

Chapter 8: Selection Bias

Contents

1	8.1 The Structure of Selection Bias (pp. 103-105)	2
1.1	Figure 8.1: Folic Acid and Cardiac Malformations	2
1.2	Figure 8.2: Extension with Parental Grief	2
1.3	Figures 8.3–8.6: Selection Bias from Differential Loss to Follow-Up	2
2	8.2 Examples of Selection Bias (pp. 105-109)	3
3	8.3 Selection Bias and Confounding (pp. 109-111)	4
3.1	The Firefighter Example	5
3.2	Healthy Worker Bias Revisited	5
4	8.4 Selection Bias and Censoring (pp. 111-113)	6
4.1	Counterfactual Notation for Censoring	6
5	8.5 How to Adjust for Selection Bias (pp. 113-115)	7
5.1	Inverse Probability Weighting for Selection Bias	7
5.2	The Wasabi Trial: IP Weighting in Practice	7
5.3	Identifiability Conditions for IP Weighting	7
5.4	Stratification vs. IP Weighting	8
6	8.6 Selection Without Bias (pp. 115-116)	8
6.1	When Collider Stratification Does Not Cause Bias	8
7	Summary	10
8	References	11

Suppose an investigator conducted a randomized experiment to answer the causal question “does one’s looking up to the sky make other pedestrians look up too?” She found a strong association between her looking up and other pedestrians’ looking up. Does this association reflect a causal effect? By definition of a randomized experiment, confounding bias is not expected. However, another potential problem existed: the analysis included only those pedestrians who, after having been part of the experiment, gave consent for their data to be used. Shy pedestrians and pedestrians in front of whom the investigator looked up were less likely to participate. Thus participating individuals in front of whom the investigator looked up are less likely to be shy and therefore more likely to look up. That is, the process of selection of individuals into the analysis guarantees that looking up is associated with other pedestrians’ looking up, regardless of whether it actually makes others look up.

An association created as a result of the process by which individuals are selected into the analysis is referred to as **selection bias**. Unlike confounding, this type of bias is not due to the presence of common causes of treatment and outcome, and can arise in both randomized experiments and observational studies. Like confounding, selection bias is just a form of lack of exchangeability between the treated and the untreated. This chapter provides a definition of selection bias and reviews the methods to adjust for it.

This chapter is based on Hernán and Robins (2020, chap. 8, pp. 103-116).

1 8.1 The Structure of Selection Bias (pp. 103-105)

The term “selection bias” encompasses various biases that arise from the procedure by which individuals are selected into the analysis. Here we focus on bias that would arise even if the treatment had a null effect on the outcome, i.e., **selection bias under the null** (as described in Section 6.5). The structure of selection bias can be represented using causal diagrams like Figure 8.1, which depicts dichotomous treatment A , outcome Y , and their common effect C .

1.1 Figure 8.1: Folic Acid and Cardiac Malformations

Suppose Figure 8.1 represents a study to estimate the effect of folic acid supplements A given to pregnant women shortly after conception on the fetus’s risk of developing a cardiac malformation Y (1: yes, 0: no) during the first two months of pregnancy. The variable C represents death before birth. A cardiac malformation increases mortality (arrow from Y to C), and folic acid supplementation decreases mortality by reducing the risk of malformations other than cardiac ones (arrow from A to C). The study was restricted to fetuses who survived until birth, i.e., conditioned on no death $C = 0$ (indicated by the box around node C).

$$A \rightarrow Y, \quad A \rightarrow C \leftarrow Y$$

The diagram shows two sources of association between treatment and outcome:

1. The open path $A \rightarrow Y$ representing the causal effect of A on Y
2. The open path $A \rightarrow C \leftarrow Y$ linking A and Y through their conditioned-on common effect C

An analysis conditioned on C will generally result in an association between A and Y . Because of this selection bias, the associational risk ratio $\Pr[Y = 1|A = 1, C = 0]/\Pr[Y = 1|A = 0, C = 0]$ does not equal the causal risk ratio $\Pr[Y^{a=1} = 1]/\Pr[Y^{a=0} = 1]$; association is not causation. If the analysis were not conditioned on the common effect (collider) C , then the only open path between treatment and outcome would be $A \rightarrow Y$, and thus the entire association between A and Y would be due to the causal effect of A on Y .

