Covariate balance for no confounding in the sufficient-cause model

Etsuji Suzuki MD, PhD[a, b], Toshihide Tsuda MD, PhD[c], Eiji Yamamoto PhD[d]

[a] Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. Email: esuzuki@hsph.harvard.edu

[b] Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, 2-5-1 Shikata-cho, Kita-ku, Okayama 700-8558, Japan. Email: etsuji-s@cc.okayama-u.ac.jp

[c] Department of Human Ecology, Graduate School of Environmental and Life Science, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan. Email: tsudatos@md.okayama-u.ac.jp

[d] Department of Information Science, Faculty of Informatics, Okayama University of Science, 1-1 Ridai-cho, Kita-ku, Okayama 700-0005, Japan. Email: eiji@mis.ous.ac.jp

Corresponding author: Etsuji Suzuki
2-5-1 Shikata-cho, Kita-ku, Okayama 700-8558, Japan
Tel: +81-86-235-7174, Fax: +81-86-235-7178
E-mail: esuzuki@hsph.harvard.edu; etsuji-s@cc.okayama-u.ac.jp

ABSTRACT

Purpose: To show conditions of covariate balance for no confounding in the sufficient-cause model and discuss its relationship with exchangeability conditions.

Methods: We consider the link between the sufficient-cause model and the counterfactual model, emphasizing that the target population plays a key role when discussing these conditions. Furthermore, we incorporate sufficient causes within the directed acyclic graph framework. We propose to use each of the background factors in sufficient causes as representing a set of covariates of interest and discuss the presence of covariate balance by comparing joint distributions of the relevant background factors between the exposed and the unexposed groups.

Results: We show conditions for partial covariate balance, covariate balance, and full covariate balance, each of which is stronger than partial exchangeability, exchangeability, and full exchangeability, respectively. This is consistent with the fact that the sufficient-cause model is a "finer" model than the counterfactual model.

Conclusions: Covariate balance is a sufficient, but not a necessary, condition for no confounding irrespective of the target population. Although our conceptualization of covariate imbalance is closely related to the recently proposed counterfactual-based definition of a confounder, the concepts of covariate balance and confounder should be clearly distinguished.

Keywords: bias; causality; confounding factors; epidemiologic methods

List of abbreviations: DAG (directed acyclic graph)

Introduction

Since the publication of the seminal paper by Greenland and Robins [1], the counterfactual approach to confounding has been widely accessible to epidemiologists, and the concept of confounding is now explained in the counterfactual framework [2-7]. Much of the literature on this topic has explained that exchangeability between the exposed and the unexposed groups is a core concept to make causal inference. In this context, covariate balance is often addressed as a key feature to control confounding in epidemiology, and many researchers have been concerned about whether covariate balance is achieved between the exposed and the unexposed groups in their analyses. Despite its significance, however, a covariate is broadly defined as a "variable that is possibly predictive of the outcome under study" [8], and the term "covariate" has been often used interchangeably with the term "confounder".

In this article, we aim to show conditions of covariate balance for no confounding in the sufficient-cause model and discuss its relationship with exchangeability conditions. In so doing, we consider the link between the sufficient-cause model and the counterfactual model, emphasizing that covariate balance depends on the target population of interest.

The link between the sufficient-cause model and the counterfactual model

The sufficient-cause model and the counterfactual model have become cornerstones for causal thinking in epidemiology [9-11], and the link between these models has been addressed [12-15]. In this section, we provide a brief overview of the link between these two fundamental causal models in a situation in which there is a binary cause E (1 = exposed, 0 = unexposed) and a binary outcome Y (1 = outcome occurred, 0 = outcome did not occur).

In the counterfactual framework, we let Y_e denote the potential outcomes for an individual if, possibly contrary to fact, there had been interventions to set $E = e$. Throughout this article, we will assume that the consistency assumption is met [16, 17], which implies that the observed outcome for an individual is the potential outcome, as a function of intervention, when the intervention is set to the observed exposure. For each individual, there would thus be two possible potential outcomes, Y_1 and Y_0, corresponding to what would have happened to that individual had he or she been exposed and unexposed, respectively. As a result, individuals can be classified into four different response types, as enumerated in Table 1 [1]. We let p_j, q_j, and r_j, j = 1–4, be proportions of response type j in the exposed group, the unexposed group, and the total population, respectively.

In the sufficient-component cause framework [11], each sufficient cause for the outcome might require the presence of E, the presence of $\overline{E}$, or may not require either, where we let $\overline{E}$ denote the complement of E in the terminology of events. We could thus enumerate three different types of sufficient causes for Y along with certain background factors C_k: C_1, C_2E, and $C_3\overline{E}$. Here, C_k denotes a set of all components or factors, other than the presence of E and $\overline{E}$, that may be required for a particular mechanism to operate. We assume that C_k is an actual or intrinsic biological factors. For simplicity, we denote the presence of these background factors as C_k = 1 and their absence as $C_k = 0$. An individual is at risk of, or susceptible to, sufficient cause k if C_k is present for that person. Note that an individual is of one, and only one, response type in the counterfactual framework, whereas an individual may be at risk of none, one, or several sufficient causes. Then, we can enumerate eight (i.e., 2^3) types of possible risk status for sufficient causes (Table 1). We let s_j, t_j, and u_j, $j = 1–8$, be proportions of risk status type j in the exposed group, the unexposed group, and the total population, respectively.

As illustrated in Table 1, the potential outcomes of Y can be described using the background factors C_k as: $Y_1 = \max\left(C_1, C_2\right)$ and $Y_0 = \max\left(C_1, C_3\right)$. These descriptions provide important implications in the following discussion.

Table 1. Correspondence between response types and risk status types under a binary exposure and a binary outcome [a]

Response types						Risk status types						
Type	**Potential outcomes**		**Proportion of types in [b]**			**Type**	**Background factors**			**Proportion of types in [b]**		
	$\boldsymbol{Y_1}$	$\boldsymbol{Y_0}$	**Exposed**	**Unexposed**	**Total population**		$\boldsymbol{C_1}$	$\boldsymbol{C_2}$	$\boldsymbol{C_3}$	**Exposed**	**Unexposed**	**Total population**
1	1	1	p_1	q_1	r_1	1[d]	1	1	1	s_1	t_1	u_1
						2	1	1	0	s_2	t_2	u_2
						3[d]	1	0	1	s_3	t_3	u_3
						4	1	0	0	s_4	t_4	u_4
						5[d]	0	1	1	s_5	t_5	u_5
2	1	0	p_2	q_2	r_2	6	0	1	0	s_6	t_6	u_6
3[c]	0	1	p_3	q_3	r_3	7[d]	0	0	1	s_7	t_7	u_7
4	0	0	p_4	q_4	r_4	8	0	0	0	s_8	t_8	u_8

[a] We consider a binary exposure E (1 = exposed, 0 = unexposed) and a binary outcome Y (1 = outcome occurred, 0 = outcome did not occur). We consider two potential outcomes, Y_e, for an individual. We consider three different sufficient causes for outcome Y along with certain binary background factors as follows: C_1, C_2E, and $C_3\overline{E}$, where we let $\overline{E}$ denote the complement of E.

[b] Note that r_j can be calculated as: $p_j \times P[E=1]+q_j \times P[E=0]$, where $P[E=e]$ represents the prevalence of E = e in the total population. Likewise, u_j can be calculated as: $s_j \times P[E=1]+t_j \times P[E=0]$.

[c] Under the assumption of (counterfactual) positive monotonicity (i.e., $Y_0 \leq Y_1$ for all individuals), this response type is excluded [15, 18].

[d] Under the assumption of no preventive action [18], or sufficient-cause positive monotonicity [19] (i.e., $C_3 = 0$ for all individuals), these risk status types are excluded [15, 20]. Note that no preventive action implies positive monotonicity.

The concept of confounding and exchangeability conditions in the counterfactual model

Before showing conditions for covariate balance, we provide a brief overview of the counterfactual approach to confounding in this section [1-6, 21-23]. (In this article, we primarily use the notion of confounding *in distribution* [24-26]. See Appendix A for further discussion.) In the counterfactual model, a causal effect is defined on the basis of contrasts between potential outcomes under different exposure status. Thus, when the exposed is the target population, we compare the incidence proportion under E = 1 in the exposed and the incidence proportion under E = 0 in the exposed. The former quantity is, by definition, observable or estimable, and it is described as $P\left[Y_1 = 1 | E = 1\right]$ or (p_1 + p_2). By contrast, the latter quantity, $P\left[Y_0 = 1 | E = 1\right]$ or (p_1 + p_3), is unobservable because it is counterfactual. Thus, we use the actual unexposed group as a substitute of what would have occurred in the actual exposed group had they been unexposed. In other words, we use the incidence proportion in the unexposed group, $P\left[Y_0 = 1 | E = 0\right]$ or (q_1 + q_3), as a substitute of $P\left[Y_0 = 1 | E = 1\right]$ or (p_1 + p_3). Thus, confounding corresponds to the difference between the desired counterfactual quantity and the observed substitute, and a sufficient and necessary condition for no confounding is given by [1]:

$$\begin{aligned}&P\left[Y_0 = y | E = 1\right] = P\left[Y_0 = y | E = 0\right] \quad (y = 0,1)\\&\Leftrightarrow Y_0 \amalg E. \qquad\qquad [\text{Eq. 1}]\\&\left(\Leftrightarrow (p_1 + p_3) = (q_1 + q_3)\right)\end{aligned}$$

Conversely, when the unexposed group is the target population, a sufficient and necessary condition for no confounding is given by:

$$\begin{aligned}&P\left[Y_1 = y | E = 1\right] = P\left[Y_1 = y | E = 0\right] \quad (y = 0,1)\\&\Leftrightarrow Y_1 \amalg E. \qquad\qquad [\text{Eq. 2}]\\&\left(\Leftrightarrow (p_1 + p_2) = (q_1 + q_2)\right)\end{aligned}$$

Finally, when the target is the total population, a sufficient and necessary condition for no confounding is given by:

$$\begin{aligned}&\left\{P\left[Y_1 = y\right] = P\left[Y_1 = y | E = 1\right]\right\} \wedge \left\{P\left[Y_0 = y\right] = P\left[Y_0 = y | E = 0\right]\right\} \quad (y = 0,1)\\&\Leftrightarrow Y_e \amalg E \quad (e = 0,1). \qquad\qquad [\text{Eq. 3}]\\&\left(\Leftrightarrow \left\{(p_1 + p_2) = (q_1 + q_2)\right\} \wedge \left\{(p_1 + p_3) = (q_1 + q_3)\right\}\right)\end{aligned}$$

If Equation 3 holds, the groups that are actually exposed and unexposed are representative of what would have occurred had the total population been exposed and unexposed, respectively. Equations 1 and 2 are referred to as partial exchangeability conditions, whereas Equation 3 is referred to as exchangeability condition [1, 27].

We should note that complete comparability of response types between the exposed and the unexposed groups (i.e., $\left(Y_0, Y_1\right) \coprod E$ or $\left(p_1, p_2, p_3, p_4\right) = \left(q_1, q_2, q_3, q_4\right)$) is a sufficient, but not a necessary, condition for no confounding in the three target populations. This is referred to as full exchangeability condition [27].

Covariate balance for no confounding in the sufficient-cause model

In this section, we propose to show conditions of covariate balance for no confounding in the sufficient-cause model. This would fit the concept of covariate balance because one innately focuses on the “factor” or “mechanism” that induces confounding when discussing it. To begin with, note that we use each of the three background factors, C_1, C_2, and C_3, as representing a set of covariates of interest. These background factors may be several combinations of variables, each of which is part of the sufficient causes. Note also that this conceptualization well describes that these factors are “predictors” of the outcome Y. To illustrate our discussion, we incorporate sufficient causes into the directed acyclic graph (DAG) framework [28, 29]. In Figure 1, we show all of the sufficient causes for outcome Y as nodes, and add an ellipse around the sufficient-cause nodes to indicate that the set of sufficient causes is determinative. In the following discussion, we do not have to assume that the background factors are marginally independent of one another, e.g., due to shared component cause(s). However, we do not show the possible dependencies between the background factors to simplify the diagrams.

In ideal randomized controlled trials with perfect adherence to assignment and no loss to follow-up, we can expect that the exposed and the unexposed groups are completely comparable and that there is no confounding [30]. Thus, Equations 1 to 3 all hold in this situation. (Strictly speaking, there is no confounding “in expectation” although there is a possibility of “realized” confounding [5, 6, 24, 25, 30]. This phenomenon has been also referred to as “random confounding” [31, 32].) From the perspective of the sufficient-cause model, we can also expect that, in ideal randomized controlled trials, all of the three background factors are distributed comparably between the exposed and the unexposed groups. Thus, there is no association between each of the background factors and the exposure, and we can expect that covariate balance is achieved. Figure 1 describes this situation; there is no backdoor path from E to Y, implying there is no confounding in this case.

By contrast, in observational studies, the exposed and the unexposed groups are usually not comparable, which implies the presence of covariate imbalance. From the perspective of the sufficient-cause model, we would expect that the distributions of the three background factors are not comparable between the two groups. Let us tentatively suppose that each of the background factors is associated with the exposure either positively or negatively. In Figure 2, we show the associations or correlations using their (unknown or unmeasured) common causes U_1, U_2, and U_3. (Note that we here assume that none of the background factors is an intermediate step in the causal path between E and Y. We discuss a related issue in Appendix B.) There are three unblocked backdoor paths between E and Y in Figure 2, which conveniently describes a situation of covariate imbalance of C_1, C_2, and C_3 between the exposed and the unexposed groups.

