Doron J. Shahar* and Eyal Shahar

A Theorem at the Core of Colliding Bias

DOI 10.1515/ijb-2016-0055

Abstract: Conditioning on a shared outcome of two variables can alter the association between these variables, possibly adding a bias component when estimating effects. In particular, if two causes are marginally independent, they might be dependent in strata of their common effect. Explanations of the phenomenon, however, do not explicitly state when dependence will be created and have been largely informal. We prove that two, marginally independent, causes will be dependent in a particular stratum of their shared outcome if and only if they modify each other's effects, on a probability ratio scale, on that value of the outcome variable. Using our result, we also qualify the claim that such causes will "almost certainly" be dependent in at least one stratum of the outcome: dependence must be created in one stratum of a binary outcome, and independence can be maintained in every stratum of a trinary outcome.

Keywords: colliding bias, independence, effect modification, causal diagram

1 Introduction

When two, marginally independent, variables affect a third variable, they might become dependent (associated) conditional on the latter [1]. If the causal structure is described by arrows (e.g., $A \rightarrow C \leftarrow B$), the shared effect (C) is called a collider on the path between its causes (A and B). In the context of causal inquiry, where effects are estimated by associations, a newly formed association after conditioning on a collider can add colliding bias – the bias that might arise from conditioning on every collider along a path between the cause and effect of interest.

For example, if $A \rightarrow C \leftarrow B$ (Figure 1, Diagram A), conditioning on C might create bias when estimating the effect (here null) of A on B (e.g., Berkson's bias). Or another example: under the M-structure (Figure 1, Diagram B) conditioning on C might create bias when estimating the effect of E on D [2, 3].

Colliding bias, known by at least half a dozen names [4], is the antithetical counterpart of confounding [5]. Both biases are well recognized in the literature on causal diagrams, and theorems based on d -separation allow for the removal of confounding without adding colliding bias [2]. Nonetheless, d -connection, which might arise after conditioning on a collider, does not necessarily result in bias, because d -connection – the opposite of d -separation – does not imply dependence.

To our knowledge no article was devoted to a basic underlying question: when does conditioning on a collider create an association between its causes? In fact, the literature contains various statements on the possible consequences of conditioning on a collider, some of which are non-specific and others sound like unproven theorems. We present here a general theorem at the core of colliding bias, the origin of which can be traced to the case-only design.

2 Notation, definitions, and basic propositions

Throughout this paper let A , B , and C be discrete (non-degenerate) random variables. Let $\{a_i\}_{i=1}^n$ and $\{b_k\}_{k=1}^m$ be the values of A and B , respectively, where n and m may be either finite or infinite. The lower case letters

*Corresponding author: **Doron J. Shahar**, Department of Mathematics, University of Arizona, 617 N. Santa Rita Ave., Tucson, AZ 85721, USA, E-mail: dshahar@math.arizona.edu

Eyal Shahar, Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, 1295 N. Martin Ave., Tucson, AZ 85724, USA, E-mail: shahar@email.arizona.edu

Figure 1: Two structures in which colliding bias might arise following conditioning on C .

a , b , and c will denote an arbitrary value of A , B , and C , respectively. For the time being, we fix a value c of C . We will consider the case where A and B are marginally independent causes of C , as may be depicted by the causal diagram $A \rightarrow C \leftarrow B$ (Figure 1, Diagram A).

For completeness, we will provide definitions of effects on C and of effect modification between A and B . In particular, we will define both effects and effect modification in terms of the *probability ratio*, because our results depend on this measure of effect. We shall assume that none of the probabilities mentioned hereafter is zero – a standard assumption under indeterminism – so the probability ratio may always be defined. Specifically, we assume that $P(A = a, B = b, C = c) \neq 0$ for all a , b , and c .

First, we define the effect of A on C under the causal structure $A \rightarrow C \leftarrow B$.

Definition 1. The effect of A (a_j vs. a_i) on $C = c$ when $B = b$ is

$$r_{ij}(b) = \frac{P(C = c \mid A = a_j, B = b)}{P(C = c \mid A = a_i, B = b)}$$

The effect of A on C could depend on the value of B . When it does, we say that B modifies A 's effect on C as defined below.