1.2 Figure 8.2: Extension with Parental Grief

The causal diagram in Figure 8.2 extends Figure 8.1 by adding a node S representing parental grief (1: yes, 0: no), which is affected by vital status at birth (arrow from C to S). Suppose the study was restricted to nongrieving parents ($S = 0$) because the others were unwilling to participate. As discussed in Chapter 6, conditioning on a variable S affected by the collider C also opens the path $A \rightarrow C \leftarrow Y$.

Both Figures 8.1 and 8.2 depict examples of selection bias in which the bias arises because of conditioning on a common effect of treatment and outcome: C in Figure 8.1 and S in Figure 8.2. This bias arises regardless of whether there is an arrow from A to Y , i.e., it is **selection bias under the null**.

1.3 Figures 8.3–8.6: Selection Bias from Differential Loss to Follow-Up

Consider the causal diagram in Figure 8.3, which represents a follow-up study of individuals with HIV infection to estimate the effect of certain antiretroviral treatment A on the 3-year risk of death Y (with no arrow from A to Y to reduce clutter). The unmeasured variable U represents high level of immunosuppression (1: yes, 0: no). Individuals with $U = 1$ have a greater risk of death. Individuals who drop out from the study or are otherwise lost to follow-up are censored ($C = 1$). Individuals with $U = 1$ are more likely to be censored because the severity of their disease prevents them from participating. The effect of U on censoring C is mediated by symptoms (fever, weight loss, diarrhea), CD4 count, and viral load, all included in L . Individuals receiving treatment are at a greater risk of experiencing side effects, which could lead them to drop out (arrow from A to C). The square around C indicates that the analysis is restricted to individuals who remained uncensored ($C = 0$).

According to the rules of d-separation, conditioning on the collider C opens the path $A \rightarrow C \leftarrow L \leftarrow U \rightarrow Y$ and thus association flows from treatment A to outcome Y , i.e., the associational risk ratio is not equal to 1 even though the causal risk ratio is equal to 1. The bias in Figure 8.3 is an example of selection bias that results from conditioning on censoring C , which is a common effect of treatment A and of a cause U of the outcome Y , rather than a common effect of treatment and outcome.

Figures 8.4–8.6 present additional causal diagrams that could lead to selection bias by differential loss to follow-up:

- **Figure 8.4:** Prior treatment A has a direct effect on symptoms L . Restricting to uncensored individuals implies conditioning on the common effect C of A and U , introducing an association between treatment and outcome.
- **Figures 8.5 and 8.6:** Variations of Figures 8.3 and 8.4 that include a common cause W of A and another measured variable.

Definition 1.1 (Selection Bias). **Selection bias** refers to all biases that arise from conditioning on a common effect of two variables, one of which is either the treatment or a cause of treatment, and the other is either the outcome or a cause of the outcome—including conditioning on a descendant of such a common effect.

Causal structures that result in bias under the null also cause bias when the treatment has a non-null effect. Both confounding (due to common causes) and selection bias (due to conditioning on common effects) are examples of bias under the null.

An important difference between confounding and selection bias: randomization protects against confounding, but not against selection bias when the selection occurs after the randomization.

2 8.2 Examples of Selection Bias (pp. 105-109)

The causal diagrams in Figures 8.3–8.6 can represent several types of selection bias:

Definition 2.1 (Differential Loss to Follow-Up). **Differential loss to follow-up**, also called bias due to **informative censoring**, arises when the variable C in Figures 8.3–8.6 represents censoring: restricting the analysis to individuals who remain uncensored ($C = 0$) induces the selection bias described in the previous section.