Like the (partial) exchangeability conditions, the target population plays a key role when discussing covariate balance for no confounding. Recall that, when the exposed group is the target population, we use the actual unexposed group as a substitute of what would have occurred in the actual exposed group had they been unexposed. Given that sufficient cause 2 (i.e., C_2E) contains exposure as a component, it can never complete when the individual is unexposed. Therefore, when the exposed group is the target population, we do not have to consider the comparability of C_2 between the exposed and the unexposed groups. In other words, C_2 represents an irrelevant set of covariates for no confounding, and we need to consider the comparability of only C_1 and C_3 between the exposed and the unexposed groups. If we compare their joint distributions between the two groups, we can obtain a sufficient condition of covariate balance for no confounding as follows:

$$P\left[\left(C_1,C_3\right)=\left(c_1,c_3\right)\mid E=1\right]=P\left[\left(C_1,C_3\right)=\left(c_1,c_3\right)\mid E=0\right] \quad \left(c_1,c_3=0,1\right)$$
$$\Leftrightarrow \left(C_1,C_3\right)\coprod E. \qquad \text{[Eq. 4]}$$
$$\left(\Leftrightarrow \left\{\left(s_1+s_3\right)=\left(t_1+t_3\right)\right\}\wedge\left\{\left(s_2+s_4\right)=\left(t_2+t_4\right)\right\}\wedge\left\{\left(s_5+s_7\right)=\left(t_5+t_7\right)\right\}\right)$$

We refer to this joint independence as *partial covariate balance*. Note that Equation 4 is stronger than the partial exchangeability condition in Equation 1, that is, a sufficient and necessary condition for no confounding. This point can be readily shown by rewriting Equation 1 as: $\max\left(C_1,C_3\right)\coprod E$ or $\left(s_1+s_2+s_3+s_4+s_5+s_7\right)=\left(t_1+t_2+t_3+t_4+t_5+t_7\right)$. When $\left(C_1,C_3\right)=\left(1,1\right)$ in the first equation of Equation 4, the left-hand side represents a proportion of subjects who are at risk of sufficient causes 1 and 3 (i.e., C_1 and $C_3\overline{E}$, respectively) in the exposed group, whereas the right-hand side represents the corresponding proportion in the unexposed group. More strictly speaking, these subjects *potentially* become at risk of sufficient causes 1 and 3 during the follow-up period. Thus, when the first equation is met, the exposed and the unexposed groups are comparable in terms of susceptibility to sufficient causes 1 and 3. An analogous discussion applies when $\left(C_1,C_3\right)=\left(1,0\right),\left(0,1\right)$, or $\left(0,0\right)$.

Conversely, when the unexposed group is the target population, C_3 represents an irrelevant set of covariates for no confounding because sufficient cause 3 (i.e., $C_3\overline{E}$) can never complete when the individual is exposed. Thus, we need to consider the comparability of joint distributions of C_1 and C_2 between the exposed and the unexposed groups, which yields a sufficient condition of covariate balance for no confounding as:

$$\begin{aligned}
&P\left[\left(C_1,C_2\right)=\left(c_1,c_2\right)\mid E=1\right]=P\left[\left(C_1,C_2\right)=\left(c_1,c_2\right)\mid E=0\right]\quad\left(c_1,c_2=0,1\right)\\
&\Leftrightarrow\left(C_1,C_2\right)\coprod E, \qquad\qquad \text{[Eq. 5]}\\
&\left(\Leftrightarrow\left\{\left(s_1+s_2\right)=\left(t_1+t_2\right)\right\}\wedge\left\{\left(s_3+s_4\right)=\left(t_3+t_4\right)\right\}\wedge\left\{\left(s_5+s_6\right)=\left(t_5+t_6\right)\right\}\right)
\end{aligned}$$

which is stronger than the partial exchangeability condition in Equation 2, that is, a sufficient and necessary condition for no confounding. This point can be readily shown by rewriting Equation 2 as: $\max\left(C_1,C_2\right)\coprod E$ or $\left(s_1+s_2+s_3+s_4+s_5+s_6\right)=\left(t_1+t_2+t_3+t_4+t_5+t_6\right)$. We also refer to Equation 5 as *partial covariate balance*.

Finally, when the target is the total population, we need to consider the comparability of joint distributions of C_1 and C_2 between the total population and the exposed group, as well as the comparability of joint distributions of C_1 and C_3 between the total population and the unexposed group. This yields a sufficient condition of covariate balance for no confounding as:

$$\begin{aligned}
&\left\{P\left[\left(C_1,C_2\right)=\left(c_1,c_2\right)\right]=P\left[\left(C_1,C_2\right)=\left(c_1,c_2\right)\mid E=1\right]\quad\left(c_1,c_2=0,1\right)\right\}\\
&\wedge\left\{P\left[\left(C_1,C_3\right)=\left(c_1,c_3\right)\right]=P\left[\left(C_1,C_3\right)=\left(c_1,c_3\right)\mid E=0\right]\quad\left(c_1,c_3=0,1\right)\right\}\\
&\Leftrightarrow\left(C_1,C_k\right)\coprod E\quad\left(k=2,3\right), \qquad\qquad \text{[Eq. 6]}\\
&\left(\begin{array}{l}\Leftrightarrow\left[\left\{\left(s_1+s_2\right)=\left(t_1+t_2\right)\right\}\wedge\left\{\left(s_3+s_4\right)=\left(t_3+t_4\right)\right\}\wedge\left\{\left(s_5+s_6\right)=\left(t_5+t_6\right)\right\}\right]\\
\wedge\left[\left\{\left(s_1+s_3\right)=\left(t_1+t_3\right)\right\}\wedge\left\{\left(s_5+s_7\right)=\left(t_5+t_7\right)\right\}\right]\end{array}\right)
\end{aligned}$$

which is stronger than the exchangeability condition in Equation 3, that is, a sufficient and necessary condition for no confounding. This point can be readily shown by rewriting Equation 3 as: $\max\left(C_1,C_k\right)\coprod E\left(k=2,3\right)$ or $\left\{\left(s_1+s_2+s_3+s_4+s_5+s_6\right)=\left(t_1+t_2+t_3+t_4+t_5+t_6\right)\right\}\wedge\left\{\left(s_1+s_2+s_3+s_4+s_5+s_7\right)=\left(t_1+t_2+t_3+t_4+t_5+t_7\right)\right\}$. Equation 6 is simply a product of Equations 4 and 5, and we refer to Equation 6 as *covariate balance* in a more limited sense. We should note that complete comparability of risk status types between the exposed and the unexposed groups (i.e., $\left(C_1,C_2,C_3\right)\coprod E$ or $\left(s_1,s_2,s_3,s_4,s_5,s_6,s_7,s_8\right)=\left(t_1,t_2,t_3,t_4,t_5,t_6,t_7,t_8\right)$) is stronger than Equations 4 to 6. We refer to this complete comparability as *full covariate balance*, which is stronger than the full exchangeability condition. Equations 4 to 6 are neither stronger nor weaker than the full exchangeability condition.

The discussion above illustrates that, irrespective of the target population, covariate balance is a sufficient, but not a necessary, condition for no confounding. In other words, although confounding implies the presence of covariate imbalance, the presence of covariate imbalance does not necessarily induce confounding. In Table 2, we summarize the relationship between exchangeability and covariate balance for no confounding. See Appendix C for further remarks under the assumption of no preventive action (i.e., $C_3 = 0$ for all individuals). When discussing covariate balance, a set of covariates can be generally divided into covariates that have no causal coaction with an exposure (i.e., C_1) and covariates that have any causal coaction with an exposure (i.e., C_2 and C_3). Consequently, the target population plays a key role when discussing covariate balance for no confounding. An illustrative example from real data is provided in Online Appendix 1.

In Online Appendix 2, we show that one can obtain weaker sufficient conditions of covariate balance for no confounding by comparing distributions of the *number* of the relevant susceptible background factors between the exposed and the unexposed groups. Furthermore, in Online Appendix 3, we discuss alternative conditions of covariate balance by considering the comparability of *marginal* distributions of each of the relevant background factors between the exposed and the unexposed groups. See Online Appendix 4 for a more subtle conceptualization of covariate balance.

Table 2. The relationship between exchangeability and covariate balance for no confounding [a]

Target population	**Counterfactual model**	**Sufficient-cause model**	
	Exchangeability	**Exchangeability in terms of background factors**	**Covariate balance**
Exposed group	$Y_0 \coprod E$	$\Leftrightarrow \max(C_1, C_3) \coprod E$	$\Leftarrow (C_1, C_3) \coprod E$
Unexposed group	$Y_1 \coprod E$	$\Leftrightarrow \max(C_1, C_2) \coprod E$	$\Leftarrow (C_1, C_2) \coprod E$
Total population	$Y_e \coprod E \quad (e = 0,1)$	$\Leftrightarrow \max(C_1, C_k) \coprod E \quad (k = 2,3)$	$\Leftarrow (C_1, C_k) \coprod E \quad (k = 2,3)$
	$\Uparrow$	$\Uparrow$	$\Uparrow$
	$(Y_0, Y_1) \coprod E$	$\Leftrightarrow (\max(C_1, C_3), \max(C_1, C_2)) \coprod E$	$\Leftarrow (C_1, C_2, C_3) \coprod E$

[a] See Table 1 for notations.

Discussion

Within the sufficient cause model, we proposed to use each of the background factors in sufficient causes as representing a set of covariates of interest and discuss the presence of covariate balance by comparing joint distributions of the relevant background factors between the exposed and the unexposed groups. By considering the link between the sufficient-cause model and the counterfactual model, we illustrated that covariate balance is a sufficient, but not a necessary, condition for no confounding irrespective of the target population. This is consistent with the fact that the sufficient-cause model is a "finer" model than the counterfactual model. We incorporated sufficient causes within the DAG framework to graphically illustrate our conceptualization of covariate balance.

As noted above, the term "covariate" is used broadly including the term "confounder" [8], and these are often used interchangeably. A confounder was traditionally explained as a factor that has the following three necessary (but not sufficient or defining) characteristics: (a) it must be a risk factor for the outcome; (b) it must be associated with the exposure; and (c) it must not be an intermediate step in the causal path between the exposure and the outcome. These points are often explained in epidemiology and statistics textbooks using a simple diagram such as that shown in Figure 3, with some variations [33-38]. Because the background factors C_1, C_2, and C_3 may contain C in Figure 3, our conceptualization of covariate balance can be extended as a mechanistic representation of this traditional "definition" of a confounder in the sufficient-cause model. Note that the background factors appear to be truly mechanistic in that they causally produce the outcome. (A confounder that blocks a backdoor path from exposure to outcome, but does not cause the outcome is not included in the background factors. We focus on a factor that causally produces the outcome on the backdoor path.) As has been well addressed, however, this traditional "definition" is not a good definition of a confounder, which may lead to inappropriate adjustment for confounding [2-7]. See Hernán and Robins [27] for further discussion.

Recently, VanderWeele and Shpitser [39] proposed that, within the counterfactual framework, a confounder be defined as a pre-exposure covariate C for which there exists a set of other covariates X such that effect of the exposure on the outcome is unconfounded conditional on (X, C) but such that for no proper subset of (X, C) is the effect of the exposure on the outcome unconfounded given the subset. Equivalently, a confounder is defined as a "member of a minimally sufficient adjustment set", which they illustrated coheres with the following two properties; (i) if one were to control for all confounders, then it would suffice to control for confounding; and (ii) control for a confounder in some sense helps to eliminate or reduce confounding. As a consequence of this so-called intervention-based perspective, the presence or absence of confounders is, by definition, an issue only when there is confounding. Indeed, the definition of a confounder is different from the definition of confounding. By contrast, our conceptualization of covariate balance is based on a so-called mechanistic perspective to identify which covariates should be adjusted for to achieve comparability between the exposed and the unexposed groups, which results in no confounding.

Thus, our conceptualization of covariate balance does not necessarily correspond to confounding. When there is confounding, however, if one were to control for all relevant background factors to achieve their comparable joint distributions between the exposed and the unexposed groups, it would suffice to control for confounding. Furthermore, control for each of the relevant background factors helps to eliminate or reduce confounding. Therefore, when there is confounding, the concept of covariate imbalance coheres to the abovementioned two properties of a confounder, which shows that covariate imbalance in the sufficient-cause model is closely related to the counterfactual-based definition of a confounder. Despite their similarities, however, the concepts of covariate balance and confounder should be clearly distinguished.

We should note that our conceptualization of covariate balance is based on a representation of a set of (observed and unobserved) covariates rather than a single covariate although one generally refers to a particular component cause (or a risk factor) as a covariate. Our discussion is also limited to a simple situation in which one considers a binary exposure and a binary outcome.

Covariate balance between the exposed and the unexposed groups has been a key issue when inferring causality, and many researchers show characteristics of study participants for each group in their analysis. We proposed a mapping between covariate balance under the sufficient-cause model and exchangeability conditions in the counterfactual model, highlighting the facts that covariate balance is a stronger condition than no confounding and that the required covariate balance depends on the target population of interest. Our formalization of the notion of covariate balance will be useful in clarifying the meaning of confounding.

Appendix A: Covariate balance and the notions of confounding

Recent literature highlights the significance of the fact that the notion of confounding can be defined with respect to both marginal distributions of potential outcomes (i.e., confounding *in distribution*) and a specific effect measure (i.e., confounding *in measure*) [2, 24-26]. According to VanderWeele [24], confounding *in distribution* is defined as follows:

We say that there is no confounding *in distribution* of the effect of E on Y conditional on C if $P\left(Y_e \mid C=c\right)=P\left(Y \mid E=e, C=c\right)$ for all e, c.

In the main text, we show sufficient and necessary conditions for no confounding when C is an empty set without loss of generality. We denote measures of interest by $\mu(\phi_1,\phi_0)$, which is a contrast of population parameters. When defining population causal effects, ϕ_e is a population parameter for the distributions of potential outcomes Y_e if E had been set to e for all in the target [40]. Then, according to VanderWeele [24], confounding *in measure* is defined as follows:

We say that there is no confounding *in measure* μ of the effect of E on Y conditional on C if $\mu\left(E\left(Y_1 \mid C=c\right), E\left(Y_0 \mid C=c\right)\right)=\mu\left(E\left(Y \mid E=1, C=c\right), E\left(Y \mid E=0, C=c\right)\right)$ for all c.

