	Investigator-specified values	Calculated value
Selection bias	K	$\Pr(Y = 1 \mid A = a, C = c, U_S = u)$	The risk of Y, given exposure, observed confounders, and an unmeasured source of selection bias	✓	
	$RR_{U_SY}(a)$	$\max(K)/\min(K)$	The ratio (at a given level of exposure, and given confounders) of the maximum risk of the outcome (across levels of the unmeasured source of selection bias) and the minimum risk of the outcome (across levels of the unmeasured source of selection bias). This is the maximum possible risk ratio contrasting levels of the variable that is a source of selection bias. If U_S is binary, this is simply the conditional risk ratio contrasting $U_S = 1$ against $U_S = 0$, or its inverse, whichever is larger.		✓
	L	$\Pr(U_S = u \mid A = a, S = a, C = c)$	Prevalence of $U_S = u$ at a given level of exposure, among those who were selected into the study (if exposed) or among those not selected into the study (if unexposed), given observed confounders	✓	
	M	$\Pr(U_S = u \mid A = a, S = 1 - a, C = c)$	Prevalence of $U_S = u$ at a given level of exposure, among those who were selected into the study (if considering the unexposed) or among those not selected into the study (if considering the exposed), given observed confounders	✓	
	$RR_{SU_S}(a)$	$\max(L/M)$	The maximum ratio by which selection into the study increases the prevalence of some value of U_S, within strata of exposure, given confounders. For binary U_S, this is the prevalence ratio for U_S, given exposure and confounders, comparing those selected into the study versus those who are not selected, or its inverse, whichever is larger.		✓
	BF_S	$g[RR_{U_SY}(a = 1), RR_{SU_S}(a = 1)] \times g[RR_{U_SY}(a = 0), RR_{SU_S}(a = 0)]$	Multiplicative bias factor, selection bias. This is interpreted as the selection-bias risk ratio or the ratio of the risk ratio under no selection bias to the risk ratio with selection bias.		✓

A, exposure of interest; C, measured confounders; OR, odds ratio; Pr, prevalence; RR, risk ratio; S, selection into study; U_C, unmeasured factor that introduces confounding bias (unmeasured confounder); U_S, unmeasured factor that introduces selection bias; Y, outcome of interest.

[a] The biases included in this particular bias analysis include (binary) differential exposure misclassification among the study population, selection bias in which bias can be envisioned as selection on a factor that results in the expected effect in the study population differing from the expected effect in the target source population, and unmeasured confounding that results from a single confounder. Note that this table presents one possible set of hypothesized biases; Smith et al. (2021) discuss a broader set of potential bias combinations for which multiple-bias bounding can be used. The function $g(a, b) = (a \times b)/(a + b - 1)$ is given by Smith et al. (2021).

Source: Smith et al. (2021).

Table A3.3. Tobacco smoking-stratified estimates of the odds ratio, and summary adjusted odds ratio (via logistic regression) inferred from Table A3.1 and results reported from Aliramaji et al. (2015)[a]

		Bladder cancer			
		Smokers		Non-smokers	
		Case participants	Control participants	Case participants	Control participants
Opium use	Yes	44	20	14	7
	No	50	11	67	137
	Total	94	31	81	144
Stratified odds ratio (smokers)				0.48	
Stratified odds ratio (non-smokers)				4.09	
Summary odds ratio				1.25	

[a] Numbers of case and control participants with exposure to both smoking and opium use, as well as marginal totals of smokers and non-smokers by case status, were given in the paper; this information could be used to complete the table.

(this may not hold well for tobacco smoking, which was strongly related with drug use in the survey, but there are insufficient data to proceed using a smoking-specific correction, given the limitations of performing bias corrections on published data). Under the assumption that exposure itself does not affect selection, given sex, prevalences from the survey can be expanded to $\Pr(A^* = a \mid U_S = u) = \Pr(A^* = a \mid U_S = u, S = s)$ and the sex-specific survey data can be used in further calculations. Some selection-bias adjustment parameters will also be based on study data, but note that the effect estimate in the study (adjusted odds ratio) is used to approximate the adjusted risk ratio from a cohort analysis, so selection parameters that rely on the study data should be estimated from only the data for control participants. In an unmatched case–control study, the control participants should represent the distribution of exposures in the source population. In this matched setting, the control participants represent a stratified sample from the source population, where the sampling proportions are derived from the distribution of sex among the case participants.