Definition 2.2 (Missing Data Bias / Nonresponse Bias). **Missing data bias**, also called **nonresponse bias**, arises when the variable C in Figures 8.3–8.6 represents missing data on the outcome for any reason, not just as a result of loss to follow-up—for example, because individuals are reluctant to provide information or because they miss study visits. Restricting the analysis to individuals with complete data ($C = 0$) may result in bias.

Definition 2.3 (Healthy Worker Bias). **Healthy worker bias** arises when estimating the effect of an occupational exposure A (e.g., a chemical) on mortality Y in a cohort of factory workers. The underlying unmeasured true health status U is a determinant of both death Y and of being at work C (1: not at work, 0: at work), and the study is restricted to individuals who are at work ($C = 0$) at the time of outcome ascertainment. Being exposed to the chemical reduces the probability of being at work in the near future, either directly (e.g., exposure can cause disabling asthma) like in Figures 8.3 and 8.4, or through a common cause W like in Figures 8.5 and 8.6.

Definition 2.4 (Self-Selection Bias / Volunteer Bias). **Self-selection bias**, also called **volunteer bias**, arises in a study in which C is agreement to participate (1: no, 0: yes)—for example, a study in which A is cigarette smoking, Y is coronary heart disease, U is family history of heart disease, and W is healthy lifestyle. Under any of the structures in Figures 8.3–8.6, selection bias may be present if the study is restricted to those who volunteered or elected to participate ($C = 0$).

Definition 2.5 (Selection Affected by Treatment Received Before Study Entry). If treatment A took place before the study started and affects the probability of being selected into the study, selection bias is expected. This bias may be present in any study that attempts to estimate the causal effect of a treatment that occurred before the study started or in which treatment includes a pre-study component.

These examples show that selection bias may occur in retrospective studies—those in which data on treatment A are collected after the outcome Y occurs—and in prospective studies—those in which data on treatment A are collected before the outcome Y occurs. Further, selection bias may occur both in observational studies and in randomized experiments (when selection happens after randomization).

No bias arises in randomized experiments from selection into the study before treatment is assigned. For example, only volunteers who agree to participate are enrolled in randomized clinical trials, but such trials are not affected by volunteer bias because participants are randomly assigned to treatment only after agreeing to participate ($C = 0$). Thus none of Figures 8.3–8.6 can represent volunteer bias in a randomized trial.

i Fine Point 8.1: Selection Bias in Case-Control Studies

Figure 8.1 can be used to represent selection bias in a case-control study. Suppose an investigator wants to estimate the effect of postmenopausal estrogen treatment A on coronary heart disease Y . The variable C indicates whether a woman in the study population is selected for the case-control study (1: no, 0: yes). The arrow from disease status Y to selection C indicates that cases in the population are more likely to be selected than noncases.

In this particular case-control study, the investigator decided to select controls ($Y = 0$) preferentially among women with a hip fracture. Because treatment A has a protective causal effect on hip fracture, the selection of controls with hip fracture implies that treatment A now has a causal effect on selection C (arrow $A \rightarrow C$). The association measure (the treatment-outcome odds ratio) is by definition conditional on having been selected into the study ($C = 0$). If individuals with hip fracture are oversampled as controls, the probability of control selection depends on a consequence of treatment A , and “inappropriate control selection” bias will occur. This bias arises because we are conditioning on a common effect C of treatment and outcome.

Other forms of selection bias in case-control studies, including some biases described by Berkson (1946) and incidence-prevalence bias, can also be represented by Figure 8.1 or modifications of it.

3 8.3 Selection Bias and Confounding (pp. 109-111)

In the previous chapter and in this chapter, we described two reasons why the treated and the untreated may not be exchangeable:

1. The presence of **common causes** of treatment and outcome (confounding)
2. **Conditioning on common effects** of treatment and outcome or causes of them (selection bias)

This structural definition provides a clear-cut classification, even though it might not coincide perfectly with the traditional terminology of some disciplines.