Definition 2. B modifies A 's effect on $C = c$ if there are two values a_i, a_j of A and two values b_k, b_l of B such that $r_{ij}(b_k) \neq r_{ij}(b_l)$

Similarly, we can define the effect of B on C and modification of B 's effect on C by A :

Definition 3. The effect of B (b_l vs. b_k) on $C = c$ when $A = a$ is

$$s_{kl}(a) = \frac{P(C = c \mid B = b_l, A = a)}{P(C = c \mid B = b_k, A = a)}$$

Definition 4. A modifies B 's effect on $C = c$ if there are two values b_k, b_l of B and two values a_i, a_j of A such that $s_{kl}(a_i) \neq s_{kl}(a_j)$

Effect modification is a symmetric property as stated in the proposition below.

Proposition 1. B modifies A 's effect on $C = c$ if and only if A modifies B 's effect on $C = c$.

Proof. Suppose that B modifies A 's effect on $C = c$. Then

$$\frac{P(C = c \mid A = a_j, B = b_k)}{P(C = c \mid A = a_i, B = b_k)} = r_{ij}(b_k) \neq r_{ij}(b_l) = \frac{P(C = c \mid A = a_j, B = b_l)}{P(C = c \mid A = a_i, B = b_l)}$$

for some i, j, k , and l . If we multiply both sides of the above inequality by $P(C = c \mid A = a_i, B = b_l) / P(C = c \mid A = a_j, B = b_k) \neq 0$, we obtain

$$\frac{P(C = c \mid A = a_i, B = b_l)}{P(C = c \mid A = a_i, B = b_k)} = s_{kl}(a_i) \neq s_{kl}(a_j) = \frac{P(C = c \mid A = a_j, B = b_l)}{P(C = c \mid A = a_j, B = b_k)}$$

That is, A modifies B 's effect on $C = c$, which proves the “only if” portion of the above proposition. The “if” portion follows by the same reasoning. $\square$

By proposition 1, we may speak of A and B modifying each other's effects on $C = c$. If A and B do not modify each other's effect on $C = c$, then $r_{ij}(b)$ and $s_{kl}(a)$ do not depend on b and a , respectively. That is, in the absence of effect modification, we may speak of A 's effect on C without noting the value of B and of B 's effect on C without noting the value of A (Definition 5).

Definition 5. If A and B do not modify each other's effects on $C = c$, then the effect of A (a_j vs. a_i) on $C = c$ is $r_{ij}(b)$ for any b , and the effect of B (b_l vs. b_k) on $C = c$ is $s_{kl}(a)$ for any a .

Next, we prove a more convenient formula for the effects of A and B on C – in the absence of effect modification.

Proposition 2. If A and B do not modify each other's effects on $C = c$, then

$$r_{ij}(b) = \frac{P(C = c \mid A = a_j)}{P(C = c \mid A = a_i)}, \quad s_{kl}(a) = \frac{P(C = c \mid B = b_l)}{P(C = c \mid B = b_k)}$$

for all b and a , respectively.

Proof. Since $r_{ij}(b)$ does not depend on b , it equals any weighted average of the $r_{ij}(b_k)$. That is, $r_{ij}(b) = \sum_{k=1}^m w_{b_k} r_{ij}(b_k)$ for any $w_{b_k} \geq 0$ with $\sum_{k=1}^m w_{b_k} = 1$. Our goal will be to find suitable weights so that $r_{ij}(b)$ can be expressed without b .

First, utilizing the independence of A and B , we will rewrite $r_{ij}(b)$ in the following way,

$$\begin{aligned} r_{ij}(b) &= \frac{P(C = c \mid A = a_j, B = b)}{P(C = c \mid A = a_i, B = b)} \\ &= \frac{P(C = c \mid A = a_j, B = b)P(B = b \mid A = a_j)}{P(C = c \mid A = a_i, B = b)P(B = b \mid A = a_i)} \\ &= \frac{P(C = c, B = b \mid A = a_j)}{P(C = c, B = b \mid A = a_i)} \end{aligned} \tag{1}$$

Given the above expression for $r_{ij}(b)$,

$$w_b = \frac{P(C = c, B = b \mid A = a_i)}{P(C = c \mid A = a_i)}$$

are natural choices for the weights. With these weights and eq. (1),

$$r_{ij}(b) = \sum_{k=1}^m w_{b_k} r_{ij}(b_k) = \sum_{k=1}^m \frac{P(C = c, B = b_k \mid A = a_j)}{P(C = c \mid A = a_i)} = \frac{P(C = c \mid A = a_j)}{P(C = c \mid A = a_i)}$$

for all b . The proof for $s_{kl}(a)$ follows the same reasoning. $\square$

This result has been referred to as collapsibility of the probability ratio [6, 7].