Because no confounding *in distribution* is a sufficient condition for no confounding *in measure* [24-26], covariate balance is a stronger condition than no confounding irrespective of whether we use either the notions of confounding *in distribution* or confounding *in measure*. Further, this relation applies in both the notions of confounding “in expectation” and “realized” confounding [25]. In line with this, we may distinguish the concepts of covariate balance “in expectation” and “realized” covariate balance. A further discussion about the notions of confounding is provided elsewhere [25].

Appendix B: Covariate balance when considering mediation

In this Appendix, we show that, when considering mediation, a distribution of an intermediate factor is a distinct issue from the conditions of covariate balance for no confounding. In so doing, we extend our discussion to illustrate the concept of covariate balance when considering mediation in the sufficient-cause model [41-43]. (Note that we do not discuss time-varying covariates in this article.) We consider a situation in which some of the effects of exposure E on outcome Y are thought to be mediated by a binary intermediate M. Then, mediation is conceptualized as a two-stage process, including both the M-stage (processes that lead to the formation of the mediator) and the Y-stage (processes that lead to the formation of the outcome). In the M-stage, we could enumerate three different types of sufficient causes for M along with certain background factors A_k: A_1, A_2E, and $A_3\overline{E}$. We let A_k denote a set of all components or factors, other than the presence of E and $\overline{E}$, that may be required for a particular mechanism to operate. In the Y-stage, we could enumerate nine (i.e., 3^2) different types of sufficient causes for Y along with certain background factors B_k: B_1, B_2E, B_3M,

$B_4\overline{E}$, $B_5\overline{M}$, B_6EM, $B_7E\overline{M}$, $B_8\overline{E}M$, and $B_9\overline{E}\overline{M}$. Here, B_k denotes a set of all components or factors, other than the presence of E, $\overline{E}$, M, and $\overline{M}$, that may be required for a particular mechanism to operate. Accordingly, we can enumerate eight (i.e., 2^3) and 512 (i.e., 2^9) patterns of possible risk status for sufficient causes in the M-stage and the Y-stage, respectively, and combining these yields 4,096 (i.e., 8 × 512) patterns of possible MY risk status [43]. Recall that we use each of the background factors (i.e., A_k and B_k) as representing a set of covariates of interest. Figure B.1 shows a causal diagram depicting mediation [42]. Note that the intermediate M is not included in the covariates of interest. Its distribution is a distinct issue from the concept of covariate balance.

As illustrated in the main text, the target population plays a key role when discussing the concept of covariate balance. To simplify the discussion, let us tentatively assume that there are only two M-mediated paths, i.e., $E \rightarrow A_2E \rightarrow M \rightarrow B_3M \rightarrow Y$ and $E \rightarrow A_2E \rightarrow M \rightarrow B_6EM \rightarrow Y$. In this case, the unexposed individuals can have neither M nor Y, and trivially, none of the exposed individuals would have had M or Y if there had been interventions to set $E = 0$. Thus, when the target population is the exposed group, there is no confounding, even if either of the abovementioned two M-mediated paths exists differentially between the exposed and the unexposed groups due to covariate imbalance. (Graphically, this situation can be depicted by adding an unmeasured common cause between E and A_2, E and B_3, and E and B_6, respectively, in Figure B1.) However, the exposed group is not an ideal representative of what would have occurred had the total population or the unexposed group been exposed. Thus, when the target population is either the total population or the unexposed group, the presence of covariate imbalance can generally lead to confounding.

Appendix C: Further remarks on covariate balance for no confounding

The relationship between exchangeability and covariate balance applies in general; for example, let us consider a situation in which one can assume no preventive action (i.e., $C_3 = 0$ for all individuals) when the target is the total population. Recall that, under the assumption of no preventive action, the individuals of risk status types 1, 3, 5, and 7 are excluded (Table 1). Recall also that no preventive action implies positive monotonicity since risk status type 7 corresponds to response type 3. Then, the exchangeability condition in Equation 3 (i.e., $Y_e \amalg E\left(e=0,1\right)$ or $\max\left(C_1,C_k\right)\amalg E\left(k=2,3\right)$) becomes:

$$\left\{C_1 \amalg E\right\} \wedge \max\left(C_1,C_2\right) \amalg E. \qquad \text{[Eq. C.1]}$$

$$\begin{pmatrix} \Leftrightarrow \left\{p_1=q_1\right\} \wedge \left\{p_2=q_2\right\} \quad \left(\because p_3=q_3=0\right) \\ \Leftrightarrow \left\{\left(s_2+s_4\right)=\left(t_2+t_4\right)\right\} \wedge \left\{s_6=t_6\right\} \quad \left(\because s_1=s_3=s_5=t_1=t_3=t_5=0\right) \end{pmatrix}$$

And, the condition of covariate balance in Equation 6 (i.e., $\left(C_1,C_k\right)\amalg E\left(k=2,3\right)$) becomes:

$$\left(C_1, C_2\right) \coprod E, \qquad \text{[Eq. C.2]}$$

$$\left(\Leftrightarrow \left(s_2 = t_2\right) \wedge \left(s_4 = t_4\right) \wedge \left(s_6 = t_6\right) \quad \left(\because s_1 = s_3 = s_5 = s_7 = t_1 = t_3 = t_5 = t_7 = 0\right)\right)$$

which is stronger than Equation C.1. Thus, when the total population is the target population, even in the presence of covariate imbalance, we may observe no confounding. This exemplifies that it would be of significance to understand the subtle difference between exchangeability and covariate balance, based on the link between the counterfactual model and the sufficient-cause model.

Although exchangeability and covariate balance are in general subtly different, they become consistent in a particular situation. As an example, let us consider the exposed group as the target population under the assumption of no preventive action. In this case, Equations 1 and 4 reduce to $C_1 \coprod E$ or $\left(s_2 + s_4\right) = \left(t_2 + t_4\right)$. This coincidence, however, would be rare in practice.

Acknowledgements

The authors thank Tyler J. VanderWeele and Ian Shrier for their helpful comments on the earlier version of this manuscript. The authors also thank Toshiharu Mitsuhashi for helpful discussions.

Funding

This work was supported by Japan Society for the Promotion of Science (KAKENHI grant numbers JP26870383, JP15K08776, and JP17K17898); Foundation for Total Health Promotion; and The Okayama Medical Foundation.

Competing interests

None declared.

References

1. Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol 1986;15(3):413-9.
2. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. Stat Sci 1999;14(1):29-46.
3. Greenland S, Morgenstern H. Confounding in health research. Annu Rev Public Health 2001;22:189-212.
4. Greenland S, Rothman KJ, Lash TL. Measures of effect and measures of association. In: Modern Epidemiology, Rothman KJ, Greenland S, Lash TL, editors. Lippincott Williams & Wilkins: Philadelphia, PA, 2008, p. 51-70.
5. Greenland S. Confounding. In: Encyclopedia of Epidemiology, Boslaugh S, editor. Sage Publications: Thousand Oaks, CA, 2008, p. 227-32.
6. Greenland S, Robins JM. Identifiability, exchangeability and confounding revisited. Epidemiol Perspect Innov 2009;6:4. doi:10.1186/1742-5573-6-4.
7. Pearl J. Causality: Models, Reasoning, and Inference. 2nd ed. Cambridge University Press: New York, NY; 2009.
8. Porta MS, editor. A Dictionary of Epidemiology. 6th ed. New York, NY: Oxford University Press; 2014.
9. Kaufman JS, Poole C. Looking back on "causal thinking in the health sciences". Annu Rev Public Health 2000;21:101-19.
10. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu Rev Public Health 2000;21:121-45.
11. Rothman KJ. Causes. Am J Epidemiol 1976;104(6):587-92.
12. Greenland S, Poole C. Invariants and noninvariants in the concept of interdependent effects. Scand J Work Environ Health 1988;14(2):125-9.
13. Flanders WD. On the relationship of sufficient component cause models with potential outcome (counterfactual) models. Eur J Epidemiol 2006;21(12):847-53.
14. VanderWeele TJ, Hernán MA. From counterfactuals to sufficient component causes and vice versa. Eur J Epidemiol 2006;21(12):855-8.
15. Suzuki E, Yamamoto E, Tsuda T. On the link between sufficient-cause model and potential-outcome model. Epidemiology 2011;22(1):131-2.
16. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? Epidemiology 2009;20(1):3-5.
17. VanderWeele TJ. Concerning the consistency assumption in causal inference. Epidemiology 2009;20(6):880-3.
18. Greenland S, Lash TL, Rothman KJ. Concepts of interaction. In: Modern Epidemiology, Rothman KJ, Greenland S, Lash TL, editors. Lippincott Williams & Wilkins: Philadelphia, PA, 2008, p. 71-83.

19. VanderWeele TJ. Attributable fractions for sufficient cause interactions. Int J Biostat 2010;6(2):5. doi:10.2202/1557-4679.1202.

20. Suzuki E, Yamamoto E, Tsuda T. On the relations between excess fraction, attributable fraction, and etiologic fraction. Am J Epidemiol 2012;175(6):567-75.

21. Maldonado G, Greenland S. Estimating causal effects. Int J Epidemiol 2002;31(2):422-9.

22. Flanders WD, Johnson CY, Howards PP, Greenland S. Dependence of confounding on the target population: a modification of causal graphs to account for co-action. Ann Epidemiol 2011;21(9):698-705.

23. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A counterfactual approach to bias and effect modification in terms of response types. BMC Med Res Methodol 2013;13:101. doi:10.1186/1471-2288-13-101.

24. VanderWeele TJ. Confounding and effect modification: distribution and measure. Epidemiol Method 2012;1(1):55-82. doi:10.1515/2161-962X.1004.

25. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A typology of four notions of confounding in epidemiology. J Epidemiol 2017;27(2):49-55.

26. Suzuki E, Yamamoto E. Further refinements to the organizational schema for causal effects. Epidemiology 2014;25(4):618-9.

27. Hernán MA, Robins JM. Causal Inference. Chapman & Hall/CRC, forthcoming: Boca Raton, FL; 2017.

28. VanderWeele TJ, Robins JM. Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. Am J Epidemiol 2007;166(9):1096-104.

29. VanderWeele TJ, Robins JM. Minimal sufficient causation and directed acyclic graphs. Ann Stat 2009;37(3):1437-65.

30. Greenland S. Randomization, statistics, and causal inference. Epidemiology 1990;1(6):421-9.

31. Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness. Eur J Epidemiol 2015;30(10):1101-10.

32. Suzuki E, Tsuda T, Mitsuhashi T, Mansournia MA, Yamamoto E. Errors in causal inference: an organizational schema for systematic error and random error. Ann Epidemiol 2016;26(11):788-93.

33. Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring III JR. Medical Epidemiology. 4th ed. Lange Medical Books/McGraw-Hill: New York, NY; 2005.

34. Gordis L. Epidemiology. 4th ed. Elsevier/Saunders: Philadelphia, PA; 2009.

35. Bhopal RS. Concepts of Epidemiology: Integrating the Ideas, Theories, Principles, and Methods of Epidemiology. 2nd ed. Oxford University Press: New York, NY; 2008.

36. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 3rd ed. Jones & Bartlett Learning: Burlington, MA; 2012.

37. Jewell NP. Statistics for Epidemiology. Chapman & Hall/CRC: Boca Raton, FL; 2004.

38. Newman SC. Biostatistical Methods in Epidemiology. John Wiley & Sons: New York, NY; 2001.
39. VanderWeele TJ, Shpitser I. On the definition of a confounder. Ann Stat 2013;41(1):196-220.
40. Flanders WD, Klein M. A general, multivariate definition of causal effects in epidemiology. Epidemiology 2015;26(4):481-9.
41. Hafeman DM. A sufficient cause based approach to the assessment of mediation. Eur J Epidemiol 2008;23(11):711-21.
42. VanderWeele TJ. Mediation and mechanism. Eur J Epidemiol 2009;24(5):217-24.
43. Suzuki E, Yamamoto E, Tsuda T. Identification of operating mediation and mechanism in the sufficient-component cause framework. Eur J Epidemiol 2011;26(5):347-57.

Figure captions

Fig. 1 Directed acyclic graph incorporating sufficient causes

We consider a binary exposure *E* and a binary outcome *Y*. We consider three different types of sufficient causes for outcome *Y* along with certain binary background factors as follows: C_1, C_2E, and $C_3\overline{E}$, where we let $\overline{E}$ denote the complement of *E*.

Fig. 2 Directed acyclic graph depicting covariate imbalance in the sufficient-cause model

U_1, U_2, and U_3 denote (unknown or unmeasured) common causes between the background factors and the exposure.

Fig. 3 Typical diagram showing confounding/confounder

E, *Y*, and *C* denote exposure, outcome, and confounder, respectively. As mentioned in the text, the traditional "definition" of a confounder may lead to inappropriate adjustment for confounding.

Fig. B1 Directed acyclic graph depicting mediation in the sufficient-cause model

We consider a binary exposure *E*, a binary intermediate *M*, and a binary outcome *Y*. For details, see the text.