3.1 The Firefighter Example

Consider a study restricted to firefighters that aims to estimate the causal effect of being physically active A on the risk of heart disease Y (Figure 8.7). For simplicity, assume that A does not cause Y . Parental socioeconomic status L affects both the risk of becoming a firefighter C and, through childhood diet, of heart disease Y . Attraction toward activities that involve physical activity (an unmeasured variable U) affects both the risk of becoming a firefighter and of being physically active (A). U does not affect Y , and L does not affect A .

According to our terminology, there is no confounding because there are no common causes of A and Y . Thus, in the full population, the associational risk ratio $\Pr[Y = 1|A = 1]/\Pr[Y = 1|A = 0]$ is expected to equal the causal risk ratio $\Pr[Y^{a=1} = 1]/\Pr[Y^{a=0} = 1] = 1$.

However, in a study restricted to firefighters ($C = 0$), the associational and causal risk ratios would differ because conditioning on a common effect C of causes of treatment and outcome induces selection bias. To the study investigators, the distinction between confounding and selection bias is moot because, regardless of nomenclature, they must adjust for L to make the treated and untreated firefighters comparable.

There are advantages to adopting a structural approach to classifying sources of non-exchangeability:

1. The structure frequently guides the choice of analytical methods to reduce or avoid the bias. In longitudinal studies with time-varying treatments, identifying the structure allows us to detect situations in which adjustment for confounding via stratification would introduce selection bias.
2. Even when understanding the structure does not have implications for data analysis, it could still help study design.
3. Selection bias resulting from conditioning on pre-treatment variables could explain why certain variables behave as “confounders” in some studies but not others.
4. Causal diagrams enhance communication among investigators.

3.2 Healthy Worker Bias Revisited

The term “healthy worker bias” is used to describe two structurally different biases:

1. The bias described in Section 8.2, which arises from conditioning on the variable C —a common effect of (a cause of) treatment and (a cause of) the outcome. This is selection bias and can be represented by Figures 8.3–8.6.
2. The bias that occurs when comparing the risk in a group of workers with that in a group of individuals from the general population. Here L represents health status, A represents membership in the group of workers, and Y represents the outcome. There are arrows from L to A and Y because being healthy affects job type and risk of subsequent outcome. This is confounding (from the common cause L) and can be represented by Figure 7.1.

The use of causal diagrams to represent the structure of the “healthy worker bias” prevents confusions that may arise from employing the same term for different sources of non-exchangeability.

i Technical Point 8.1: The Built-in Selection Bias of Hazard Ratios

The causal DAG in Figure 8.8 describes a randomized experiment of the effect of heart transplant A on death at times 1 (Y_1) and 2 (Y_2). The arrow from A to Y_1 represents that transplant decreases the risk of death at time 1. The lack of an arrow from A to Y_2 indicates that A has no direct effect on death at time 2. The unmeasured haplotype U decreases the individual’s risk of death at all times.

The time-specific hazard ratio at time 2 is $\Pr[Y_2 = 1|A = 1, Y_1 = 0]/\Pr[Y_2 = 1|A = 0, Y_1 = 0]$,

which conditions on having survived past time 1 (the square around Y_1). Treated survivors of time 1 are less likely than untreated survivors of time 1 to have the protective haplotype U (because treatment can explain their survival) and therefore are more likely to die at time 2. Thus, the hazard ratio at time 1 is less than 1, whereas the hazard ratio at time 2 is greater than 1, i.e., the hazards have crossed.

The hazard ratio at time 2 is a biased estimate of the direct effect of treatment on mortality at time 2. The bias is selection bias arising from conditioning on a common effect Y_1 of treatment A and of U , which is a cause of Y_2 that opens the associational path $A \rightarrow Y_1 \leftarrow U \rightarrow Y_2$.