Now, we will define what it means for A and B to have a null effect on C .

Definition 6. A has a null effect on $C = c$ if $r_{ij}(b) = 1$ for all i, j and b . Similarly, B has a null effect on $C = c$ if $s_{kl}(a) = 1$ for all k, l and a .

Which naturally extends to the following definition:

Definition 7. A has a null effect on C if A has a null effect on $C = c$ for all c . Similarly, B has a null effect on C if B has a null effect on $C = c$ for all c .

Lastly, we will introduce the notion of an effect matrix.

Definition 8. The effect matrix of A on $C = c$ when $B = b$ is the matrix $R(b) = (r_{ij}(b))$. The effect matrix of B on $C = c$ when $A = a$ is the matrix $S(a) = (s_{kl}(a))$.

Note that $R(b)$ (or $S(a)$) will be an infinite matrix when A (or B) has an infinite number of values. With the notion of an effect matrix, we can more succinctly describe the properties of effect modification and null effects. Specifically, B does not modify A 's effect on $C = c$ iff $R(b)$ does not depend on b . And A has a null effect on $C = c$ iff for every b , $R(b)$ is the matrix whose entries are all 1.

3 Main theorem

Theorem 1. If A and B are marginally independent causes of C , then A and B are dependent conditional on $C = c$ if and only if A and B modify each other's effects on $C = c$.

Proof. We will prove both the “if” and “only if” directions of the statement by proving their contrapositives. Note that A and B are independent conditional on $C = c$ if and only if $P(A = a_i | B = b, C = c) = P(A = a_i | C = c)$ for all a_i and b . Below we shall write both $P(A = a_i | B = b, C = c)$ and $P(A = a_i | C = c)$ in terms of effects and compare the results. Specifically, $P(A = a_i | B = b, C = c)$ will be written in terms of the $r_{ij}(b)$, and $P(A = a_i | C = c)$ will be written in terms of $P(C = c | A = a_j)/P(C = c | A = a_i)$, which would be an effect if B didn't modify A 's effect on $C = c$ (Proposition 2).

$$\begin{aligned}
 P(A = a_i | B = b, C = c) &= \frac{P(C = c | A = a_i, B = b)P(A = a_i | B = b)}{P(C = c | B = b)} \\
 &= \frac{P(C = c | A = a_i, B = b)P(A = a_i | B = b)}{\sum_{j=1}^n P(C = c | A = a_j, B = b)P(A = a_j | B = b)} \\
 &= \frac{P(C = c | A = a_i, B = b)P(A = a_i)}{\sum_{j=1}^n P(C = c | A = a_j, B = b)P(A = a_j)} \\
 &= P(A = a_i) \left[\sum_{j=1}^n r_{ij}(b)P(A = a_j) \right]^{-1}
 \end{aligned} \tag{2}$$

$$\begin{aligned}
P(A = a_i \mid C = c) &= \frac{P(C = c \mid A = a_i)P(A = a_i)}{P(C = c)} \\
&= \frac{P(C = c \mid A = a_i)P(A = a_i)}{\sum_{j=1}^n P(C = c \mid A = a_j)P(A = a_j)} \\
&= P(A = a_i) \left[\sum_{j=1}^n \frac{P(C = c \mid A = a_j)}{P(C = c \mid A = a_i)} P(A = a_j) \right]^{-1}
\end{aligned} \tag{3}$$

If A and B do not modify each other's effects on $C = c$, then, utilizing proposition 2, the last line in eq. (2) equals the last line in eq. (3). Therefore, $P(A = a_i \mid B = b, C = c) = P(A = a_i \mid C = c)$ for all a_i and b . That is, A and B are independent conditional on $C = c$, which proves the “only if” direction.