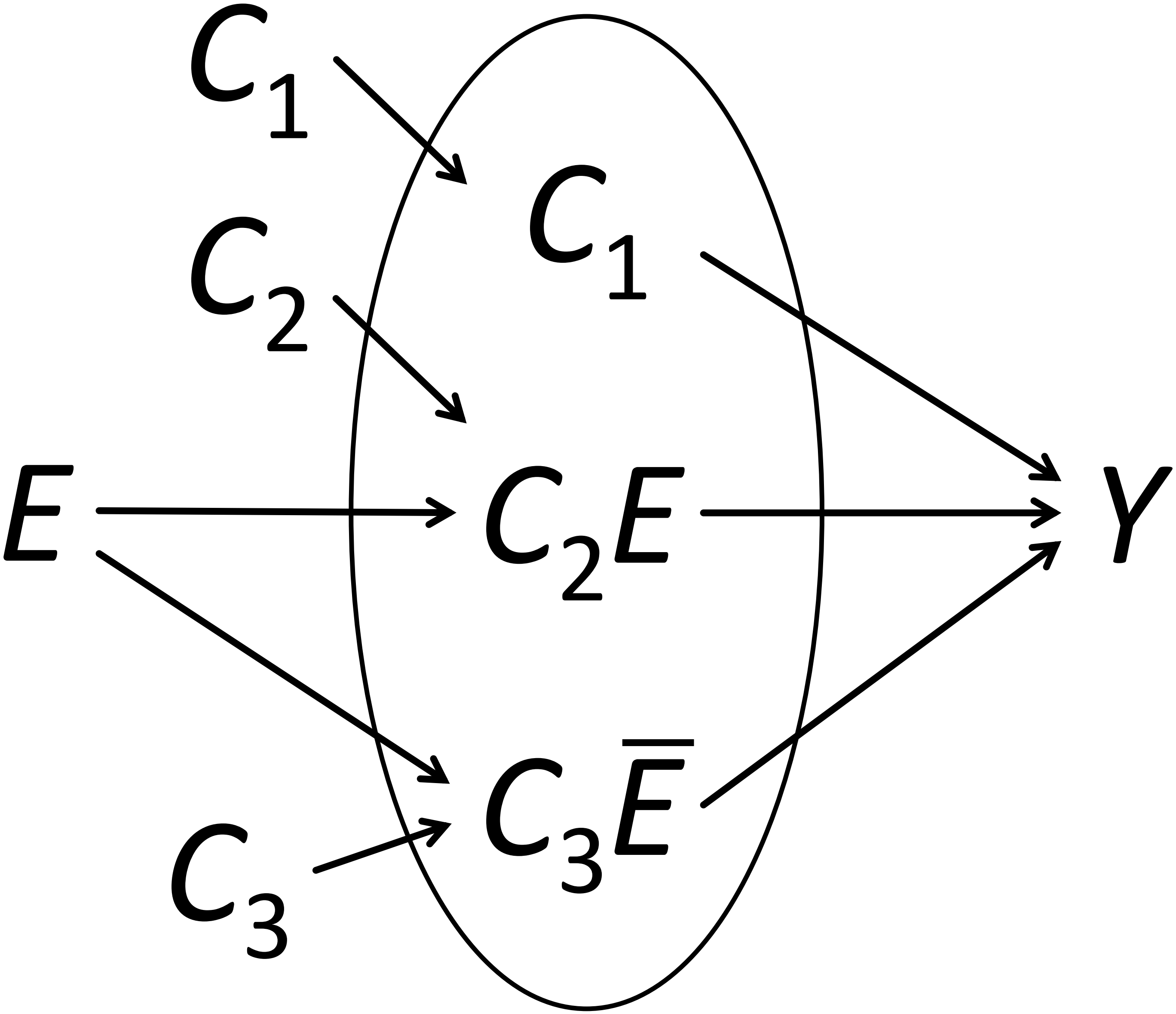

Fig. 1

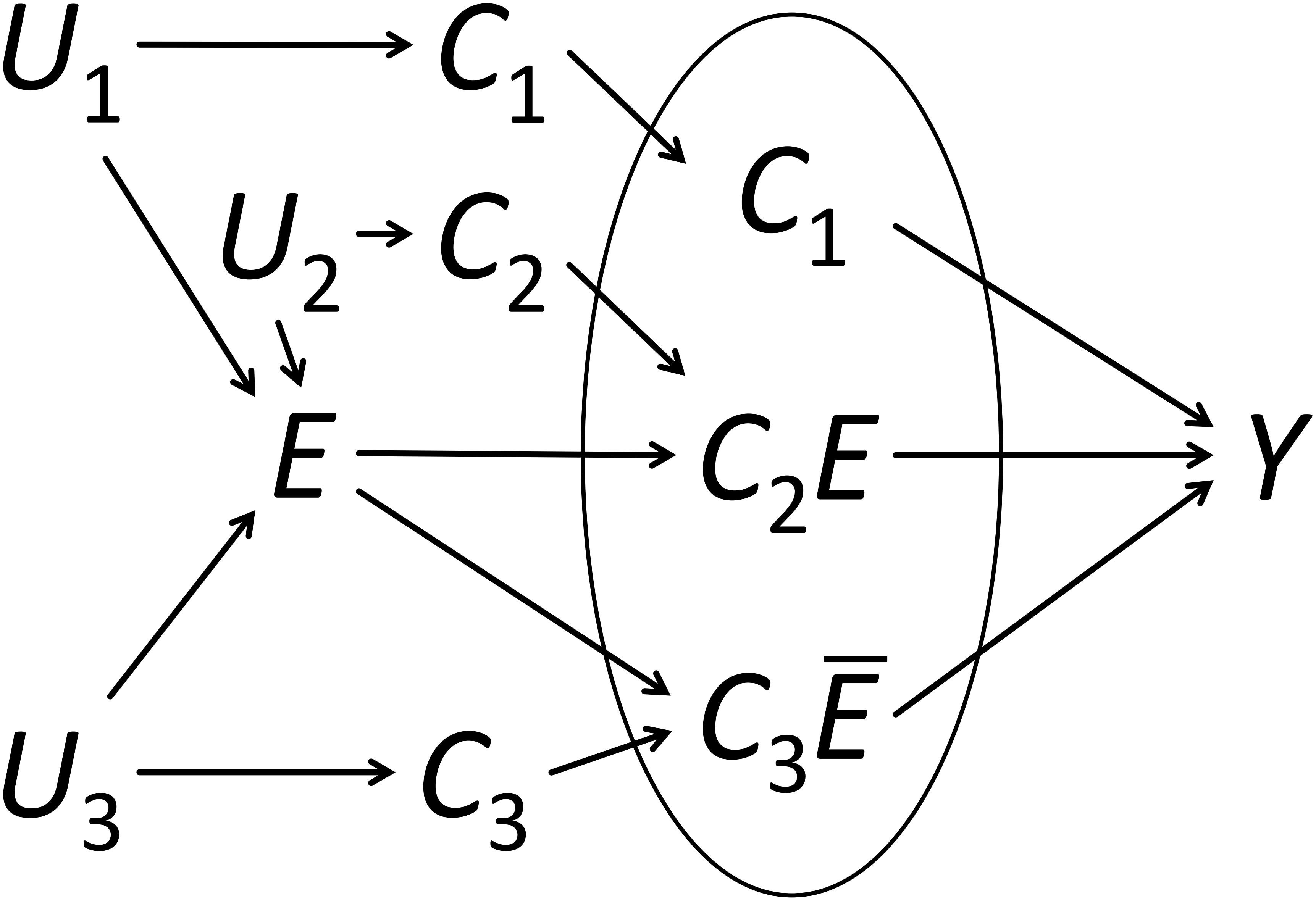

Fig. 2

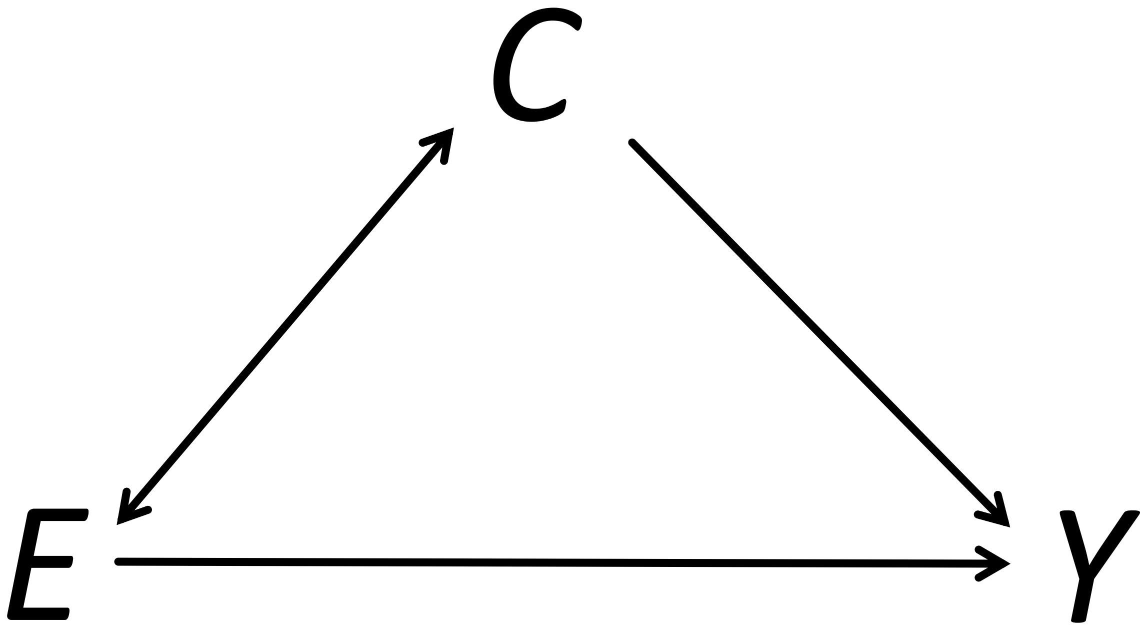

Fig. 3

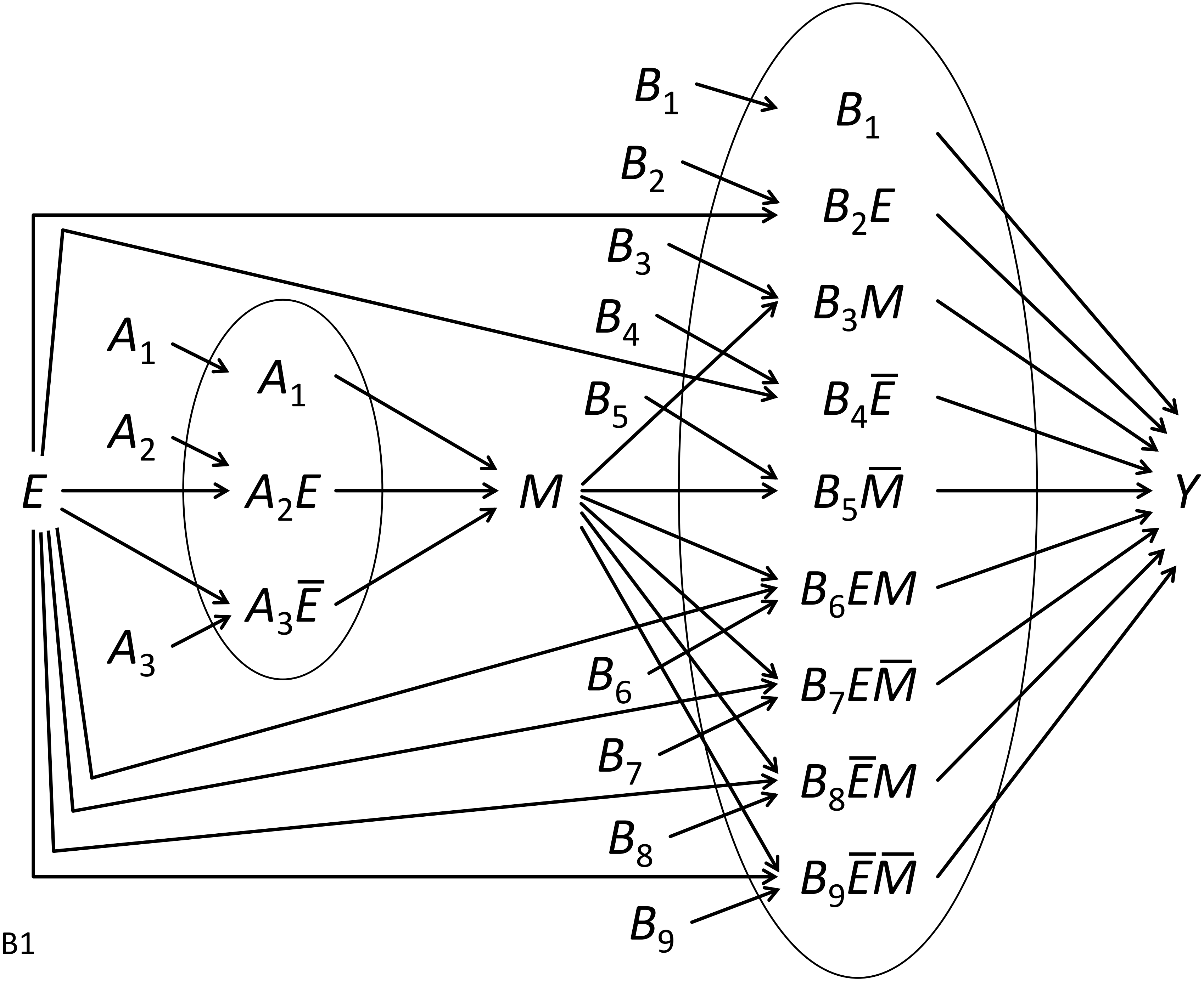
B_1
B_1
B_2
B_2E
B_3
B_3M
B_4
A_1
A_1
$B_4\overline{E}$
B_5
A_2
E
A_2E
M
$B_5\overline{M}$
Y
B_6EM
$A_3\overline{E}$
A_3
B_6
$B_7E\overline{M}$
B_7
$B_8\overline{E}M$
B_8
$B_9\overline{E}\overline{M}$
B_9

Fig. B1

Covariate balance for no confounding in the sufficient-cause model

Online Appendix 1: An example of the Western Collaborative Group Study

To illustrate the methods, we discuss the relationship between the concepts of confounding and covariate balance using real data from the Western Collaborative Group Study (WCGS) [1]. This data has often been used in epidemiology textbooks [2, 3], and we describe it briefly here. The WCGS data consist of 3,154 middle-aged men (aged 39 to 59) recruited from ten California companies during the years 1960–1961. The exposure of interest was behavior type (type A vs. type B). Type A behavior is characterized by aggressiveness and competitiveness, whereas type B behavior is characterized by a relaxed, noncompetitive, less hurried personality. A total of 1,589 type-A and 1,565 type-B individuals were identified using tape-recorded interviews. The outcome of interest was the occurrence of coronary heart disease (CHD), which was determined by expert diagnosis. During the mean of 8.5 years of follow-up, CHD occurred in 257 subjects. Online Table 1 summarizes the relationship between behavior type and CHD. The risk of CHD among type-A individuals (i.e., the exposed group) was 0.112, while the risk of CHD among type-B individuals (i.e., the unexposed group) was 0.051.