4 8.4 Selection Bias and Censoring (pp. 111-113)

Suppose an investigator conducted a marginally randomized experiment to estimate the average causal effect of wasabi intake on the one-year risk of death ($Y = 1$). Half of the 60 study participants were randomly assigned to eating meals supplemented with wasabi ($A = 1$) until the end of follow-up or death, whichever occurred first. The other half were assigned to meals that contained no wasabi ($A = 0$). After 1 year, 17 individuals died in each group. That is, the associational risk ratio $\Pr[Y = 1|A = 1]/\Pr[Y = 1|A = 0]$ was 1. Because of randomization, the causal risk ratio $\Pr[Y^{a=1} = 1]/\Pr[Y^{a=0} = 1]$ is also expected to be 1.

Unfortunately, the investigator could not observe the 17 deaths that occurred in each group because many patients were lost to follow-up, or censored, before the end of the study. The proportion of censoring ($C = 1$) was higher among patients with heart disease ($L = 1$) at the start of the study and among those assigned to wasabi supplementation ($A = 1$). In fact, only 9 individuals in the wasabi group and 22 individuals in the other group were not lost to follow-up. The investigator observed 4 deaths in the wasabi group and 11 deaths in the other group. That is, the associational risk ratio $\Pr[Y = 1|A = 1, C = 0]/\Pr[Y = 1|A = 0, C = 0]$ was $(4/9)/(11/22) = 0.89$ among the uncensored. The risk ratio of 0.89 in the uncensored differs from the causal risk ratio of 1 in the entire population: there is selection bias due to conditioning on the common effect C .

The causal diagram in Figure 8.3 depicts the relation between the variables L , A , C , and Y in the randomized trial of wasabi. U represents atherosclerosis, an unmeasured variable, that affects both heart disease L and death Y . Figure 8.3 shows there are no common causes of A and Y , as expected in a marginally randomized experiment, and thus there is no need to adjust for confounding to compute the causal effect of A on Y .

The backdoor criterion says that adjustment for the causal effect of censoring C on Y (which is null in Figure 8.3) is possible because the measured variable L can be used to block the backdoor path $C \leftarrow L \leftarrow U \rightarrow Y$.

4.1 Counterfactual Notation for Censoring

The causal contrast of interest needs to be modified in the presence of censoring. Because selection bias would not exist if everybody had been uncensored, we want to consider a causal contrast that reflects what would have happened in the absence of censoring.

Definition 4.1 (Causal Contrast Under Censoring). Let $Y^{a=1,c=0}$ be an individual's counterfactual outcome if they had received treatment $A = 1$ and had remained uncensored $C = 0$. Similarly, let $Y^{a=0,c=0}$ be an individual's counterfactual outcome if they had not received treatment $A = 0$ and had remained uncensored $C = 0$. Our causal contrast of interest is:

$$\Pr[Y^{a=1,c=0} = 1] \quad \text{versus} \quad \Pr[Y^{a=0,c=0} = 1]$$

By conceptualizing the causal contrast of interest in terms of $Y^{a,c=0}$, we can think of censoring C as just another treatment. The goal of the analysis is to compute the causal effect of a joint intervention on A and C . To eliminate selection bias for the effect of treatment A , we need to adjust for confounding for the effect of treatment C .

Since censoring C is now viewed as a treatment, we need to ensure that the identifiability conditions of exchangeability, positivity, and consistency hold for C as well as for A , and use analytical methods appropriate to both.

5 8.5 How to Adjust for Selection Bias (pp. 113-115)

Though selection bias can sometimes be avoided by an adequate design (see Fine Point 8.1), it is often unavoidable. Loss to follow-up, self-selection, and missing data leading to bias can occur no matter how careful the investigator. In those cases, the selection bias needs to be explicitly corrected in the analysis.