If A and B are independent conditional on $C = c$, then $P(A = a_i \mid B = b, C = c) = P(A = a_i \mid C = c)$ for all a_i and b . Hence, by eq. (2), $\sum_{j=1}^n r_{ij}(b)P(A = a_j)$ does not depend on b for any i . Using the effect matrix, we can say that $R(b)\vec{w}$ does not depend on b where $\vec{w} = (w_1, \dots, w_n)$ and $w_j = P(A = a_j)$. In lemma 1 we show that if $R(b)\vec{w}$ does not depend on b , then $R(b)$ does not depend on b either. That is, B does not modify A 's effect on $C = c$, which proves the “if” direction. $\square$

Lemma 1. *If $R(b)\vec{w}$ does not depend on b , where $\vec{w} = (w_1, \dots, w_n)$ and $w_j = P(A = a_j)$, then $R(b)$ does not depend on b either.*

Proof. Let b_k, b_l be distinct values of B . Our goal will be to show that $R(b_k) = R(b_l)$. Since b_k and b_l are arbitrary distinct values of B , it will follow that $R(b)$ does not depend on b .

Let $R_i(b)$ denote the i th row of $R(b)$. Note the following three properties concerning the rows of $R(b)$:

1. $R_i(b_k)\vec{w} = R_i(b_l)\vec{w}$ for all i
2. $R_i(b) = r_{ij}(b)R_j(b)$ for all i, j , and b
3. $R_i(b)\vec{w} > 0$ for all i and b

Since $R(b)\vec{w}$ does not depend on b , $R_i(b_k)\vec{w} = R_i(b_l)\vec{w}$ for all i , which is property (1). Property (2) follows directly from the definition of $r_{ij}(b)$, and property (3) follows from the fact that $R(b)$ contains only positive entries. Using these properties, we will show that $R(b_k) = R(b_l)$.

By properties (1) and (2),

$$r_{ij}(b_k)R_j(b_k)\vec{w} = r_{ij}(b_l)R_j(b_l)\vec{w}$$

for all i, j . Then by properties (1) and (3), we may divide both sides of the previous equation by $R_j(b_k)\vec{w} = R_j(b_l)\vec{w}$ to obtain

$$r_{ij}(b_k) = r_{ij}(b_l)$$

for all i, j . That is, $R(b_k) = R(b_l)$. $\square$

In appendix A, we give an alternative proof of theorem 1.

In proving the theorem, we assumed that $P(A = a, B = b, C = c) \neq 0$ for all a, b , and c . It may be the case, though, that A and B are already restricted to take certain values. (For example, when studying the effect of A on C we may fix A to only two values, often labeled “exposed” and “unexposed”.) Theorem 1 will still

hold under such restrictions if we replace the phrase “modify each other’s effects on $C = c$ ” with the phrase “conditionally modify each other’s effects on $C = c$ ” where the latter phrase is defined as follows:

Definition 9. *If we restrict A and B to take only values a_i and b_k , respectively, with $i \in U \subseteq \{1, 2, \dots, n\}$ and $k \in V \subseteq \{1, 2, \dots, m\}$, then A and B conditionally modify each other’s effects on $C = c$ if for some $i, j \in U$ and $k, l \in V$, $r_{ij}(b_k) \neq r_{ij}(b_l)$ (or equivalently, $s_{kl}(a_i) \neq s_{kl}(a_j)$)*

Note that when n (or m) is infinite the notation $\{1, 2, \dots, n\}$ (or $\{1, 2, \dots, m\}$) denotes the positive integers.

4 Some special cases

Some versions of the following statement may be found in the literature:

“If the effects of A and B on C are not null, then A and B are dependent in at least one stratum of C .”

Using theorem 1, we will prove that the statement is true when C is binary, and we will provide a counterexample when C is not binary.

That the above statement holds when C is binary follows from the following proposition:

Proposition 3. *If C is a binary variable (with values denoted 0 and 1) and the effects of A and B on C are not null, then A and B modify each other’s effects on at least one value of C .*

Proof. Let $p_{ik} = P(C = 0 \mid A = a_i, B = b_k)$. Suppose that A and B do not modify each other’s effects on either value of C . We will then show that A or B has a null effect on C .