Online Table 2 shows a possible distribution of response types and risk status types in the WCGS data. Various risk factors were also measured in the study (e.g., smoking, blood pressure, cholesterol level), and each of the three background factors, C_1, C_2, and C_3, represents a set of measured and unmeasured covariates. Consistent with the results in Online Table 1, the incidence proportion in the exposed group, $\left(p_1+p_2\right)$, is 0.112 and the incidence proportion in the unexposed group, $\left(q_1+q_3\right)$, is 0.051. The associational risk difference is calculated as: $\left(p_1+p_2\right)-\left(q_1+q_3\right)=0.061$. Equations 1 to 3 are not met, and there is confounding irrespective of the target population, which is likely in most observational studies. The causal risk differences in the exposed group, the unexposed group, and the total population are 0.017, -0.004, and 0.007, respectively. Furthermore, Equations 4 to 6 are not met, and there is covariate imbalance irrespective of the target population. Recall that confounding implies the presence of covariate imbalance. For example, a proportion of subjects who are at risk of sufficient causes 1 and 3 in the exposed group is calculated as: $s_1+s_3=0.020$, and the corresponding proportion in the unexposed group is calculated as: $t_1+t_3=0.010$. Once covariate balance is achieved for a specific target population, we can obtain unconfounded estimates for the population.

As another possible example, suppose that proportions of response types 3 and 4 (or risk status types 7 and 8) in the exposed group are 0.006 and 0.882, respectively, in Online Table 2 (i.e., $p_3=s_7=0.006$ and $p_4=s_8=0.882$). Then, Equation 1 is met, and there is no confounding when the exposed group is the target population. The causal risk difference in the exposed group is now calculated as: $\left(p_1+p_2\right)-\left(p_1+p_3\right)=0.061$, which is equal to the associational risk difference. However, Equation 4 is not met, and there is covariate imbalance. Thus, even in the presence of covariate imbalance, we may observe no confounding. When the target population is the unexposed group or the total population, there is confounding and covariate imbalance; Equations 2, 3, 5, and 6 are not met.

A deeper understanding about etiology helps researchers to achieve covariate balance to obtain unconfounded estimates of causal parameters.

Online Appendix 2: Weaker sufficient conditions of covariate balance for no confounding

In the main text, we show sufficient conditions of covariate balance for no confounding by comparing joint distributions of relevant background factors between the exposed and the unexposed groups (see Equations 4 to 6). In this Online Appendix, we show that one can obtain weaker sufficient conditions of covariate balance for no confounding by comparing distributions of the *number* of the relevant susceptible background factors between the exposed and the unexposed groups.

Recall that, when the exposed group is the target population, we do not have to consider the comparability of C_2 between the exposed and the unexposed groups. In other words, C_2 represents an irrelevant set of covariates when considering covariate balance for no confounding, and we need to consider the comparability of only C_1 and C_3 between the exposed and the unexposed groups. If we compare distributions of the number of relevant susceptible background factors between the two groups, we can obtain a sufficient condition of covariate balance for no confounding as follows:

$$\begin{aligned}&P\left[C_1+C_3=c \mid E=1\right]=P\left[C_1+C_3=c \mid E=0\right] \quad \left(c=0,1,2\right)\\ &\Leftrightarrow \left(C_1+C_3\right)\amalg E, \qquad\qquad \text{[Eq. A1]}\\ &\left(\Leftrightarrow \left\{\left(s_1+s_3\right)=\left(t_1+t_3\right)\right\}\wedge\left\{\left(s_2+s_4+s_5+s_7\right)=\left(t_2+t_4+t_5+t_7\right)\right\}\right)\end{aligned}$$

which is weaker than the partial covariate balance in Equation 4 (i.e., $(C_1,C_3)\amalg E$). When $C_1+C_3=1$ in the first equation of Equation A1, the left-hand side represents a proportion of subjects who are at risk of one of the relevant background factors (i.e., sufficient causes 1 and 3) in the exposed group, whereas the right-hand side represents the corresponding proportion in the unexposed group. In other words, we do not distinguish those who are at risk of only sufficient cause 1 from those who are at risk of only sufficient cause 3 in Equation A1, although we distinguish them in Equation 4. Thus, Equation A1 is weaker than Equation 4. Note, however, that Equation A1 is stronger than the partial exchangeability condition in Equation 1, that is, a sufficient and necessary condition for no confounding (i.e., $Y_0\amalg E$ or $\max(C_1,C_3)\amalg E$).

Conversely, when the unexposed group is the target population, C_3 represents an irrelevant set of covariates for no confounding because sufficient cause 3 (i.e., $C_3\overline{E}$) can never complete when the individual is exposed. Thus, we need to consider the comparability of only C_1 and C_2 between the exposed and the unexposed groups. If we compare distributions of the number of relevant susceptible background factors between the two groups, we can obtain a sufficient condition of covariate balance for no confounding as follows:

$$\begin{aligned}&P\left[C_1+C_2=c\mid E=1\right]=P\left[C_1+C_2=c\mid E=0\right]\quad\left(c=0,1,2\right)\\&\Leftrightarrow\left(C_1+C_2\right)\coprod E, \qquad\qquad\qquad\qquad \text{[Eq. A2]}\\&\left(\Leftrightarrow\left\{\left(s_1+s_2\right)=\left(t_1+t_2\right)\right\}\wedge\left\{\left(s_3+s_4+s_5+s_6\right)=\left(t_3+t_4+t_5+t_6\right)\right\}\right)\end{aligned}$$

which is weaker than the partial covariate balance in Equation 5 (i.e., $(C_1,C_2)\coprod E$). Note that Equation A2 is stronger than the partial exchangeability condition in Equation 2, that is, a sufficient and necessary condition for no confounding (i.e., $Y_1\coprod E$ or $\max(C_1,C_2)\coprod E$).

Finally, when the target is the total population, we need to consider distributions of the number of susceptible background factors among sufficient causes 1 and 2 between the total population and the exposed group, as well as distributions of the number of susceptible background factors among sufficient causes 1 and 3 between the total population and the unexposed group. This yields a sufficient condition of covariate balance for no confounding as:

$$\begin{aligned}&\left\{P\left[C_1+C_2=c\right]=P\left[C_1+C_2=c\mid E=1\right]\quad\left(c=0,1,2\right)\right\}\\&\wedge\left\{P\left[C_1+C_3=c\right]=P\left[C_1+C_3=c\mid E=0\right]\quad\left(c=0,1,2\right)\right\}\\&\Leftrightarrow\left(C_1+C_k\right)\coprod E\quad\left(k=2,3\right), \qquad\qquad\qquad \text{[Eq. A3]}\\&\left(\begin{aligned}&\Leftrightarrow\left[\left\{\left(s_1+s_2\right)=\left(t_1+t_2\right)\right\}\wedge\left\{\left(s_3+s_4+s_5+s_6\right)=\left(t_3+t_4+t_5+t_6\right)\right\}\right]\\&\wedge\left[\left\{\left(s_1+s_3\right)=\left(t_1+t_3\right)\right\}\wedge\left\{\left(s_2+s_4+s_5+s_7\right)=\left(t_2+t_4+t_5+t_7\right)\right\}\right]\end{aligned}\right)\end{aligned}$$

which is weaker than the covariate balance in Equation 6 (i.e., $(C_1,C_k)\coprod E\;(k=2,3)$). Note that Equation A3, which is simply a product of Equations A1 and A2, is stronger than the exchangeability condition in Equation 3, that is, a sufficient and necessary condition for no confounding (i.e., $Y_e\coprod E\;(e=0,1)$ or $\max(C_1,C_k)\coprod E\;(k=2,3)$).

In the main text, we refer to the complete comparability of risk status types between the exposed and the unexposed groups (i.e., $(C_1,C_2,C_3)\coprod E$ or $(s_1,s_2,s_3,s_4,s_5,s_6,s_7,s_8)=(t_1,t_2,t_3,t_4,t_5,t_6,t_7,t_8)$) as full covariate balance. If we compare distributions of the number of relevant susceptible background factors between the exposed and the unexposed groups, a corresponding condition can be given as $\left((C_1+C_3),(C_1+C_2)\right)\coprod E$, which is stronger than Equations A1 to A3. Note that, in the corresponding condition, we do not distinguish those who are at risk of only sufficient cause 1 (i.e., $(C_1,C_2,C_3)=(1,0,0)$ or risk status type 4) from those who are at risk of only sufficient causes 2 and 3 (i.e., $(C_1,C_2,C_3)=(0,1,1)$ or risk status type 5). Therefore, the corresponding condition can be written using the proportions of risk status types as: $(s_1,s_2,s_3,(s_4+s_5),s_6,s_7,s_8)=(t_1,t_2,t_3,(t_4+t_5),t_6,t_7,t_8)$, which shows that it is weaker than the full covariate balance. Note, however, that the corresponding condition is stronger than the full exchangeability condition. Equations A1 to A3 are neither stronger nor weaker than the full exchangeability condition.

We summarize the relationship between exchangeability, the weaker conditions of covariate balance, and covariate balance in Online Table 3.

Online Appendix 3: Alternative conditions of marginal covariate balance

In the main text, we show conditions of covariate balance by comparing joint distributions of relevant background factors between the exposed and the unexposed groups (see Equations 4 to 6). Alternatively, one may compare *marginal* distributions of each of the relevant background factors between the exposed and the unexposed groups. Here we show alternative conditions of marginal covariate balance and their implications.

Recall that, when the exposed group is the target population, we do not have to consider the comparability of C_2 between the exposed and the unexposed groups. In other words, C_2 represents an irrelevant set of covariates when considering covariate balance. Thus, we need to consider the comparability of only C_1 and C_3 between the exposed and the unexposed groups. If we compare their marginal distributions between the two groups, we can obtain a condition of covariate balance as:

$$\begin{aligned}
&\left\{P\left[C_1 = c_1 \mid E = 1\right] = P\left[C_1 = c_1 \mid E = 0\right]\right\} \wedge \left\{P\left[C_3 = c_3 \mid E = 1\right] = P\left[C_3 = c_3 \mid E = 0\right]\right\} \quad \left(c_1, c_3 = 0,1\right)\\
&\Leftrightarrow C_k \coprod E \quad \left(k = 1,3\right), \qquad\qquad \text{[Eq. A4]}\\
&\left(\Leftrightarrow \left\{\left(s_1 + s_2 + s_3 + s_4\right) = \left(t_1 + t_2 + t_3 + t_4\right)\right\} \wedge \left\{\left(s_1 + s_3 + s_5 + s_7\right) = \left(t_1 + t_3 + t_5 + t_7\right)\right\}\right)
\end{aligned}$$

which is weaker than the partial covariate balance in Equation 4 (i.e., $(C_1, C_3) \coprod E$). Note, in the first equation of Equation A4, the left-hand side represents a proportion of subjects who are at risk of sufficient cause 1 (i.e., C_1) in the exposed group, whereas the right-hand side represents the corresponding proportion in the unexposed group. More strictly speaking, these subjects *potentially* become at risk of sufficient cause 1 during the follow-up period. Thus, when the first equation is met, the exposed and the unexposed groups are comparable in terms of susceptibility to sufficient cause 1. An analogous discussion applies to the second equation with regard to sufficient cause 3 (i.e., $C_3 \overline{E}$).

Conversely, when the unexposed group is the target population, C_3 represents an irrelevant set of covariates when considering covariate balance, and we need to consider the comparability of C_1 and C_2 between the exposed and the unexposed groups. If we compare their marginal distributions between the two groups, we can obtain a condition of covariate balance as:

$$\begin{aligned}
&\left\{P\left[C_1 = c_1 \mid E = 1\right] = P\left[C_1 = c_1 \mid E = 0\right]\right\} \wedge \left\{P\left[C_2 = c_2 \mid E = 1\right] = P\left[C_2 = c_2 \mid E = 0\right]\right\} \quad \left(c_1, c_2 = 0,1\right)\\
&\Leftrightarrow C_k \coprod E \quad \left(k = 1,2\right), \qquad\qquad \text{[Eq. A5]}\\
&\left(\Leftrightarrow \left\{\left(s_1 + s_2 + s_3 + s_4\right) = \left(t_1 + t_2 + t_3 + t_4\right)\right\} \wedge \left\{\left(s_1 + s_2 + s_5 + s_6\right) = \left(t_1 + t_2 + t_5 + t_6\right)\right\}\right)
\end{aligned}$$

which is weaker than the partial covariate balance in Equation 5 (i.e., $(C_1, C_2) \coprod E$).

Finally, when the target is the total population, we need to consider the comparability of C_1 and C_2 between the total population and the exposed group, as well as the comparability of C_1 and C_3 between the total population and the unexposed group. If we compare their marginal distributions, we can obtain a condition of covariate balance as:

$$\begin{aligned}
&\left[\left\{P\left[C_1=c_1\right]=P\left[C_1=c_1 \mid E=1\right]\right\}\wedge\left\{P\left[C_2=c_2\right]=P\left[C_2=c_2 \mid E=1\right]\right\}\right]\\
&\wedge\left[\left\{P\left[C_1=c_1\right]=P\left[C_1=c_1 \mid E=0\right]\right\}\wedge\left\{P\left[C_3=c_3\right]=P\left[C_3=c_3 \mid E=0\right]\right\}\right] \quad \left(c_1,c_2,c_3=0,1\right)\\
&\Leftrightarrow C_k \amalg E \quad \left(k=1,2,3\right), \qquad\qquad\qquad\qquad \text{[Eq. A6]}\\
&\left(\begin{aligned}&\Leftrightarrow\left\{\left(s_1+s_2+s_3+s_4\right)=\left(t_1+t_2+t_3+t_4\right)\right\}\wedge\left\{\left(s_1+s_2+s_5+s_6\right)=\left(t_1+t_2+t_5+t_6\right)\right\}\\ &\wedge\left\{\left(s_1+s_3+s_5+s_7\right)=\left(t_1+t_3+t_5+t_7\right)\right\}\end{aligned}\right)
\end{aligned}$$

which is weaker than the covariate balance in Equation 6 (i.e., $(C_1, C_k) \amalg E \; (k=2,3)$).