The sex-specific prevalences of drug use from the survey can then be transformed to yield prevalences of each sex in each category of exposure, using Bayes' theorem. This is given as:

$$\Pr(U_S = u \mid A^* = a, S = s) = \Pr(A^* = a \mid U_S = u, S = s) \times \Pr(U_S = u \mid S = s)/\Pr(A^* = a \mid S = s) \quad \text{(A3.1)}$$

The multiplicative factor $\Pr(U_S = u \mid S = s)/\Pr(A^* = a \mid S = s)$ can be estimated from study data, demographic data, and the population distribution of sex (here, a 1:1 female:male ratio is assumed). The sex-specific proportions in the control data are $\Pr(U_S = u \mid S = 1)$, and are given as 87.4% for men and 12.6% for women. The distribution of exposure in the control participants is given as $\Pr(A^* = a \mid S = 1)$ (15.4% exposed, 84.6% unexposed). The sex-standardized survey estimate of drug use prevalence (24.1% × 0.5 + 2.2% × 0.5 = 13.2%) is used as the assumed exposure prevalence in the target control population, and can be given as

$$\Pr(A^* = a) = \Pr(S = 1)\Pr(A^* = a \mid S = 1) + \Pr(S = 0)\Pr(A^* = a \mid S = 0) \quad \text{(A3.2)}$$

This enables solving for $\Pr(A^* = a \mid S = 0)$, noting that the selection probabilities are derived by dividing the number of control participants by the expected number of control participants if sex had not been used as a matching factor, (153 men + 22 women)/(153 + 153) = 57%, and $\Pr(A^* = 1 \mid S = 0) = 10.1\%$, which supports the idea that the unselected population will have less exposure than the selected control participants (prevalence = 15.4%), because the unselected population will include women who were omitted from the study as a consequence of matched sampling. However, this percentage is substantially higher than the female-specific prevalence of drug use in the survey data of 2.2%, which is an alternative value that could be used in a sensitivity analysis.

Finally, note that $\Pr(A^* = a \mid U_S = u, S = s) = \Pr(A^* = a \mid U_S = u)$, through the assumptions of Fig. A3.2, because selection and exposure are independent, given sex. Thus, exposure

prevalence by sex can be taken directly from the survey data estimates of 24.1% among men and 2.2% among women. Calculations of $\Pr(U_S = u \mid A^* = a, S = s)$ can then be made by application of Bayes' theorem, as before.

The risk parameters are used in the bias-factor method by taking the maximum ratio (at each level of exposure) by which U_S could increase the risk. Consequently, the exact risks are not crucial, but the risk ratio comparing bladder cancer across levels of U_S is crucial. For non-binary U_S, the risk ratio would be calculated for the lowest risk value of U_S against the highest risk value of U_S. Here, it is possible simply to fill in an arbitrary (valid) value for the risks for unexposed (or exposed) men and use the risk ratio for being female compared with that of being male among the unexposed (or exposed) participants to calculate the second set of risks. Ideally, these values could be informed through regression coefficients for sex from a study in which sex was included in a model for the risk of bladder cancer, given opium use. This would be different from a crude risk ratio contrasting men and women, because opium use is a potential mediator between sex and bladder cancer, and the parameter needed in this case is the risk ratio for sex with adjustment for opium use as a mediator. Such coefficients may not be available in the literature because they may not be a central feature of interest in a regression analysis. For example, Hadji et al. (2022) estimate an opium-adjusted odds ratio for sex but do not report the coefficient for sex in the model. Consequently, the crude odds ratio estimate for sex, 0.33, given by Hadji et al. (2022) is used to approximate the opium-conditional odds ratio. Negligible effect measure modification by sex occurred for opium use, suggesting that similar risk ratios can be used for men and women. After filling in arbitrary values for exposed and unexposed men of 0.08 and 0.02, respectively, and letting the odds ratio estimate for sex (0.33) stand in for the estimated risk ratio, risk estimates for exposed and unexposed women of 0.026 and 0.0066, respectively, are used. Note that the method is not sensitive to the absolute values of risk or to the ratio of risks between exposed groups.