Since B does not modify A ’s effect on $C = 0$,

$$\frac{p_{jk}}{p_{ik}} = \frac{p_{jl}}{p_{il}} \quad \left(\text{or equivalently, } \frac{p_{il}}{p_{ik}} = \frac{p_{jl}}{p_{jk}} \right)$$

for all i, j, k, l . And since B does not modify A ’s effect on $C = 1$ either,

$$\frac{1 - p_{jk}}{1 - p_{ik}} = \frac{1 - p_{jl}}{1 - p_{il}} \quad \left(\text{or equivalently, } \frac{1 - p_{il}}{1 - p_{ik}} = \frac{1 - p_{jl}}{1 - p_{jk}} \right)$$

for all i, j, k, l . Therefore,

$$p_{jk}p_{il} = p_{jl}p_{ik}, \quad (1 - p_{jk})(1 - p_{il}) = (1 - p_{jl})(1 - p_{ik})$$

Multiplying out both sides of the right equation and utilizing the left equation, we obtain the equality

$$p_{jk} - p_{ik} = p_{jl} - p_{il}$$

which can also be written as

$$\left(\frac{p_{jk}}{p_{ik}} - 1 \right) = \frac{p_{il}}{p_{ik}} \left(\frac{p_{jl}}{p_{il}} - 1 \right)$$

Since $p_{jk}/p_{ik} = p_{jl}/p_{il}$, the above equality implies that $p_{jk}/p_{ik} = 1$ or $p_{il}/p_{ik} = 1$ for all i, j, k , and l . (Because A and B do not modify each other’s effects on $C = 0$, $p_{jk}/p_{ik} = 1$ iff $p_{jl}/p_{il} = 1$ for all l , and $p_{il}/p_{ik} = 1$ iff $p_{jl}/p_{jk} = 1$ for all j .) It follows that for all i, j , and l , $p_{jl}/p_{il} = 1$ or for all k, l , and j , $p_{jl}/p_{jk} = 1$. That is, A has a null effect on $C = 0$ or B has a null effect on $C = 0$.

If A has a null effect on $C = 0$, then A also has a null effect on $C = 1$, because C is binary. Therefore, A has a null effect on C . Likewise, if B has a null effect on $C = 0$, then B also has a null effect on C . In any case, A or B has a null effect on C . $\square$

Proposition 3 and theorem 1 can be combined to give the following corollary.

Corollary 1. *If A and B are marginally independent causes of a binary variable C , and if the effects of A and B on C are not null, then A and B are dependent in at least one stratum of C .*

Corollary 1 does not hold when C is not binary. In that case, it is possible that A and B have non-null effects on C and do not modify each other's effects on any value of C . It then follows from theorem 1 that A and B will be independent in every stratum of C . Our counterexample will be based on this reasoning.

Below we give an example of a trinary variable C (with values 1, 2, 3) and binary A and B (with values 0, 1) such that A and B have non-null effects on C . Consider the probabilities below, which were found by guessing values that satisfy the conditions for the absence of effect modification (Example 1).

Example 1.

$$\begin{array}{lll} P(C = 1 \mid A = 0, B = 0) = \frac{1}{6} & P(C = 2 \mid A = 0, B = 0) = \frac{1}{6} & P(C = 3 \mid A = 0, B = 0) = \frac{4}{6} \\ P(C = 1 \mid A = 1, B = 0) = \frac{3}{10} & P(C = 2 \mid A = 1, B = 0) = \frac{1}{10} & P(C = 3 \mid A = 1, B = 0) = \frac{6}{10} \\ P(C = 1 \mid A = 0, B = 1) = \frac{2}{9} & P(C = 2 \mid A = 0, B = 1) = \frac{3}{9} & P(C = 3 \mid A = 0, B = 1) = \frac{4}{9} \\ P(C = 1 \mid A = 1, B = 1) = \frac{2}{5} & P(C = 2 \mid A = 1, B = 1) = \frac{1}{5} & P(C = 3 \mid A = 1, B = 1) = \frac{2}{5} \end{array}$$

The effects of A (1 vs. 0) on $C = 1, 2$, and 3 when $B = 0$ are 1.8, 0.6, and 0.9, respectively.

The effects of A (1 vs. 0) on $C = 1, 2$, and 3 when $B = 1$ are 1.8, 0.6, and 0.9, respectively.

The effects of B (1 vs. 0) on $C = 1, 2$, and 3 when $A = 0$ are 1.3, 2, and 0.6, respectively.

The effects of B (1 vs. 0) on $C = 1, 2$, and 3 when $A = 1$ are 1.3, 2, and 0.6, respectively.

Therefore, A and B do not modify each other's effects on any value of C nor do they have null effects on C . Appendix B includes tables with counts, illustrating that A and B are independent in every stratum of C .