5.1 Inverse Probability Weighting for Selection Bias

This correction can sometimes be accomplished by IP weighting (or by standardization), which is based on assigning a weight W^C to each selected individual ($C = 0$) so that she accounts in the analysis not only for herself, but also for those like her—with the same values of L and A —who were not selected ($C = 1$). The IP weight W^C is the inverse of the probability of her selection $\Pr[C = 0 | L, A]$:

$$W^C = \frac{1}{\Pr[C = 0 | L, A]}$$

5.2 The Wasabi Trial: IP Weighting in Practice

Consider again the wasabi randomized trial. The tree graph in Figure 8.10 presents the trial data. Of the 60 individuals in the trial, 40 had ($L = 1$) and 20 did not have ($L = 0$) heart disease at the time of randomization. Regardless of their L status, all individuals had a 50/50 chance of being assigned to wasabi supplementation ($A = 1$). Thus 10 individuals in the $L = 0$ group and 20 in the $L = 1$ group received treatment $A = 1$.

The probability of remaining uncensored varies across branches in the tree. For example:

- 50% of individuals without heart disease assigned to wasabi ($L = 0, A = 1$) remained uncensored
- 60% of individuals with heart disease assigned to no wasabi ($L = 1, A = 0$) remained uncensored

Look at the bottom of the tree. There are 20 individuals with heart disease ($L = 1$) who were assigned to wasabi supplementation ($A = 1$). Of these, 4 remained uncensored and 16 were lost to follow-up. The conditional probability of remaining uncensored in this group is $\Pr[C = 0 | L = 1, A = 1] = 4/20 = 0.2$. In an IP-weighted analysis the 16 censored individuals receive a zero weight, whereas the 4 uncensored individuals each receive a weight of $5 = 1/0.2$ —representing themselves and 4 others like them. IP weighting replaces the 20 original individuals by 5 copies of each of the 4 uncensored individuals.

The same procedure can be repeated for all branches of the tree (Figure 8.11) to construct a pseudo-population of the same size as the original study population but in which nobody is lost to follow-up. The associational risk ratio in the pseudo-population is 1, the same as the causal risk ratio $\Pr[Y^{a=1, c=0} = 1] / \Pr[Y^{a=0, c=0} = 1]$ that would have been computed in the original population if nobody had been censored.

When both confounding and selection bias exist, the product weight $W^A W^C$ can be used to adjust simultaneously for both biases, where $W^A = 1/f(A|L)$ adjusts for confounding and $W^C = 1/\Pr[C = 0 | A, L]$ adjusts for selection bias.

5.3 Identifiability Conditions for IP Weighting

The association measure in the pseudo-population equals the effect measure in the original population if three identifiability conditions are met:

1. **Exchangeability:** The average outcome in the uncensored individuals must equal the unobserved average outcome in the censored individuals with the same values of A and L . This requires that the variables in A and L are sufficient to block all backdoor paths between C and Y .
2. **Positivity:** All conditional probabilities of being uncensored given A and the variables in L must be greater than zero, i.e., $\Pr[C = 0|A, L] > 0$ for all relevant values of A and L . (Note this positivity condition is required for the probability of being uncensored, not for the probability of being censored.)
3. **Consistency:** The interventions are sufficiently well-defined. IP weighting creates a pseudo-population in which censoring C has been abolished, and in which the effect of the treatment A is the same as in the original population. This is relatively well defined when censoring is the result of loss to follow-up or nonresponse, but may be problematic when censoring is defined as the occurrence of a competing event (such as death from causes other than the outcome of interest).

5.4 Stratification vs. IP Weighting

One might attempt to remove selection bias by stratification (estimating the effect conditional on L) rather than by IP weighting. Stratification could yield unbiased conditional effect measures within levels of L because conditioning on L is sufficient to block the backdoor path from C to Y in Figure 8.3:

$$\Pr[Y = 1|A = 1, C = 0, L = l] / \Pr[Y = 1|A = 0, C = 0, L = l]$$

However, stratification would not work under the causal structures depicted in Figures 8.4 and 8.6. In Figure 8.4, conditioning on L blocks the backdoor path from C to Y but also opens the path $A \rightarrow L \leftarrow U \rightarrow Y$ from A to Y because L is a collider on that path. In contrast, IP weighting appropriately adjusts for selection bias under Figures 8.3–8.6 because it estimates unconditional effect measures after reweighting the individuals.