Finally, given that the prevalence and risk parameters have been fully enumerated, the bias-factor calculation leads to a selection-bias factor of BF_S = 1.40; after adjustment for selection bias, the lower bound of the adjusted risk ratio is 1.24/1.40 = 0.89. At this point one might stop, if the goal is to determine whether the plausible lower bound moves across the null from the study estimates, because further adjustments will only decrease this bound. One might also refine selection-bias adjustment by calculating an additional value of BF_S for the impact of selection bias by recruiting hospitalized control participants or matching on age; the lower bound of 0.89 would be divided by this additional factor to obtain a new lower bound. This use of the survey data from Moradinazar et al. (2020) demonstrates that data from outside sources can inform bias analysis in useful ways, even if the bias parameters that are needed for analysis are not estimated directly in the study, provided that additional calculations can be performed, as was true here.

A3.2.3 Selection bias adjustment using pseudo-data

The selection bias induced by matching on sex and age means that in a study sample without this selection a different distribution of these two factors would be observed. A simple (and long-used) approach to estimate the effect of selection bias for an odds ratio (which is how the impact of opium use on bladder cancer was estimated by Aliramaji et al., 2015) is to multiply the odds ratio by the selection odds, which are calculated using the probability of selection into the study for the four combinations of case or control status and exposed or unexposed status. Fox et al. (2021) show that this is equivalent to inverse odds-of-selection weighting in this simple case of four selection parameters. Inverse odds-of-selection weighting could be extended further, to account for selection bias that occurs specifically as a result of matching on sex and age, but such an approach would rely on having much more refined estimates than are available in the study of Aliramaji et al. (2015). Thus, the simpler approach to weighting is chosen here; this is equivalent to the selection-odds approach.

The selection probabilities for the combination of case or control status and exposed or unexposed status can be inferred partly by the study design. Because the concern for selection bias is matching of the control participants, there is no issue (in this situation) with selection of the case participants in terms of bias. It would be expected that the proportion of exposed case participants in the study is equal to the proportion of exposed case participants in the source population, such that the selection probability of case participants

can be considered to be 1.0, i.e. there are no additional case participants who would have been observed in the population if matching had not been used (however, this would not be the situation if exact matching led to the exclusion of some unmatched case participants).

The selection proportions for the exposed and unexposed control participants can then be informed by the same survey data as before, given the rarity of bladder cancer in the population. These probabilities are given by $\Pr(S = 1 \mid Y = 0, A^* = 1)$ and $\Pr(S = 1 \mid Y = 0, A^* = 0)$, which are not directly given by the data or in validation data. However, these can be expanded to include sex by noting that

$$\Pr(S = 1 \mid Y = 0, A^* = a) = \sum_u \Pr(S = 1 \mid Y = 0, A^* = a, U_S = u) \Pr(U_S = u \mid A^* = a, Y = 0) \quad \text{(A3.3)}$$

which are the sex-specific (and exposure-specific) probabilities of selection and the population exposure-specific probabilities of reporting sex as male or female.

First, it is necessary to find $\Pr(S = 1 \mid Y = 0, A^* = a, U_S = u)$. Because selection into the study did not depend on exposure, conditional on sex, $\Pr(S = 1 \mid Y = 0, A^* = a, U_S = u)$ is equal to $\Pr(S = 1 \mid Y = 0, U_S = u)$. Next, using Bayes' theorem:

$$\Pr(S = 1 \mid Y = 0, U_S = u) = \Pr(U_S = u \mid S = 1, Y = 0)\Pr(S = 1 \mid Y = 0)/\Pr(U_S = u \mid Y = 0) \quad \text{(A3.4)}$$