5 Related work and possible extensions

Heuristic arguments for binary variables are often used to explain why conditioning on a common effect sometimes creates an association between its causes [8]. For instance, if each of two drugs, A and B , is a deterministic cause of bleeding ($C = 1$), and we are told that John did not take drug B ($B = 0$), we are inclined to guess that John took drug A ($A = 1$) – once we are informed that John suffered bleeding ($C = 1$). That rational guess intuitively points to a conditional association between A and B . Notice, however, that the deterministic story above hides extreme effect modification between A and B :

$$\frac{P(C = 1 \mid A = 1, B = 0)}{P(C = 1 \mid A = 0, B = 0)} \gg 1, \quad \frac{P(C = 1 \mid A = 1, B = 1)}{P(C = 1 \mid A = 0, B = 1)} = 1$$

Previous work has extended the heuristic explanation to a formal deterministic model when effects are monotonic (e.g., taking drug A cannot cause the death of John and prevent the death of Jane) and all variables are binary [9]. Under these constraints and others, it is possible to infer the sign of the conditional covariance between two causes of C in the strata of C [9]. Although a zero covariance is equivalent to independence for binary variables, most of the paper's results show that the covariance will be non-negative or non-positive, guaranteeing neither dependence nor independence. Result 2 part 6 in that paper, however, is a (weaker)

formulation of the following statement: For binary A , B , and C , if A and B don't modify each other's effects on $C = c$ (on the probability ratio scale), then A and B will be independent conditional on $C = c$. A recent article [10] implicitly proved our theorem 1 in the restricted case of binary variables.

The case of binary variables was also explored, in retrospect, in the context of the case-only design [2]. It was shown that it is possible to estimate the magnitude of a multiplicative interaction between A and B (in our notation) in the stratum of cases ($C = 1$), assuming that A and B are independent in the population from which the cases arose [11–14]. Stated generically, two marginally independent binary causes do not modify each other's effect on one value of a binary outcome (on the probability ratio scale), if and only if they are independent conditional on that value. That statement, though not articulated [12–14], is a specific case of our general theorem.

Notwithstanding the case-only design, the link between effect modification on the multiplicative scale and the consequences of conditioning on a shared effect did not receive much attention, perhaps because the very concept of effect modification is still debated [15–18]. Moreover, methodological literature tends to favor a deterministic model, which downplays the concept of effect modification and is tightly connected to measuring effects on the additive scale. Whether the preference for difference measures of effect is justified is an open question, but the connection between effect modification and colliding bias may be one reason to prefer the multiplicative scale [19]. As shown here, if we are worried about colliding bias, we should look for heterogeneity of the probability ratio – not the probability difference – even if we eventually choose to estimate probability differences.

Some authors have suggested that A and B (in our notation) will “almost certainly” be associated in at least one stratum of C . Our results provide a formal justification of that phrase and a deeper insight into its meaning. The exceptions are instances in which there is precisely no effect modification on any value of C (e.g., Appendix B). Since no effect modification is just one point (the null) within an infinite range of possibilities, the phrase “almost certainly” is appropriate. Nonetheless, it may be improved. In general, we may say that A and B will “almost certainly” be associated in *every* stratum of C . And when C is binary, A and B will *certainly* be associated in at least one stratum of C (Corollary 1). Still, the phrase “almost certainly” is relevant only under a Bayesian framework, conveying a belief that effect modification is almost always present. (The same reasoning may be used to claim that effects are “almost certainly” never null.) Under a non-Bayesian framework, the state of affairs is fixed and independent of our beliefs: either effect modification is present or it is not.

The consequences of conditioning on a collider through regression do not immediately follow from our theorem. In linear regression, for instance, effect modification is defined by the presence of an interaction term, which is not comparable to our definition of effect modification except under a log probability model. Furthermore, it is unclear how conditioning on a continuous variable through regression corresponds to other methods of conditioning (e.g. restriction and stratification).

A couple of extensions of our work are possible: (1) The main result applies to discrete variables but should hold for continuous variables by replacing probabilities with probability densities (assuming a joint density function) and converting sums to integrals. It may be more complicated to prove the theorem when some variables are continuous and others are discrete. (2) The results are limited to the independent-dependent dichotomy. We conjecture that if A and B are marginally dependent (due to a common cause, for instance), the association between them will be altered upon conditioning if effect modification is present.

A Alternative proof of theorem 1

The proof below formalizes and extends the results for case-only studies [12, 13], utilizing a measure of effect modification that equals a measure of a conditional association between the colliding variables. The connection between two such measures was recently alluded to elsewhere [20].