This is the first time we discuss a situation in which stratification cannot be used to validly compute the causal effect of treatment, even if the three conditions of exchangeability, positivity, and consistency hold. We will discuss other situations with a similar structure in Part III when considering the effect of time-varying treatments.

6 8.6 Selection Without Bias (pp. 115-116)

The causal diagram in Figure 8.12 represents a hypothetical study with dichotomous variables surgery A , certain genetic haplotype E , and death Y . According to the rules of d-separation, surgery A and haplotype E are:

1. Marginally independent: the probability of receiving surgery is the same for people with and without the genetic haplotype
2. Associated conditionally on Y : the probability of receiving surgery varies by haplotype when the study is restricted to, say, the survivors ($Y = 0$)

Indeed, conditioning on the common effect Y of two independent causes A and E always induces a conditional association between A and E in at least one of the strata of Y .

6.1 When Collider Stratification Does Not Cause Bias

Suppose A and E affect survival through totally independent mechanisms such that E cannot possibly modify the effect of A on Y , and vice versa. For example, suppose surgery A affects survival through the removal of a tumor, whereas the haplotype E affects survival through increasing LDL-cholesterol levels resulting in an increased risk of heart attack (whether or not a tumor is present).

We can consider three cause-specific mortality variables: death from tumor Y_A , death from heart attack Y_E , and death from any other causes Y_O . The observed mortality variable $Y = 1$ (death) when Y_A or Y_E or Y_O is 1, and $Y = 0$ (survival) when all three equal 0.

The causal diagram in Figure 8.13 (an expansion of Figure 8.12) represents this causal structure. Because the arrows from Y_A , Y_E , and Y_O to Y are deterministic, conditioning on observed survival ($Y = 0$) is equivalent to simultaneously conditioning on $Y_A = 0$, $Y_E = 0$, and $Y_O = 0$. As a consequence, A and E are conditionally independent given $Y = 0$: conditioning on collider $Y = 0$, the path between A and E through Y is blocked by conditioning on the non-colliders Y_A , Y_E , and Y_O .

When the data can be summarized by Figure 8.13, we say that the data follow a multiplicative survival model (see Technical Point 8.2). In other words, collider stratification is not always a source of selection bias.

Definition 6.1 (Multiplicative Survival Model). When the conditional probability of survival $\Pr[Y = 0|E = e, A = a]$ given A and E is equal to a product $g(e)h(a)$ of functions of e and a , we say that a **multiplicative survival model** holds:

$$\Pr[Y = 0|E = e, A = a] = g(e)h(a)$$

This is equivalent to a model that assumes the survival ratio $\Pr[Y = 0|E = e, A = a]/\Pr[Y = 0|E = e, A = 0]$ does not depend on e and is equal to $h(a)$. The data follow a multiplicative survival model when there is no interaction between A and E for $Y = 0$ on the multiplicative scale.

In contrast, A and E will not be independent conditionally on $Y = 0$ when any of the following situations occur (Figures 8.14–8.16):

- If A and E affect survival through a common mechanism, there will exist an arrow either from A to Y_E or from E to Y_A (Figure 8.14)
- If Y_A and Y_E are not independent because of a common cause V (Figure 8.15)
- If the causes Y_A and Y_O , and Y_E and Y_O , are not independent because of common causes W_1 and W_2 (Figure 8.16)

In summary, conditioning on a collider always induces an association between its causes, but this association could be restricted to certain levels of the common effect. Selection on a common effect does not always result in selection bias when the analysis is restricted to a single level of the common effect.

i Technical Point 8.2: Multiplicative Survival Model

A proof that Figure 8.13 represents a multiplicative survival model (defined in the main text):

$$\Pr[Y = 0|E = e, A = a] = \Pr[Y_A = 0, Y_E = 0, Y_O = 0|E = e, A = a] = \Pr[Y_O = 0] \Pr[Y_A = 0|A = a] \Pr[Y_E = 0|E = e]$$

where the first equality is by determinism and the second by the DAG factorization. Setting $g(e) = \Pr[Y_E = 0|E = e]$ and $h(a) = \Pr[Y_O = 0] \Pr[Y_A = 0|A = a]$ confirms the multiplicative structure.