In the main text, we show that covariate balance is a sufficient, but not a necessary, condition for no confounding, irrespective of the target population. In other words, although confounding implies the presence of covariate imbalance, the presence of covariate imbalance does not necessarily induce confounding. If one uses the alternative conditions of marginal covariate balance, however, covariate balance is neither a necessary condition nor a sufficient condition for no confounding, and vice versa. To illustrate, we use two numerical examples below.

First, consistent with the findings in the main text, we may observe no confounding even in the presence of covariate imbalance if one uses the alternative conditions. Online Table 4 shows a numerical example when the prevalence of exposure in the total population is 0.2. Note that Equations 1 to 3 are all met, and there is no confounding irrespective of whether we use the exposed group, the unexposed group, or the total population as the target population. If we use risk difference as a measure of interest, the associational risk difference in Online Table 4 is calculated as: $\left(p_1+p_2\right)-\left(q_1+q_3\right)=-0.05$, which is equal to the causal risk difference in the exposed group, the unexposed group, and the total population. However, the marginal distributions of C_1, C_2, and C_3 are not comparable between the exposed and the unexposed groups. In the exposed group, proportions of subjects who are at risk of C_1, C_2, and C_3 are 0.40, 0.40, and 0.55, respectively, and the corresponding proportions in the unexposed group are 0.35, 0.45, and 0.50, respectively. Thus, Equations A4 to A6 are not met, which implies the presence of covariate imbalance irrespective of the target population. Therefore, if one uses the alternative conditions, no confounding is not a sufficient condition for covariate balance, which is consistent with the findings in the main text.

Meanwhile, if one uses the alternative conditions of marginal covariate balance, even if there is covariate balance, confounding may occur in some situations. Online Table 5 also shows a numerical example when the prevalence of exposure in the total population is 0.2. Note that the marginal distributions of C_1, C_2, and C_3 are comparable between the exposed and the unexposed groups in Online Table 5; proportions of subjects who are at risk of C_1, C_2, and C_3 are 0.35, 0.40, and 0.45, respectively, in the two groups. Thus, Equations A4 to A6 are all met, and there is covariate balance irrespective of whether we use the exposed group, the unexposed group, or the total population as the target population. However, Equations 1 to 3 are not met. Thus, contrary to intuitive expectations, there is confounding (or strictly speaking, confounding *in distribution*) irrespective of the target population. If we use risk difference as a measure of interest, the associational risk difference in Online Table 5 is calculated as: $\left(p_1+p_2\right)-\left(q_1+q_3\right)=-0.05$, whereas the causal risk differences in the exposed group, the unexposed group, and the total population are 0.00, -0.10, and -0.08, respectively. (Note that confounding *in distribution* does not necessarily imply confounding *in measure* when the target is the total population [4-6].) Therefore, if one uses the alternative conditions, covariate balance is not a sufficient condition for no confounding, which is inconsistent with the findings in the main text.

Although this may seem a subtle point, it is a consequence of comparing marginal distributions of each of the three background factors between the exposed and the unexposed groups. If instead we compare joint distributions of the relevant background factors between the exposed and the unexposed groups, this point does not occur. Equations 4 to 6 in the main text are not met in Online Table 5, which implies the presence of covariate imbalance irrespective of the target population. In other words, this numerical example shows that, covariate balance is a sufficient, but not a necessary, condition for no confounding, irrespective of the target population. This would fit with the frequently used argument that covariate balance is a key feature to control confounding. In conclusion, it would be appropriate to conceptualize the concept of covariate balance by comparing not marginal but joint distributions of relevant background factors between the exposed and the unexposed groups. In Online Table 3, we summarize the relationship between covariate balance and marginal covariate balance.

Online Appendix 4: A more subtle conceptualization of covariate balance

When considering the link between the sufficient-cause model and the counterfactual model in the main text, we focused on the correspondence between the four response types and the eight risk status types to simplify our discussion. Accordingly, we obtained sufficient conditions of covariate balance for no confounding (i.e., Equations 4 to 6), assuming that individuals can be “at risk” of sufficient causes even after they experience the outcome. In other words, Equations 4 to 6 are based on the comparability of potential completion of each sufficient cause. This would practically work well when one considers disease incidence as an outcome of interest if its latent period (i.e., the time from irreversible disease occurrence to detection) is relatively long (e.g., cancer). In this case, it would be admissible to conceptualize background factors in sufficient causes that complete after the disease occurrence as set of covariates. However, the assumption does not hold in some situations; for example, when considering all-cause mortality as an outcome of interest, the deceased cannot be “at risk” of sufficient causes, and thus it would be unnatural to use the assumption. In this Online Appendix, we aim to show a more subtle conceptualization of covariate balance by taking into account the potential completion time of each sufficient cause [7].

We let d_ϕ, d_e, and $d_{\bar{e}}$ denote the potential completion times of sufficient causes C_1, C_2E, and $C_3\overline{E}$ at which outcome would occur in an individual, respectively. We also let d_1 and d_0 denote the potential outcome occurrence time of an individual when exposed (E = 1) and unexposed (E = 0), respectively. In other words, we denote $d_1=\min(d_\phi, d_e)$ and $d_0=\min(d_\phi, d_{\bar{e}})$. Further, h denotes a maximum follow-up time for an individual. Note that the potential outcomes of Y can be described as: $Y_1=I(d_1\le h)$ and $Y_0=I(d_0\le h)$. Further, the background factors can be described as: $C_1=I(d_\phi\le h)$, $C_2=I(d_e\le h)$, and

$C_3 = I(d_{\bar{e}} \leq h)$. We assume that each potential completion time is different. Furthermore, we let $d_1 = \infty$ if the outcome would never occur for an individual when E = 1, and similarly define $d_0 = \infty$. Thus, when considering a binary exposure and a binary outcome, individuals can be classified into 24 (i.e., 4!) sequence types (Online Table 6) [7]. We let v_j, w_j, and x_j, j = 1–24, be proportions of sequence type j in the exposed group, the unexposed group, and the total population, respectively. Note that x_j can be calculated as: $v_j \times P[E=1] + w_j \times P[E=0]$. In some cases, we may assume that d_1 is always less than or equal to d_0 for all individuals, that is, $d_1 \leq d_0$ for all individuals. Suzuki et al. [7] referred to this assumption as "no preventive sequence", which excludes sequence types 5, 6, 10, 14, 17, 18, 23, and 24.

By considering the 24 sequence types, we can relax the assumption used in the main text. To develop a more subtle conceptualization of covariate balance, we focus on the comparability of background factors contained in the first completed sufficient cause between groups. Recall that, when the exposed is the target population, we need to consider the comparability of joint distributions of C_1 and C_3 between the exposed and the unexposed groups. Thus, we can obtain a more subtle sufficient condition of covariate balance for no confounding as:

$$\begin{aligned}
&\left\{P\left[d_\phi < d_{\bar{e}} \leq h \mid E=1\right] = P\left[d_\phi < d_{\bar{e}} \leq h \mid E=0\right]\right\}\\
&\wedge\left\{P\left[d_{\bar{e}} < d_\phi \leq h \mid E=1\right] = P\left[d_{\bar{e}} < d_\phi \leq h \mid E=0\right]\right\}\\
&\wedge\left\{P\left[d_\phi \leq h < d_{\bar{e}} \mid E=1\right] = P\left[d_\phi \leq h < d_{\bar{e}} \mid E=0\right]\right\}\\
&\wedge\left\{P\left[d_{\bar{e}} \leq h < d_\phi \mid E=1\right] = P\left[d_{\bar{e}} \leq h < d_\phi \mid E=0\right]\right\}\\
&\wedge\left\{P\left[h < d_\phi < d_{\bar{e}} \mid E=1\right] = P\left[h < d_\phi < d_{\bar{e}} \mid E=0\right]\right\}\\
&\wedge\left\{P\left[h < d_{\bar{e}} < d_\phi \mid E=1\right] = P\left[h < d_{\bar{e}} < d_\phi \mid E=0\right]\right\}\\
&\Leftrightarrow \left\{v_1 + v_2 + v_3 + v_9 = w_1 + w_2 + w_3 + w_9\right\}\\
&\wedge\left\{v_4 + v_5 + v_6 + v_{10} = w_4 + w_5 + w_6 + w_{10}\right\}\\
&\wedge\left\{v_7 + v_8 + v_{11} + v_{12} = w_7 + w_8 + w_{11} + w_{12}\right\}\\
&\wedge\left\{v_{13} + v_{14} + v_{17} + v_{18} = w_{13} + w_{14} + w_{17} + w_{18}\right\}\\
&\wedge\left\{v_{15} + v_{19} + v_{20} + v_{21} = w_{15} + w_{19} + w_{20} + w_{21}\right\}\\
&\wedge\left\{v_{16} + v_{22} + v_{23} + v_{24} = w_{16} + w_{22} + w_{23} + w_{24}\right\}, \qquad \text{[Eq. A7]}
\end{aligned}$$

which is stronger than Equation 4. Recall that, when (C_1, C_3) = (1, 1) in the first equation of Equation 4, we consider the comparability of subjects who are at risk of sufficient causes 1 and 3 (i.e., C_1 and $C_3\overline{E}$, respectively) between the exposed and the unexposed groups. Among these subjects of risk status types 1 and 3, both sufficient causes 1 and 3 *potentially* complete by the end of follow-up time. In the first two equations of Equation A7, however, we "decompose" the first equation of Equation 4 by taking into account the potential completion times of sufficient causes 1 and 3. For example, in the first equation of

Equation A7, the left-hand side represents a proportion of subjects who would have experienced the outcome because of sufficient cause 1 during the follow-up period (though sufficient cause 3 *potentially* completes by the end of follow-up time) in the exposed group had they been unexposed, whereas the right-hand side represents the corresponding proportion in the unexposed group. Note that the former quantity is, by definition, unobservable or counterfactual. The latter quantity is, though theoretically observable or actual, unable to be estimated unless we can understand the “etiology” of the cases [7]. Among these subjects of sequence types 1, 2, 3, and 9, the potential completion time of sufficient cause 1 is shorter than that of sufficient cause 3, both of which *potentially* complete by the end of follow-up time. An analogous discussion applies to the second equation of Equation A7. To summarize, when Equation A7 is met, the exposed and the unexposed groups are comparable in terms of the “etiologic mechanism” because of sufficient causes 1 and 3.

Conversely, when the unexposed group is the target population, we need to consider the comparability of joint distributions of C_1 and C_2 between the exposed and the unexposed groups, which yields a more subtle sufficient condition of covariate balance for no confounding as:

$$\begin{aligned}
&\left\{P\left[d_\phi < d_e \le h \mid E=1\right] = P\left[d_\phi < d_e \le h \mid E=0\right]\right\} \\
&\wedge\left\{P\left[d_e < d_\phi \le h \mid E=1\right] = P\left[d_e < d_\phi \le h \mid E=0\right]\right\} \\
&\wedge\left\{P\left[d_\phi \le h < d_e \mid E=1\right] = P\left[d_\phi \le h < d_e \mid E=0\right]\right\} \\
&\wedge\left\{P\left[d_e \le h < d_\phi \mid E=1\right] = P\left[d_e \le h < d_\phi \mid E=0\right]\right\} \\
&\wedge\left\{P\left[h < d_\phi < d_e \mid E=1\right] = P\left[h < d_\phi < d_e \mid E=0\right]\right\} \\
&\wedge\left\{P\left[h < d_e < d_\phi \mid E=1\right] = P\left[h < d_e < d_\phi \mid E=0\right]\right\} \\
&\Leftrightarrow \left\{v_1 + v_2 + v_5 + v_7 = w_1 + w_2 + w_5 + w_7\right\} \\
&\wedge\left\{v_3 + v_4 + v_6 + v_8 = w_3 + w_4 + w_6 + w_8\right\} \\
&\wedge\left\{v_9 + v_{10} + v_{11} + v_{12} = w_9 + w_{10} + w_{11} + w_{12}\right\} \\
&\wedge\left\{v_{13} + v_{14} + v_{15} + v_{16} = w_{13} + w_{14} + w_{15} + w_{16}\right\} \\
&\wedge\left\{v_{17} + v_{19} + v_{20} + v_{23} = w_{17} + w_{19} + w_{20} + w_{23}\right\} \\
&\wedge\left\{v_{18} + v_{21} + v_{22} + v_{24} = w_{18} + w_{21} + w_{22} + w_{24}\right\}, \qquad \text{[Eq. A8]}
\end{aligned}$$

which is stronger than Equation 5. Note that, in the first two equations of Equation A8, we “decompose” the first equation of Equation 5 when (C_1, C_2) = (1, 1), by taking into account the potential completion times of sufficient causes 1 and 2 (i.e., C_1 and C_2E, respectively). When Equation A8 is met, the exposed and the unexposed groups are comparable in terms of the “etiologic mechanism” because of sufficient causes 1 and 2.