The value of $\Pr(U_S = u \mid S = 1, Y = 0)$ is given by the study data as 87.4% for men and 12.6% for women. $\Pr(U_S = u \mid Y = 0)$ is assumed to be 50% (1:1 female:male ratio in the source population, and a rare outcome, such that the ratio in the non-case participants will be very similar). $\Pr(S = 1 \mid Y = 0)$ is the overall selection proportion for non-case participants, for whom direct data are not available but can be derived using the idea that a control group with no selection bias ought to have a 1:1 sex ratio, and would thus be expected to include 153 men and 153 women. This means that the probability of selection is (153 + 22)/(153 + 153) = 57.1%. As an example calculation for (exposed and unexposed) women, $\Pr(S = 1 \mid Y = 0, A^* = a, U_S = \text{female})$ is given as 0.126 × 0.571/0.5 = 0.143; for men, this is 0.874 × 0.571/0.5 = 0.998.

To complete the selection probabilities, it is also necessary to find $\Pr(U_S = u \mid A^* = a, Y = 0)$. Unfortunately, these probabilities are not given in the study data, but noting that for a rare disease $\Pr(U_S = u \mid A^* = a, Y = 0) \approx \Pr(U_S = u \mid A^* = a)$, this quantity can be estimated from survey data and (again) Bayes' theorem. First,

$$\Pr(U_S = u \mid A^* = a) = \Pr(A^* = a \mid U_S = u)\Pr(U_S = u)/\Pr(A^* = a) \quad \text{(A3.5)}$$

As before, exposure prevalence by sex can be taken directly from the survey data estimates of $\Pr(A^* = a \mid U_S = \text{male}) = 24.1\%$ and $\Pr(A^* = a \mid U_S = \text{female}) = 2.2\%$. A 1:1 sex ratio yields $\Pr(U_S = u) = 50\%$, and (for example) the marginal probability of misclassified exposure is given by the sex-standardized probability of exposure:

$$\sum_u \Pr(U_S = u)\Pr(A^* = 1 \mid U_S = \text{male}) = 0.5 \times 0.022 + 0.5 \times 0.241 = 0.1315 \quad \text{(A3.6)}$$

The full calculation yields selection probabilities of 0.93 for exposed control participants and 0.52 for unexposed control participants, which yields a selection-bias-adjusted relative risk of 4.9 (Table A3.4).

Although the selection-bias-adjusted relative risk of 4.9 obtained using the pseudo-data method and the lower bound relative risk of 0.89 obtained using the bias-factor method seem to give conflicting results, there are important caveats to note. First, the bias-factor method is focused on extreme circumstances, such that even if, in expectation, a bias might be downwards, the bias-factor method

Table A3.4. Selection-bias-adjusted pseudo-data, selection probabilities, and calculated selection-bias-adjusted risk ratio

		Bladder cancer		Case odds/OR
		Case participants	Control participants	
Opium use	Yes	58	29.08	1.99
	No	117	285.76	0.41
Crude/apparent odds ratio				4.87
Selection probabilities	Exposed	1	0.928	
	Unexposed	1	0.518	

OR, odds ratio.
Source: Aliramaji et al. (2015).

focuses on extrema of the bias, which could be in opposing directions from the expectation. Second, the process of adjusting for selection bias from matching using the pseudo-data method reintroduced confounding by sex (which was presumably what matching was intended to solve). Because being male is strongly positively associated with both opium use and bladder cancer, this induced confounding is expected to be upwards, such that a confounding-bias-adjusted relative risk would be expected to be less than 4.87. The topic of confounding by sex will be revisited in Section A3.2.6.

A3.2.4 Exposure misclassification bias adjustment using bias factors

Exposure misclassification is a special concern in studies of illicit drug use when self-report is used to determine drug use and in any study in which recall periods are long for defining exposure. In the study of Aliramaji et al. (2015), a hospital-based study in the Islamic Republic of Iran, patients were considered exposed if opium consumption was noted in their files from the pathology department, hospital archives, and phone calls (although scant details are given, each of these records presumably originates from self-report or physician report, rather than routine biological test results). Exposure was defined as reported duration of use greater than or equal to 1 year.