Proof. Since A and B are marginally independent, the following holds for all i, j, k , and l ,

$$\begin{aligned}
 r_{ij}(b_l)/r_{ij}(b_k) &= \frac{P(C = c \mid A = a_j, B = b_l)}{P(C = c \mid A = a_i, B = b_l)} \bigg/ \frac{P(C = c \mid A = a_j, B = b_k)}{P(C = c \mid A = a_i, B = b_k)} \\
 &= \frac{P(C = c \mid A = a_j, B = b_l)}{P(C = c \mid A = a_i, B = b_l)} \bigg/ \frac{P(C = c \mid A = a_j, B = b_k)}{P(C = c \mid A = a_i, B = b_k)} \\
 &\quad \times \frac{P(A = a_j, B = b_l)}{P(A = a_i, B = b_l)} \bigg/ \frac{P(A = a_j, B = b_k)}{P(A = a_i, B = b_k)} \tag{4} \\
 &= \frac{P(C = c, A = a_j, B = b_l)}{P(C = c, A = a_i, B = b_l)} \bigg/ \frac{P(C = c, A = a_j, B = b_k)}{P(C = c, A = a_i, B = b_k)} \\
 &= \frac{P(A = a_j \mid B = b_l, C = c)}{P(A = a_i \mid B = b_l, C = c)} \bigg/ \frac{P(A = a_j \mid B = b_k, C = c)}{P(A = a_i \mid B = b_k, C = c)}
 \end{aligned}$$

$r_{ij}(b_l)/r_{ij}(b_k) = 1$ for all i, j, k , and l iff B and does not modify A 's effect on $C = c$. In the lemma below, we show that

$$\frac{P(A = a_j \mid B = b_l, C = c)}{P(A = a_i \mid B = b_l, C = c)} \bigg/ \frac{P(A = a_j \mid B = b_k, C = c)}{P(A = a_i \mid B = b_k, C = c)} = 1$$

for all i, j, k , and l iff A and B are independent conditional on $C = c$. It follows then by eq. (4) that A and B are independent conditional on $C = c$ iff A and B do not modify each other's effects on $C = c$. $\square$

Lemma 2.

$$\frac{P(A = a_j \mid B = b_l, C = c)}{P(A = a_i \mid B = b_l, C = c)} \bigg/ \frac{P(A = a_j \mid B = b_k, C = c)}{P(A = a_i \mid B = b_k, C = c)} = 1$$

for all i, j, k , and l if and only if A and B are independent conditional on $C = c$.

Proof. The “if” direction follows immediately upon realizing that $P(A = a \mid B = b, C = c) = P(A = a \mid C = c)$ for all a and b , if A and B are independent conditional on $C = c$.

For the “only if” direction,

$$\frac{P(A = a_j \mid B = b_l, C = c)}{P(A = a_i \mid B = b_l, C = c)} = \frac{P(A = a_j \mid B = b_k, C = c)}{P(A = a_i \mid B = b_k, C = c)} \tag{5}$$

for all i, j, k , and l . We can then sum both sides of eq. (5) over j to obtain

$$\frac{1}{P(A = a_i \mid B = b_l, C = c)} = \frac{1}{P(A = a_i \mid B = b_k, C = c)} \tag{6}$$

for all i, k , and l . Since i, k , and l are arbitrary, eq. (6) implies that $P(A = a \mid B = b, C = c)$ does not depend on b for any a . Therefore, $P(A = a \mid B = b, C = c)$ equals any weighted average of the $P(A = a \mid B = b_k, C = c)$. Specifically, for the weights $w_{b_k} = P(B = b_k \mid C = c)$

$$P(A = a \mid B = b, C = c) = \sum_{k=1}^m w_{b_k} P(A = a \mid B = b_k, C = c) = P(A = a \mid C = c)$$

for all a and b . That is, A and B are independent conditional on $C = c$. □

B Tables illustrating counterexample

We show below tables illustrating that A and B , two causes of C , can be independent in every stratum of C . Consider the next table, in which A and B are marginally independent.

Table 1: Marginal independence.

	$B = 0$	$B = 1$	Total
$A = 0$	90	90	180
$A = 1$	90	90	180
Total	180	180	

The following tables depict the effects of A on C , stratified on B , and the effects of B on C , stratified on A . The counts and effects correspond to the probabilities mentioned in the text.