Finally, when the target is the total population, we need to consider the comparability of joint distributions of C_1 and C_2 between the total population and the exposed group, as well as the comparability of joint distributions of C_1 and C_3 between the total population and the unexposed group. Thus, we can obtain a more subtle sufficient condition of covariate balance for no confounding as:

$$\begin{array}{l}
\left[\begin{array}{l}
\left\{ P\left[d_\phi < d_e \le h \right] = P\left[d_\phi < d_e \le h \mid E = 1 \right] \right\} \\
\wedge \left\{ P\left[d_e < d_\phi \le h \right] = P\left[d_e < d_\phi \le h \mid E = 1 \right] \right\} \\
\wedge \left\{ P\left[d_\phi \le h < d_e \right] = P\left[d_\phi \le h < d_e \mid E = 1 \right] \right\} \\
\wedge \left\{ P\left[d_e \le h < d_\phi \right] = P\left[d_e \le h < d_\phi \mid E = 1 \right] \right\} \\
\wedge \left\{ P\left[h < d_\phi < d_e \right] = P\left[h < d_\phi < d_e \mid E = 1 \right] \right\} \\
\wedge \left\{ P\left[h < d_e < d_\phi \right] = P\left[h < d_e < d_\phi \mid E = 1 \right] \right\}
\end{array} \right] \wedge \left[\begin{array}{l}
\left\{ P\left[d_\phi < d_{\bar{e}} \le h \right] = P\left[d_\phi < d_{\bar{e}} \le h \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_{\bar{e}} < d_\phi \le h \right] = P\left[d_{\bar{e}} < d_\phi \le h \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_\phi \le h < d_{\bar{e}} \right] = P\left[d_\phi \le h < d_{\bar{e}} \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_{\bar{e}} \le h < d_\phi \right] = P\left[d_{\bar{e}} \le h < d_\phi \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[h < d_\phi < d_{\bar{e}} \right] = P\left[h < d_\phi < d_{\bar{e}} \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[h < d_{\bar{e}} < d_\phi \right] = P\left[h < d_{\bar{e}} < d_\phi \mid E = 0 \right] \right\}
\end{array} \right] \\
\Leftrightarrow \left[\begin{array}{l}
\left\{ v_1 + v_2 + v_5 + v_7 = w_1 + w_2 + w_5 + w_7 \right\} \\
\wedge \left\{ v_3 + v_4 + v_6 + v_8 = w_3 + w_4 + w_6 + w_8 \right\} \\
\wedge \left\{ v_9 + v_{10} + v_{11} + v_{12} = w_9 + w_{10} + w_{11} + w_{12} \right\} \\
\wedge \left\{ v_{13} + v_{14} + v_{15} + v_{16} = w_{13} + w_{14} + w_{15} + w_{16} \right\} \\
\wedge \left\{ v_{17} + v_{19} + v_{20} + v_{23} = w_{17} + w_{19} + w_{20} + w_{23} \right\} \\
\wedge \left\{ v_{18} + v_{21} + v_{22} + v_{24} = w_{18} + w_{21} + w_{22} + w_{24} \right\}
\end{array} \right] \wedge \left[\begin{array}{l}
\left\{ v_1 + v_2 + v_3 + v_9 = w_1 + w_2 + w_3 + w_9 \right\} \\
\wedge \left\{ v_4 + v_5 + v_6 + v_{10} = w_4 + w_5 + w_6 + w_{10} \right\} \\
\wedge \left\{ v_7 + v_8 + v_{11} + v_{12} = w_7 + w_8 + w_{11} + w_{12} \right\} \\
\wedge \left\{ v_{13} + v_{14} + v_{17} + v_{18} = w_{13} + w_{14} + w_{17} + w_{18} \right\} \\
\wedge \left\{ v_{15} + v_{19} + v_{20} + v_{21} = w_{15} + w_{19} + w_{20} + w_{21} \right\} \\
\wedge \left\{ v_{16} + v_{22} + v_{23} + v_{24} = w_{16} + w_{22} + w_{23} + w_{24} \right\}
\end{array} \right],
\end{array} \quad \text{[Eq. A9]}$$

which is stronger than Equation 6. Note that Equation A9 is simply a product of Equations A7 and A8.

It is worth mentioning that Equations A7, A8, and A9 can be conceived as “decomposition” of Equations 4, 5, and 6 in the main text, respectively. This point can be readily shown by rewriting Equations 4, 5, and 6 using the potential completion times of sufficient causes as:

$$\begin{array}{l}
\left\{ P\left[\max(d_\phi, d_{\bar{e}}) \le h \mid E = 1 \right] = P\left[\max(d_\phi, d_{\bar{e}}) \le h \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_\phi \le h < d_{\bar{e}} \mid E = 1 \right] = P\left[d_\phi \le h < d_{\bar{e}} \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_{\bar{e}} \le h < d_\phi \mid E = 1 \right] = P\left[d_{\bar{e}} \le h < d_\phi \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[h < \min(d_\phi, d_{\bar{e}}) \mid E = 1 \right] = P\left[h < \min(d_\phi, d_{\bar{e}}) \mid E = 0 \right] \right\},
\end{array} \quad \text{[Eq. A10]}$$

$$\begin{array}{l}
\left\{ P\left[\max(d_\phi, d_e) \le h \mid E = 1 \right] = P\left[\max(d_\phi, d_e) \le h \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_\phi \le h < d_e \mid E = 1 \right] = P\left[d_\phi \le h < d_e \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[d_e \le h < d_\phi \mid E = 1 \right] = P\left[d_e \le h < d_\phi \mid E = 0 \right] \right\} \\
\wedge \left\{ P\left[h < \min(d_\phi, d_e) \mid E = 1 \right] = P\left[h < \min(d_\phi, d_e) \mid E = 0 \right] \right\},
\end{array} \quad \text{[Eq. A11]}$$

and

$$\left[\begin{array}{l}\left\{P\left[\max(d_{\phi},d_{e})\le h\right]=P\left[\max(d_{\phi},d_{e})\le h\mid E=1\right]\right\}\\ \wedge\left\{P\left[d_{\phi}\le h<d_{e}\right]=P\left[d_{\phi}\le h<d_{e}\mid E=1\right]\right\}\\ \wedge\left\{P\left[d_{e}\le h<d_{\phi}\right]=P\left[d_{e}\le h<d_{\phi}\mid E=1\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{e})\right]=P\left[h<\min(d_{\phi},d_{e})\mid E=1\right]\right\}\end{array}\right]$$

$$\wedge\left[\begin{array}{l}\left\{P\left[\max(d_{\phi},d_{\bar{e}})\le h\right]=P\left[\max(d_{\phi},d_{\bar{e}})\le h\mid E=0\right]\right\}\\ \wedge\left\{P\left[d_{\phi}\le h<d_{\bar{e}}\right]=P\left[d_{\phi}\le h<d_{\bar{e}}\mid E=0\right]\right\}\\ \wedge\left\{P\left[d_{\bar{e}}\le h<d_{\phi}\right]=P\left[d_{\bar{e}}\le h<d_{\phi}\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{\bar{e}})\right]=P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\}\end{array}\right], \qquad \text{[Eq. A12]}$$

respectively. Furthermore, Equations 4, 5, and 6 can be also conceived as “decomposition” of Equations A1, A2, and A3 in Online Appendix 2, respectively. This point can be readily shown by rewriting Equations A1, A2, and A3 using the potential outcome occurrence time as:

$$\begin{array}{l}\left\{P\left[\max(d_{\phi},d_{\bar{e}})\le h\mid E=1\right]=P\left[\max(d_{\phi},d_{\bar{e}})\le h\mid E=0\right]\right\}\\ \wedge\left\{P\left[\min(d_{\phi},d_{\bar{e}})\le h<\max(d_{\phi},d_{\bar{e}})\mid E=1\right]=P\left[\min(d_{\phi},d_{\bar{e}})\le h<\max(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=1\right]=P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\},\end{array} \qquad \text{[Eq. A13]}$$

$$\begin{array}{l}\left\{P\left[\max(d_{\phi},d_{e})\le h\mid E=1\right]=P\left[\max(d_{\phi},d_{e})\le h\mid E=0\right]\right\}\\ \wedge\left\{P\left[\min(d_{\phi},d_{e})\le h<\max(d_{\phi},d_{e})\mid E=1\right]=P\left[\min(d_{\phi},d_{e})\le h<\max(d_{\phi},d_{e})\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{e})\mid E=1\right]=P\left[h<\min(d_{\phi},d_{e})\mid E=0\right]\right\},\end{array} \qquad \text{[Eq. A14]}$$

and

$$\left[\begin{array}{l}\left\{P\left[\max(d_{\phi},d_{e})\leq h\right]=P\left[\max(d_{\phi},d_{e})\leq h\mid E=1\right]\right\}\\ \wedge\left\{P\left[\min(d_{\phi},d_{e})\leq h<\max(d_{\phi},d_{e})\right]=P\left[\min(d_{\phi},d_{e})\leq h<\max(d_{\phi},d_{e})\mid E=1\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{e})\right]=P\left[h<\min(d_{\phi},d_{e})\mid E=1\right]\right\}\end{array}\right]$$
$$\wedge\left[\begin{array}{l}\left\{P\left[\max(d_{\phi},d_{\bar{e}})\leq h\right]=P\left[\max(d_{\phi},d_{\bar{e}})\leq h\mid E=0\right]\right\}\\ \wedge\left\{P\left[\min(d_{\phi},d_{\bar{e}})\leq h<\max(d_{\phi},d_{\bar{e}})\right]=P\left[\min(d_{\phi},d_{\bar{e}})\leq h<\max(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{\bar{e}})\right]=P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\}\end{array}\right], \qquad \text{[Eq. A15]}$$

respectively. Finally, Equations A1, A2, and A3 can be conceived as “decomposition” of Equations 1, 2, and 3 in the main text, respectively. This point can be readily shown by rewriting Equations 1, 2, and 3 using the potential outcome occurrence time as:

$$\begin{array}{l}P\left[d_{0}\leq h\mid E=1\right]=P\left[d_{0}\leq h\mid E=0\right]\\ \Leftrightarrow\left\{P\left[\min(d_{\phi},\,d_{\bar{e}})\leq h\mid E=1\right]=P\left[\min(d_{\phi},\,d_{\bar{e}})\leq h\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=1\right]=P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\},\end{array} \qquad \text{[Eq. A16]}$$

$$\begin{array}{l}P\left[d_{1}\leq h\mid E=1\right]=P\left[d_{1}\leq h\mid E=0\right]\\ \Leftrightarrow\left\{P\left[\min(d_{\phi},\,d_{e})\leq h\mid E=1\right]=P\left[\min(d_{\phi},\,d_{e})\leq h\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{e})\mid E=1\right]=P\left[h<\min(d_{\phi},d_{e})\mid E=0\right]\right\},\end{array} \qquad \text{[Eq. A17]}$$

and

$$\begin{array}{l}\left\{P\left[d_{1}\leq h\right]=P\left[d_{1}\leq h\mid E=1\right]\right\}\wedge\left\{P\left[d_{0}\leq h\right]=P\left[d_{0}\leq h\mid E=0\right]\right\}\\ \Leftrightarrow\left[\begin{array}{l}\left\{P\left[\min(d_{\phi},\,d_{e})\leq h\right]=P\left[\min(d_{\phi},\,d_{e})\leq h\mid E=1\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{e})\right]=P\left[h<\min(d_{\phi},d_{e})\mid E=1\right]\right\}\end{array}\right]\\ \wedge\left[\begin{array}{l}\left\{P\left[\min(d_{\phi},\,d_{\bar{e}})\leq h\right]=P\left[\min(d_{\phi},\,d_{\bar{e}})\leq h\mid E=0\right]\right\}\\ \wedge\left\{P\left[h<\min(d_{\phi},d_{\bar{e}})\right]=P\left[h<\min(d_{\phi},d_{\bar{e}})\mid E=0\right]\right\}\end{array}\right],\end{array} \qquad \text{[Eq. A18]}$$

respectively. Recall that d_1 and d_0 are defined as $\min(d_{\phi},\,d_{e})$ and $\min(d_{\phi},\,d_{\bar{e}})$, respectively. Comparison of these rewritten equations would facilitate understanding of the more subtle conceptualization of covariate balance in this Online Appendix.

In conclusion, by considering potential completion time of each sufficient cause, we have developed a more subtle conceptualization of covariate balance. Even under this conceptualization, covariate balance is a sufficient, but not a necessary, condition for no confounding, irrespective of the target population. This point applies even under the assumption of no preventive sequence. Given that our main conclusions do not vary, we do not consider the sequence types to simplify our discussion in the main text.

References

1. Rosenman RH, Brand RJ, Jenkins D, Friedman M, Straus R, Wurm M. Coronary heart disease in Western Collaborative Group Study: final follow-up experience of 8 1/2 years. JAMA 1975;233(8):872-7.
2. Jewell NP. Statistics for Epidemiology. Chapman & Hall/CRC: Boca Raton, FL; 2004.
3. Selvin S. Statistical Tools for Epidemiologic Research. Oxford University Press: New York, NY; 2011.
4. VanderWeele TJ. Confounding and effect modification: distribution and measure. Epidemiol Method 2012;1(1):55-82. doi:10.1515/2161-962X.1004.
5. Suzuki E, Yamamoto E. Further refinements to the organizational schema for causal effects. Epidemiology 2014;25(4):618-9.
6. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A typology of four notions of confounding in epidemiology. J Epidemiol 2017;27(2):49-55.
7. Suzuki E, Yamamoto E, Tsuda T. On the relations between excess fraction, attributable fraction, and etiologic fraction. Am J Epidemiol 2012;175(6):567-75.

Online Table 1. Behavior type and occurrence of coronary heart disease in the Western Collaborative Group Study, 1960–1969

	Type A	**Type B**	**Total**
CHD	178	79	257
No CHD	1,411	1,486	2,897
Total	1,589	1,565	3,154

Abbreviation: CHD; coronary heart disease

Online Table 2. A possible distribution of response types and risk status types in the Western Collaborative Group Study [a]

Response types						Risk status types						
Type	Potential outcomes		Proportion of types in [b]			Type	Background factors			Proportion of types in [b]		
	Y_1	Y_0	Exposed (type A)	Unexposed (type B)	Total population		C_1	C_2	C_3	Exposed (type A)	Unexposed (type B)	Total population
1	1	1	$p_1 = 0.045$	$q_1 = 0.036$	$r_1 = 0.041$	1	1	1	1	$s_1 = 0.010$	$t_1 = 0.005$	$u_1 = 0.008$
						2	1	1	0	$s_2 = 0.005$	$t_2 = 0.010$	$u_2 = 0.007$
						3	1	0	1	$s_3 = 0.010$	$t_3 = 0.005$	$u_3 = 0.008$
						4	1	0	0	$s_4 = 0.010$	$t_4 = 0.010$	$u_4 = 0.010$
						5	0	1	1	$s_5 = 0.010$	$t_5 = 0.006$	$u_5 = 0.008$
2	1	0	$p_2 = 0.067$	$q_2 = 0.011$	$r_2 = 0.039$	6	0	1	0	$s_6 = 0.067$	$t_6 = 0.011$	$u_6 = 0.039$
3	0	1	$p_3 = 0.050$	$q_3 = 0.015$	$r_3 = 0.033$	7	0	0	1	$s_7 = 0.050$	$t_7 = 0.015$	$u_7 = 0.033$
4	0	0	$p_4 = 0.838$	$q_4 = 0.938$	$r_4 = 0.888$	8	0	0	0	$s_8 = 0.838$	$t_8 = 0.938$	$u_8 = 0.888$

[a] See Table 1 for notations.