Note that when A and E are conditionally independent given $Y = 0$, they will be conditionally dependent given $Y = 1$.

i Fine Point 8.2: The Strength and Direction of Selection Bias

We have referred to selection bias as an “all or nothing” issue: either bias exists or it doesn’t. In practice, however, it is important to consider the expected direction and magnitude of the bias.

The direction of the conditional association between two marginally independent causes A

and E within strata of their common effect Y depends on how A and E interact to cause Y . For example, suppose that, in the presence of an undiscovered background factor U that is unassociated with A or E , having either $A = 1$ or $E = 1$ is sufficient and necessary to cause death (an “or” mechanism), but that neither A nor E causes death in the absence of U . Then among those who died ($Y = 1$), A and E will be negatively associated, because it is more likely that an individual with $A = 0$ had $E = 1$ (since the absence of A increases the chance that E was the cause of death).

Alternatively, suppose that having both $A = 1$ and $E = 1$ is sufficient and necessary to cause death (an “and” mechanism). In this case, among those who died ($Y = 1$), A and E will be positively associated: knowing that an individual with $A = 1$ died implies that $E = 1$ must also have been present for the death to occur.

7 Summary

This chapter examined **selection bias**, a form of lack of exchangeability that arises from conditioning on common effects rather than common causes.

Key concepts:

1. **Structure of selection bias:** Arises from conditioning on a common effect (collider) of treatment and outcome, or a variable affected by such a common effect (Figures 8.1–8.6)
2. **Examples** (Figures 8.3–8.6):
 - Differential loss to follow-up and informative censoring
 - Missing data and nonresponse bias
 - Healthy worker bias
 - Self-selection and volunteer bias
 - Selection affected by pre-study treatment
3. **Selection bias vs. confounding:**
 - Confounding: common causes ($A \leftarrow L \rightarrow Y$)
 - Selection bias: conditioning on common effects ($A \rightarrow C \leftarrow Y$)
 - Randomization prevents confounding but not selection bias occurring after randomization
4. **Censoring as selection:** In the wasabi trial, restricting to uncensored individuals ($C = 0$) gives a risk ratio of 0.89 instead of the true causal value of 1. The counterfactual of interest becomes $Y^{a,c=0}$, and censoring C is treated as another treatment.
5. **Adjustment:** IP weighting with $W^C = 1/\Pr[C = 0|L, A]$ creates a pseudo-population free of selection bias. Stratification adjusts for selection bias only in some causal structures (Figures 8.3, 8.5) but not others (Figures 8.4, 8.6).
6. **Selection without bias:** Collider stratification does not always induce selection bias. When data follow a multiplicative survival model (Figure 8.13), conditioning on survivors does not create a spurious association.

Practical implications:

1. **Study design:** Minimize selection bias by maximizing follow-up and reducing loss to follow-up.
2. **Analysis:** Use causal diagrams (DAGs) to identify the structure of selection bias and determine appropriate adjustment methods.
3. **Key warning:** Stratification can remove selection bias in simple structures (Figure 8.3) but can introduce additional bias in other structures (Figure 8.4) by opening a collider path on the treatment variable. IP weighting is more general.

Looking ahead:

- **Chapter 9:** Measurement bias

- **Chapters 12–15:** Advanced estimation methods including censoring and time-varying treatments

8 References

Hernán, Miguel A, and James M Robins. 2020. *Causal Inference: What If*. Chapman & Hall/CRC.
<https://miguelhernan.org/whatifbook>.