Correction for exposure misclassification via bias factors requires values of specific sensitivity and specificity for case and control participants. Ideally, a validation study to adjust the estimates of Aliramaji et al. (2015) would be able to provide estimates of each of these parameters. The most relevant study that could be identified was that of Rashidian et al. (2017), who conducted an assay-based validation study of self-reported opioid use (primarily raw opium) among patients in a hospital in the Islamic Republic of Iran (also the study population setting and country of origin for the analysis of Aliramaji et al., 2015). Self-reported regular use of opioids for 6 months or longer in the user's lifetime was selected as the target variable, which was validated by two measures: self-reported use in the previous 72 hours and, among those who did not self-report use in the previous 72 hours, immunoassay by thin-layer chromatography from urine samples taken at interview. Sensitivity and specificity were estimated as 0.775 and 0.921, respectively, among the hospital patients.

This validation study is not ideal; recent exposure could genuinely disagree with longer-term use without being a false-positive or a false-negative, and no case-specific estimates were given. However, in the validation study, investigators also noted that sensitivity was higher among hospital patients than among healthy control participants drawn from other visitors to the hospital (0.775 vs 0.688), suggesting that the similar settings of Aliramaji et al. (2015) and the validation study are a strength (although an alternative explanation is that the conditions requiring hospitalization may have increased recent opioid use among regular users). Nonetheless, the sensitivity and specificity in this study are within the range of previous studies of illicit substance use (Harrison et al., 2007).

The values of sensitivity and specificity were used to calculate a bias factor for exposure misclassification, which was assumed to be non-differential, because of the lack of information about bladder cancer status in the validation study. Notably, the bias factor used here is valid when the target parameter is an odds ratio or in situations in which the target parameter estimates an odds ratio (e.g. the risk ratio estimates the odds ratio with a rare outcome), representing a limitation of the bias-factor approach to exposure misclassification. Adjustment for outcome misclassification is not subject to a similar caveat. A further shortcoming of this approach is that some misclassification parameter values will be incompatible with the data (e.g. may result in implausible values for exposure prevalence). The misclassification bias factor was 1.0, which is a result of the observation that non-differential misclassification of a binary exposure will lead to bias away from the null, so that non-differential misclassification will not result in a reduction of the bound of plausible parameter values that are consistent with the data and bias parameters. The selection-bias- and exposure-misclassification-adjusted lower bound relative risk is equal to

$$RR_{app}/(BF_S \times BF_M) = 1.24/(1.40 \times 1.0) = 0.89 \quad \text{(A3.7)}$$

A3.2.5 Exposure misclassification bias adjustment using pseudo-data

Again using the validation data from Rashidian et al. (2017), sensitivity and specificity were estimated as 0.775 and 0.921, respectively. Fox et al. (2021) give a formula for creating pseudo-data from a 2 × 2 table (binary exposure, binary outcome), which is used in the R code

provided in Annex 2 (online only; available from: https://publications.iarc.who.int/634#supmat) (Fox et al., 2021). This yielded pseudo-data adjusted for selection bias and exposure misclassification (Table A3.5), which notably resulted in an adjusted odds ratio of 28.2 (which, it should be noted, is verging on implausible and relies on a corrected count of exposed control participants of only 6.2).

A3.2.6 Unmeasured confounding bias adjustment using bias factors

After correcting for selection bias by matching on sex, there will be residual confounding by sex in the study. One approach to this residual confounding is to treat sex as an unmeasured confounder and conduct a bias analysis. Bias analysis for unmeasured confounding through bias factors is operationally similar to that for selection bias, in that parameters for the conditional probability of the unmeasured confounder and the outcome must both be specified. That is, for an unmeasured confounder U_c with discrete levels, prevalence parameters given by $\Pr(U_c = u \mid A^* = a, C = c)$ are required; these are used to quantify the maximal relation between the confounder and exposure and risk parameters given by $\Pr(Y = 1 \mid A^* = a, C = c, U_c = u)$, which are used to quantify the maximal relation between the confounder and the outcome. This approach is quite general, because U_c can be binary, categorical, or continuous; it is identical for many scenarios, and a full distribution of the confounder does not have to be specified.