[b] As shown in Online Table 1, r_j can be calculated as: $p_j \times 1,589/3,154 + q_j \times 1,565/3,154$. Likewise, u_j can be calculated as: $s_j \times 1,589/3,154 + t_j \times 1,565/3,154$. Proportions in the total population do not add to 1 due to rounding.

Online Table 3. The relationship between exchangeability, weaker covariate balance, covariate balance, and marginal covariate balance [a]

Target population	**Counterfactual model**	**Sufficient-cause model**			
	Exchangeability	**Exchangeability in terms of background factors**	**Weaker covariate balance** [b]	**Covariate balance**	**Marginal covariate balance** [c]
Exposed group	$Y_0 \coprod E$	$\Leftrightarrow\ \max(C_1, C_3) \coprod E$	$\Leftarrow\ (C_1 + C_3) \coprod E$	$\Leftarrow\ (C_1, C_3) \coprod E$	$\Rightarrow\ C_k \coprod E \quad (k = 1, 3)$
Unexposed group	$Y_1 \coprod E$	$\Leftrightarrow\ \max(C_1, C_2) \coprod E$	$\Leftarrow\ (C_1 + C_2) \coprod E$	$\Leftarrow\ (C_1, C_2) \coprod E$	$\Rightarrow\ C_k \coprod E \quad (k = 1, 2)$
Total population	$Y_e \coprod E \quad (e = 0, 1)$	$\Leftrightarrow\ \max(C_1, C_k) \coprod E \quad (k = 2, 3)$	$\Leftarrow\ (C_1 + C_k) \coprod E \quad (k = 2, 3)$	$\Leftarrow\ (C_1, C_k) \coprod E \quad (k = 2, 3)$	$\Rightarrow\ C_k \coprod E \quad (k = 1, 2, 3)$
	$\Uparrow$	$\Uparrow$	$\Uparrow$	$\Uparrow$	
	$(Y_0, Y_1) \coprod E$	$\Leftrightarrow\ (\max(C_1, C_3), \max(C_1, C_2)) \coprod E$	$\Leftarrow\ ((C_1 + C_3), (C_1 + C_2)) \coprod E$	$\Leftarrow\ (C_1, C_2, C_3) \coprod E$	

[a] See Table 1 for notations.

[b] See Online Appendix 2.

[c] See Online Appendix 3.

Online Table 4. A numerical example of no confounding when there is covariate imbalance [a]

Response types						Risk status types						
Type	Potential outcomes		Proportion of types in [b]			Type	Background factors			Proportion of types in [b]		
	Y_1	Y_0	Exposed	Unexposed	Total population		C_1	C_2	C_3	Exposed	Unexposed	Total population
1	1	1	$p_1 = 0.60$	$q_1 = 0.50$	$r_1 = 0.52$	1	1	1	1	$s_1 = 0.10$	$t_1 = 0.05$	$u_1 = 0.06$
						2	1	1	0	$s_2 = 0.05$	$t_2 = 0.10$	$u_2 = 0.09$
						3	1	0	1	$s_3 = 0.15$	$t_3 = 0.10$	$u_3 = 0.11$
						4	1	0	0	$s_4 = 0.10$	$t_4 = 0.10$	$u_4 = 0.10$
						5	0	1	1	$s_5 = 0.20$	$t_5 = 0.15$	$u_5 = 0.16$
2	1	0	$p_2 = 0.05$	$q_2 = 0.15$	$r_2 = 0.13$	6	0	1	0	$s_6 = 0.05$	$t_6 = 0.15$	$u_6 = 0.13$
3	0	1	$p_3 = 0.10$	$q_3 = 0.20$	$r_3 = 0.18$	7	0	0	1	$s_7 = 0.10$	$t_7 = 0.20$	$u_7 = 0.18$
4	0	0	$p_4 = 0.25$	$q_4 = 0.15$	$r_4 = 0.17$	8	0	0	0	$s_8 = 0.25$	$t_8 = 0.15$	$u_8 = 0.17$

[a] We consider a binary exposure E (1 = exposed, 0 = unexposed) and a binary outcome Y (1 = outcome occurred, 0 = outcome did not occur). See Table 1 for notations.

[b] We consider a situation in which the prevalence of exposure in the total population is 0.2. Thus, r_j can be calculated as: $p_j \times 0.2 + q_j \times 0.8$. Likewise, u_j can be calculated as: $s_j \times 0.2 + t_j \times 0.8$.

Online Table 5. A numerical example of confounding when there is covariate balance under the alternative conditions [a]

Response types						Risk status types						
Type	**Potential outcomes**		**Proportion of types in** [b]			**Type**	**Background factors**			**Proportion of types in** [b]		
	$\boldsymbol{Y_1}$	$\boldsymbol{Y_0}$	**Exposed**	**Unexposed**	**Total population**		$\boldsymbol{C_1}$	$\boldsymbol{C_2}$	$\boldsymbol{C_3}$	**Exposed**	**Unexposed**	**Total population**
1	1	1	$p_1 = 0.55$	$q_1 = 0.50$	$r_1 = 0.51$	1	1	1	1	$s_1 = 0.10$	$t_1 = 0.10$	$u_1 = 0.10$
						2	1	1	0	$s_2 = 0.05$	$t_2 = 0.10$	$u_2 = 0.09$
						3	1	0	1	$s_3 = 0.10$	$t_3 = 0.05$	$u_3 = 0.06$
						4	1	0	0	$s_4 = 0.10$	$t_4 = 0.10$	$u_4 = 0.10$
						5	0	1	1	$s_5 = 0.20$	$t_5 = 0.15$	$u_5 = 0.16$
2	1	0	$p_2 = 0.05$	$q_2 = 0.05$	$r_2 = 0.05$	6	0	1	0	$s_6 = 0.05$	$t_6 = 0.05$	$u_6 = 0.05$
3	0	1	$p_3 = 0.05$	$q_3 = 0.15$	$r_3 = 0.13$	7	0	0	1	$s_7 = 0.05$	$t_7 = 0.15$	$u_7 = 0.13$
4	0	0	$p_4 = 0.35$	$q_4 = 0.30$	$r_4 = 0.31$	8	0	0	0	$s_8 = 0.35$	$t_8 = 0.30$	$u_8 = 0.31$

[a] We consider a binary exposure E (1 = exposed, 0 = unexposed) and a binary outcome Y (1 = outcome occurred, 0 = outcome did not occur). See Table 1 for notations.

[b] We consider a situation in which the prevalence of exposure in the total population is 0.2. Thus, r_j can be calculated as: $p_j \times 0.2 + q_j \times 0.8$. Likewise, u_j can be calculated as: $s_j \times 0.2 + t_j \times 0.8$.

Online Table 6. Correspondence between response types, risk status types, and sequence types under a binary exposure and a binary outcome [a]

Response types						**Risk status types**							**Sequence types**						
Type	**Potential outcomes**		**Proportion of types in** [b]			**Type**	**Background factors**			**Proportion of types in** [b]			**Type**	**Sequence of potential completion time**	**Potential outcome occurrence time**		**Proportion of types in** [b]		
	$\boldsymbol{Y_1}$	$\boldsymbol{Y_0}$	**Exposed**	**Unexposed**	**Total population**		$\boldsymbol{C_1}$	$\boldsymbol{C_2}$	$\boldsymbol{C_3}$	**Exposed**	**Unexposed**	**Total population**			$\boldsymbol{d_1}$	$\boldsymbol{d_0}$	**Exposed**	**Unexposed**	**Total population**
1	1	1	p_1	q_1	r_1	1[d]	1	1	1	s_1	t_1	u_1	1	$d_\phi < d_e < d_{\bar{e}} \le h$	d_ϕ	d_ϕ	v_1	w_1	x_1
													2	$d_\phi < d_{\bar{e}} < d_e \le h$	d_ϕ	d_ϕ	v_2	w_2	x_2
													3	$d_e < d_\phi < d_{\bar{e}} \le h$	d_e	d_ϕ	v_3	w_3	x_3
													4	$d_e < d_{\bar{e}} < d_\phi \le h$	d_e	$d_{\bar{e}}$	v_4	w_4	x_4
													5[e]	$d_{\bar{e}} < d_\phi < d_e \le h$	d_ϕ	$d_{\bar{e}}$	v_5	w_5	x_5
													6[e]	$d_{\bar{e}} < d_e < d_\phi \le h$	d_e	$d_{\bar{e}}$	v_6	w_6	x_6
						2	1	1	0	s_2	t_2	u_2	7	$d_\phi < d_e \le h < d_{\bar{e}}$	d_ϕ	d_ϕ	v_7	w_7	x_7
													8	$d_e < d_\phi \le h < d_{\bar{e}}$	d_e	d_ϕ	v_8	w_8	x_8
						3[d]	1	0	1	s_3	t_3	u_3	9	$d_\phi < d_{\bar{e}} \le h < d_e$	d_ϕ	d_ϕ	v_9	w_9	x_9
													10[e]	$d_{\bar{e}} < d_\phi \le h < d_e$	d_ϕ	$d_{\bar{e}}$	v_{10}	w_{10}	x_{10}
						4	1	0	0	s_4	t_4	u_4	11	$d_\phi \le h < d_e < d_{\bar{e}}$	d_ϕ	d_ϕ	v_{11}	w_{11}	x_{11}
													12	$d_\phi \le h < d_{\bar{e}} < d_e$	d_ϕ	d_ϕ	v_{12}	w_{12}	x_{12}
						5[d]	0	1	1	s_5	t_5	u_5	13	$d_e < d_{\bar{e}} \le h < d_\phi$	d_e	$d_{\bar{e}}$	v_{13}	w_{13}	x_{13}
													14[e]	$d_{\bar{e}} < d_e \le h < d_\phi$	d_e	$d_{\bar{e}}$	v_{14}	w_{14}	x_{14}
2	1	0	p_2	q_2	r_2	6	0	1	0	s_6	t_6	u_6	15	$d_e \le h < d_\phi < d_{\bar{e}}$	d_e	d_ϕ	v_{15}	w_{15}	x_{15}
													16	$d_e \le h < d_{\bar{e}} < d_\phi$	d_e	$d_{\bar{e}}$	v_{16}	w_{16}	x_{16}
3[c]	0	1	p_3	q_3	r_3	7[d]	0	0	1	s_7	t_7	u_7	17[e]	$d_{\bar{e}} \le h < d_\phi < d_e$	d_ϕ	$d_{\bar{e}}$	v_{17}	w_{17}	x_{17}
													18[e]	$d_{\bar{e}} \le h < d_e < d_\phi$	d_e	$d_{\bar{e}}$	v_{18}	w_{18}	x_{18}
4	0	0	p_4	q_4	r_4	8	0	0	0	s_8	t_8	u_8	19	$h < d_\phi < d_e < d_{\bar{e}}$	d_ϕ	d_ϕ	v_{19}	w_{19}	x_{19}
													20	$h < d_\phi < d_{\bar{e}} < d_e$	d_ϕ	d_ϕ	v_{20}	w_{20}	x_{20}
													21	$h < d_e < d_\phi < d_{\bar{e}}$	d_e	d_ϕ	v_{21}	w_{21}	x_{21}
													22	$h < d_e < d_{\bar{e}} < d_\phi$	d_e	$d_{\bar{e}}$	v_{22}	w_{22}	x_{22}
													23[e]	$h < d_{\bar{e}} < d_\phi < d_e$	d_ϕ	$d_{\bar{e}}$	v_{23}	w_{23}	x_{23}
													24[e]	$h < d_{\bar{e}} < d_e < d_\phi$	d_e	$d_{\bar{e}}$	v_{24}	w_{24}	x_{24}

[a] We consider a binary exposure E (1 = exposed, 0 = unexposed) and a binary outcome Y (1 = outcome occurred, 0 = outcome did not occur). We consider two potential outcomes, Y_e, for an individual. We consider three different types of sufficient causes for outcome Y along with certain binary background factors as follows: C_1, C_2E, and $C_3\overline{E}$, where we let $\overline{E}$ denote the complement of E. We let d_ϕ, d_e, and $d_{\bar{e}}$ denote the potential completion times of sufficient causes C_1, C_2E, and $C_3\overline{E}$ at which outcome would occur in an individual, respectively. We also let d_1 and d_0 denote the potential outcome occurrence time of an individual when exposed and unexposed, respectively. In other words, we denote $d_1 = \min(d_\phi, d_e)$ and $d_0 = \min(d_\phi, d_{\bar{e}})$. Further, h denotes a maximum follow-up time of an individual.

[b] Note that r_j can be calculated as: $p_j \times P[E=1] + q_j \times P[E=0]$, where $P[E=e]$ represents the prevalence of $E = e$ in the total population. Likewise, u_j can be calculated as: $s_j \times P[E=1] + t_j \times P[E=0]$, and x_j can be calculated as: $v_j \times P[E=1] + w_j \times P[E=0]$.

[c] Under the assumption of (counterfactual) positive monotonicity (i.e., $Y_0 \le Y_1$ for all individuals), this response type is excluded.

[d] Under the assumption of no preventive action, or sufficient-cause positive monotonicity (i.e., $C_3 = 0$ for all individuals), these risk status types are excluded.

[e] Under the assumption of no preventive sequence (i.e., $d_1 \le d_0$ for all individuals), these sequence types are excluded.