Table 2: Effects of A on C when $B = 0$

Effects of A on C when $B = 1$

	$C = 1$	$C = 2$	$C = 3$	Total		$C = 1$	$C = 2$	$C = 3$	Total
$A = 0$	15	15	60	90	$A = 0$	20	30	40	90
$A = 1$	27	9	54	90	$A = 1$	36	18	36	90
Effects	1.8	0.6	0.9		Effects	1.8	0.6	0.9	

Table 3: Effects of B on C when $A = 0$

Effects of B on C when $A = 1$

	$C = 1$	$C = 2$	$C = 3$	Total		$C = 1$	$C = 2$	$C = 3$	Total
$B = 0$	15	15	60	90	$B = 0$	27	9	54	90
$B = 1$	20	30	40	90	$B = 1$	36	18	36	90
Effects	1.3	2	0.6		Effects	1.3	2	0.6	

As can be seen, there is no effect modification between A and B on any value of C . Using the tables above, we can depict the relation (counts, column percentages) between A and B in each stratum of C :

Table 4: Conditional independence

$C = 1$			$C = 2$			$C = 3$		
	$B = 0$	$B = 1$		$B = 0$	$B = 1$		$B = 0$	$B = 1$
$A = 0$	15 (36%)	20 (36%)	$A = 0$	15 (62.5%)	30 (62.5%)	$A = 0$	60 (53%)	40 (53%)
$A = 1$	27 (64%)	36 (64%)	$A = 1$	9 (37.5%)	18 (37.5%)	$A = 1$	54 (47%)	36 (47%)
Total	42	56	Total	24	48	Total	114	76

Evidently, A and B , which were marginally independent, are conditionally independent as well. No association was created between A and B in any stratum of C .

References

1. Greenland S. Quantifying biases in causal models: Classical confounding vs collider-stratification bias. *Epidemiology* 2003;14:300–6.
2. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37–48.
3. Ding P, Miratrix LW. To adjust or not to adjust? Sensitivity analysis of *M*-bias and butterfly-bias. *J Causal Infer* 2015;3:41–57.
4. Elwert F. Graphical causal models. In Morgan SL, editor. *Handbook of causal analysis for social research*. New York: Springer, 2013;250–1.
5. Shahar E, Shahar DJ. Causal diagrams and three pairs of biases. In Lunet N, editor. *Epidemiology – current perspectives on research and practice*, 2012:31–62. Available at: www.intechopen.com/books/epidemiology-current-perspectives-on-research-and-practice.
6. Greenland S, Pearl J, Robins JM. Confounding and collapsibility in causal inference. *Stat Sci* 1999;14:29–46.
7. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*, 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 2008.
8. Pearl J. *Causality: models, reasoning, and inference*. New York: Cambridge University Press, 2000:17.
9. VanderWeele TJ, Robins JM. Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol* 2007;166:1096–104.
10. Nguyen TQ, Dafoe A, Ogburn EL. Collider bias in binary variable structures, 2016. [arXiv:1609.00606](https://arxiv.org/abs/1609.00606).
11. Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat Med* 1994;13:153–62.
12. Yang Q, Khoury MJ, Sun F, Flanders WD. Case-only design to measure gene-gene interaction. *Epidemiology* 1999;10:167–70.
13. Schmidt S, Schaid DJ. Potential misinterpretation of the case-only study to assess gene-environment interaction. *Am J Epidemiol* 1999;150:878–85.
14. Jiang Z, Ding P. The directions of selection bias, 2017. [arXiv:1609.07834v2](https://arxiv.org/abs/1609.07834v2).
15. VanderWeele TJ. On the distinction between interaction and effect modification. *Epidemiology* 2009;20:863–71.
16. Shahar E, Shahar DJ. On the definition of effect modification. *Epidemiology* 2010;21:587.
17. Lawlor DA. Biological interaction: time to drop the term? *Epidemiology* 2011;22:148–50.
18. Weinberg CR. Interaction and exposure modification: Are we asking the right questions? *Am J Epidemiol* 2012;175:602–5.
19. Shahar DJ. Deciding on a measure of effect under indeterminism. *Open J Epidemiol* 2016;6:198–232.
20. Ding P, VanderWeele TJ. Sharp sensitivity bounds for mediation under unmeasured mediator-outcome confounding. *Biometrika* 2016;103:483–90.