As with selection bias, the prevalence parameters are used to quantify the maximum risk ratio that contrasts prevalence values across exposure values. For example, if U_c is sex, the prevalence ratio contrasts the prevalence of being male (or female) across levels of exposure, and takes the maximum of those two prevalence ratios. Here, the only parameter of crucial interest (for sex as a binary confounder) is

$$\Pr(U_c = \text{male} \mid A^* = 1, C = c) / \Pr(U_c = \text{male} \mid A^* = 0, C = c) \quad \text{(A3.8)}$$

for which it is assumed that the measured covariates C are not crucial to the problem (e.g. the confounder–exposure relation does not change substantially after adjusting for C, and the necessary parameters can be simplified to $\Pr(U_c = u \mid A^* = a)$.)

These parameters can be drawn from survey data. As in the adjustment for selection bias, parameters are taken from the study by Moradinazar et al. (2020), which is used to present one conceptual issue: for selection bias, exposure from that study is treated as a mismeasured exposure, whereas for confounding bias it is necessarily treated as a gold standard exposure. Regardless, it is unlikely that survey data could be identified using a better measure of opium use than self-report, and an assumption that there is no unmeasured confounding can be much stronger than the assumptions inherent in bias analysis.

To calculate the bias parameters, the first calculation is

$$\Pr(U_c = u \mid A = a) = \Pr(A = a \mid U_c = u) \Pr(U_c = u) / \Pr(A = a) \quad \text{(A3.9)}$$

As before, exposure prevalence by sex can be taken directly from the survey data estimates of $\Pr(A^* = a \mid U_S = \text{male}) = 24.1\%$ and $\Pr(A^* = a \mid U_S = \text{female}) = 2.2\%$. A 1:1 sex ratio yields $\Pr(U_S = u) = 50\%$, and (for example) the marginal probability of misclassified exposure is given by the sex-standardized probability of exposure:

$$\begin{aligned} &\textstyle\sum_u \Pr(U_S = u) \Pr(A^* = 1 \mid U_S = \text{male}) \\ &= 0.5 \times 0.022 + 0.5 \times 0.241 \\ &= 0.1315 \end{aligned} \quad \text{(A3.10)}$$

Table A3.5. Selection-bias- and exposure-misclassification-adjusted pseudo-data, exposure misclassification parameters from Rashidian et al. (2017), and the calculated selection-bias-adjusted risk ratio from Aliramaji et al. (2015)

		Bladder cancer		Case odds/OR
		Case participants	Control participants	
Opium use	Yes	63.49	6.24	10.18
	No	111.51	308.61	0.36
Crude/apparent odds ratio				28.18
Misclassification	Sensitivity	0.775	0.775	
	Specificity	0.921	0.921	

OR, odds ratio.
Sources: Aliramaji et al. (2015); Rashidian et al. (2017).

The full calculation yields a prevalence of being male of 0.92 among the exposed participants and 0.44 among the unexposed participants (and can be used to calculate the same prevalences of being female).

As before, for a binary confounder, the key aspect for the risk parameters is the risk ratio comparing the risk of the outcome for men versus that for women, which was given before as 0.33. These bias parameters yielded an unmeasured-confounding bias factor of 1.53. Thus, the selection-bias-, exposure-misclassification-, and unmeasured-confounding-bias-adjusted lower bound relative risk is equal to

$$RR_{app}/(BF_S \times BF_M \times BF_C)$$
$$= 1.25/(1.40 \times 1.0 \times 1.54) = 0.58 \quad \text{(A3.11)}$$

Thus, a true odds ratio of 0.58 is a lower bound of the true odds ratio that is consistent with the smoking-adjusted odds ratio of 1.25 presented in the study of Aliramaji et al. (2015). That is, after adjustment for selection bias, exposure misclassification, and unmeasured confounding, the study results are consistent with odds ratios as low as 0.58.

A3.2.7 Unmeasured confounding bias adjustment using pseudo-data

The bias parameters used for the pseudo-data approach also include the risk ratio, comparing the risk of outcomes for men versus women and the prevalence of being male (or female), given exposure. These parameters resulted in a selection-bias-, exposure-misclassification-, and unmeasured-confounding-bias-adjusted relative risk of 18.6 (Table A3.6). Notably, this approach assumes that the odds ratio is the same across levels of the covariate (no effect measure modification for the odds ratio), as demonstrated by the sex-specific odds ratios of 18.6.

Unlike the bias-factor approach, the pseudo-data approach uses an apparent relative risk that is adjusted for confounding by smoking. Comparing the crude relative risk with the smoking-adjusted relative risk yields a measured-confounding bias of 2.71/1.25 = 2.17, indicating that the crude estimate is too high. The selection-bias-, exposure-misclassification-, and unmeasured-confounding-bias-adjusted odds ratio is further divided by this bias factor. This yields a final adjusted odds ratio of 8.6, which is adjusted for selection bias, exposure misclassification, unmeasured-confounding bias, and measured-confounding bias. This last calculation ignores the fact that estimates of confounding bias will change on adjustment for selection bias and exposure misclassification bias, but it is relatively simple to implement, and estimates are used directly from the data. Unlike the bias-factor approach, which yielded a worst-case odds ratio estimate of 0.58, the pseudo-data approach provides a best-guess odds ratio estimate of 8.6. These results are consistent with each other because they are interpreted differently. The bias-factor estimate indicates that it is possible (but not necessarily likely) that the positive study result could have occurred due to bias alone. The pseudo-data estimate indicates that the positive study result is nonetheless most likely an underestimate of the true odds ratio.

Table A3.6. Selection-bias-, exposure-misclassification-, and unmeasured-confounding-bias-adjusted pseudo-data and risk ratio derived from Aliramaji et al. (2015) and validation studies noted in the text

		Bladder cancer			
		Women		Men	
		Case participants	Control participants	Case participants	Control participants
Opium use	Yes	1.87	0.52	61.62	5.71
	No	33.42	173.76	78.09	134.85
	Total	35.29	174.28	139.71	140.56
Stratified odds ratio (women)		18.63			
Stratified odds ratio (men)				18.63	
Summary odds ratio					18.63

Source: Aliramaji et al. (2015).

A3.3 Sensitivity analysis

As explained elsewhere in this book (e.g. the multidimensional analysis in Section 4.3.2), it is useful to assess how reasonable departures from chosen parameters may influence results. The following scenarios were assessed using the pseudo-data approach: (i) measured-confounding bias only (to assess the accuracy of correcting for measured confounding as a last step); (ii) no false-positive exposures (often assumed where exposure may carry stigma); (iii) false-positive exposures among case participants only; (iv) false-positive exposures among control participants only; (v) a stronger unmeasured-confounder–outcome relation; (vi) selection bias arising from the recruitment of hospital-based control participants; and (vii) alternative exposure misclassification parameters obtained from a study by Abnet et al. (2004). Notably, point estimates from this sensitivity analysis ranged from 1.28 (false-positive exposures among case participants only) to 15.4 (additional selection bias arising from the recruitment of hospital-based control participants), but none of the point estimates was below the null (Table A3.7).

A3.4 A potential probabilistic multiple-bias analysis strategy

Either the bias-factor approach or the pseudo-data approach could be amenable to a probabilistic bias analysis, wherein the fixed values of the bias parameters given are replaced with values drawn from appropriate distributions (Table A3.2). Examples of such an approach are given elsewhere in this book (e.g. Example 4.21), and there are no additional complications to applying those approaches to multiple-bias analysis, so an explicit example of probabilistic multiple-bias analysis is omitted here. However, the R code provided in Annex 2 (online only; available from: https://publications.iarc.who.int/634#supmat) gives an example of how such an analysis could be carried out using the same functions used to conduct the multiple-bias analysis with pseudo-data discussed in Section A3.2. Crucially, the parameter distributions used in the code were arbitrarily chosen because reasonable parameter distributions could not be obtained for the example in this annex. Nonetheless, the code may be used to facilitate probabilistic bias analysis when reasonable and informative distributions can be specified over the bias parameters.

Table A3.7. Sensitivity analysis results with the pseudo-data approach using alternative bias parameters

	No bias adjustment	Bias adjustment scenario						
		Base analysis	No false-positive exposures	False-positives among cases participants only	False-positives among control participants only	Stronger confounder–outcome relation	Additional selection from hospital-based control participants	Misclassification parameters from Abnet et al. (2004)
Selection bias								
Selection probability, exposed case participants	1	1	1	1	1	1	1	1
Selection probability, unexposed case participants	1	1	1	1	1	1	1	1
Selection probability, exposed control participants	1	0.93	0.93	0.93	0.93	0.93	1.00	1.00
Selection probability, unexposed control participants	1	0.52	0.52	0.52	0.52	0.52	0.52	0.52
Exposure misclassification								
Case sensitivity	1	0.78	0.78	0.78	0.78	0.78	0.78	0.90
Case specificity	1	0.92	1.00	0.92	1.00	0.92	0.92	0.93
Control sensitivity	1	0.78	0.78	0.78	0.78	0.78	0.78	0.90
Control specificity	1	0.92	1.00	1.00	0.92	0.92	0.92	0.93
Unmeasured binary confounder								
RR($U \rightarrow Y$)	1	0.33	0.33	0.33	0.33	0.16	0.33	0.33
Pr($U = 1 \mid$ exposed)	1	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Pr($U = 1 \mid$ unexposed)	1	0.56	0.56	0.56	0.56	0.56	0.56	0.56
Measured-confounding bias	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17
Summary odds ratio	1.25	8.58	1.68	1.28	11.24	7.42	15.40	4.95

Sources: Abnet et al. (2004); Aliramaji et al. (2015).

References

Abnet CC, Saadatian-Elahi M, Pourshams A, Boffetta P, Feizzadeh A, Brennan P, et al. (2004). Reliability and validity of opiate use self-report in a population at high risk for esophageal cancer in Golestan, Iran. Cancer Epidemiol Biomarkers Prev. 13(6):1068–70. doi:10.1158/1055-9965.1068.13.6 PMID:15184266

Aliramaji A, Kaseean A, Yousefnia Pasha YR, Shafi H, Kamali S, Safari M, et al. (2015). Age distribution types of bladder cancers and their relationship with opium consumption and smoking. Caspian J Intern Med. 6(2):82–6. PMID:26221505

Fox MP, MacLehose R, Lash TL (2021). Applying quantitative bias analysis to epidemiologic data. Statistics for biology and health. 2nd ed. Cham, Switzerland: Springer. doi:10.1007/978-3-030-82673-4

Hadji M, Rashidian H, Marzban M, Naghibzadeh-Tahami A, Gholipour M, Mohebbi E, et al. (2022). Opium use and risk of bladder cancer: a multi-centre case-referent study in Iran. Int J Epidemiol. 51(3):830–8. doi:10.1093/ije/dyac031 PMID:35244716

Harrison LD, Martin SS, Enev T, Harrington D (2007). Comparing drug testing and self-report of drug use among youths and young adults in the general population. Rockville (MD), USA: Office of Applied Studies, Substance Abuse and Mental Health Services Administration, United States Department of Health and Human Services. doi:10.1037/e637122007-001

Mansournia MA, Jewell NP, Greenland S (2018). Case-control matching: effects, misconceptions, and recommendations. Eur J Epidemiol. 33(1):5–14. doi:10.1007/s10654-017-0325-0 PMID:29101596

Moradinazar M, Najafi F, Jalilian F, Pasdar Y, Hamzeh B, Shakiba E, et al. (2020). Prevalence of drug use, alcohol consumption, cigarette smoking and measure of socioeconomic-related inequalities of drug use among Iranian people: findings from a national survey. Subst Abuse Treat Prev Policy. 15(1):39. doi:10.1186/s13011-020-00279-1 PMID:32503660

Rashidian H, Hadji M, Marzban M, Gholipour M, Rahimi-Movaghar A, Kamangar F, et al. (2017). Sensitivity of self-reported opioid use in case-control studies: healthy individuals versus hospitalized patients. PLoS One. 12(8):e0183017. doi:10.1371/journal.pone.0183017 PMID:28854228

Smith LH, Mathur MB, VanderWeele TJ (2021). Multiple-bias sensitivity analysis using bounds. Epidemiology. 32(5):625–34. doi:10.1097/EDE.0000000000001380 PMID:34224471