

Colliding bias (part 1): misnomers and the missing dashed line

To set the stage for a close look at the literature on colliding bias, I will start with a series of questions and answers:

Q: What is colliding bias?

A: It is the bias that might arise after conditioning on a collider – a variable at which causal arrows collide (for example, $A \rightarrow C \leftarrow B$).

Q: Why “might arise”? Are there specific conditions under which the bias arises?

A: Yes, but the literature on this key point is surprisingly sketchy and sometimes inaccurate.

Q: What is conditioning on a variable?

A: Any alteration of the distribution of a variable, which also means gaining some knowledge about its value(s).

Q: What is explicit conditioning?

A: Restricting to one value, or stratifying on the values.

Q: What is implicit conditioning?

A: Categorization; truncating the distribution (deleting values but stopping short of restriction); or adding the variable to a regression model.

Q: Is this a common definition of “conditioning”?

A: No. The common definition is restricting a variable to one of its values (or stratifying on all values). I think, however, that the broader definition above is helpful. Any alteration of the distribution of the collider, including implicit conditioning through regression, can result in colliding bias. It does not have to be restriction or stratification.

Q: Are there subtypes of colliding bias?

A: Yes, two: bi-path ($A \rightarrow C \leftarrow B$) and uni-path ($A \rightarrow B \rightarrow C$). We may also distinguish between sampling colliding bias (where conditioning is done by physical selection of the study sample), and analytical colliding bias (where conditioning is done by computation).

Q: Examples?

A: Sampling colliding bias might arise when you recruit only stroke-free people. (Stroke victims are not allowed to participate in the study.) Analytical colliding bias might arise when you ignore stroke status at recruitment and then do one of the following: 1) restrict the analysis to the subset (stratum) of stroke-free people, or to the subset (stratum) of stroke victims; 2) compute a weighted average across the two strata; 3) fit a regression model to the entire sample and add stroke status as an independent variable.

Commentary

Q: What is the relation between confounding bias and colliding bias?

A: They make up a pair of biases with antithetical properties,¹ but most writers

continue to ignore that interesting feature. Some writers even subsume paths of colliding bias under confounding bias. (Years ago I have made the very same pedagogical mistake.)

Misnomers

In a footnote to a book chapter,² Elwert lists more than half a dozen alternative names for colliding bias. His list does not mention “colliding bias”, but includes an innovation of his own: “endogenous selection bias”.³ He writes:

⁶Terminology is in flux. The name “endogenous selection bias” highlights that the problem originates from conditioning on an endogenous variable. Others prefer “selection bias” (Hernán et al. 2004), “collider stratification bias” (Greenland 2003), “M-bias” (Greenland 2003), “Berkson’s [1946] bias,” “explaining away effect” (Kim and Pearl 1983), or “conditioning bias” (Morgan and Winship 2007). Simpson’s paradox (Hernán et al. 2011) and the Monty-Hall dilemma (Burns and Wieth 2004) involve specific examples of endogenous selection. The shared structure of some examples of

I argue that most of the names do not describe the multifaceted nature of the bias, and some of them are misnomers.

- “Selection bias” not only ignores the key phenomenon (colliding), but also perpetuates a misconception that the bias arises only due to sample selection.
- Elwert doesn’t like “selection bias”, but I have no idea what he had in mind when he added “endogenous”. Is there another type called “exogenous selection bias”? His “endogenous variable” is just a long name for “collider”.
- “Collider-stratification bias” obviously misses the point that the bias can arise by implicit conditioning. (There is no “stratification” when a continuous collider is added to a regression model.)
- “M-bias” or “M-collider bias” describes one of the structures of colliding bias in a causal diagram, and not even the simplest one. Where is the M-shape in the structure $A \rightarrow C \leftarrow B$ (with conditioning on C)?
- “Berkson’s bias” appeals to the first description of sampling colliding bias (1946), but historical credit is a weak justification.
- “Conditioning bias” is close to the target because the bias arises by conditioning. Another bias, however, can also arise by conditioning (causal-pathway bias¹). Furthermore, conditioning is also a method to *remove* colliding bias. An associational path that was created by conditioning on a collider can be blocked by conditioning on an intermediary.

Commentary

- “explaining away effect [bias]” in Elwert’s footnote is attributed to Kim and Pearl (1983),⁴ but that phrase does not show up in the 4-page article.^a The closest I found was the following sentence: “...the radio announcement [about an earthquake] reduces the likelihood of a burglary, as it ‘explains away’ the alarm sound.” That is fairly distant from a proposal to name the bias “explaining away effect”. In fact, the focus of that 1983 paper is not bias at all, but rather updating beliefs about deterministic causes, given the value of a shared effect. We will soon discover that poor citation is not uncommon in writing about colliding bias.

So, why is “colliding bias” my preferred name?

First, it alludes to the culprit: colliding arrows. Second, it is a natural analogue to confounding bias: If a confounder is the reason for confounding bias, then a collider is the reason for colliding bias. Yes, I know. It’s not the collider alone but rather “the collider and conditioning on the collider”. But it’s not the confounder alone, either. It’s “the confounder and not conditioning on the confounder”. Now, if you still hold to “collider-stratification bias”, please be consistent and rename confounding in a parallel fashion: “confounder-non-stratification bias”. Or, if you prefer “collider bias”,⁵ be sure to use “confounder bias” instead of “confounding bias”. Am I creating the impression that bizarre stubbornness prevents writers from adopting the most natural name for the bias that might arise from conditioning on a collider? I hope so.

^a Speaking of accuracy, Elwert (and a few other authors) also cite the title of the article incorrectly (adding the word “combined”):

Kim, J.H., & Pearl, J. (1983). A computational model for combined causal and diagnostic reasoning in inference systems. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence* (pp. 190–193). Karlsruhe.

Here is the actual title page of the article in question:

In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*,
Karlsruhe, West Germany, pp. 190-193, August 8-12, 1983.

TECHNICAL REPORT
R-25
May 1983

A COMPUTATIONAL MODEL FOR CAUSAL AND DIAGNOSTIC REASONING IN INFERENCE SYSTEMS¹

Jin H. Kim² and Judea Pearl

Cognitive Systems Laboratory
University of California, Los Angeles

Elwert’s citation replicates a citation on a website (http://bayes.cs.ucla.edu/csl_papers.html) that contains the same mistake:

(R-25) Kim, J. H. & Pearl, J., "A Computational Model for Combined Causal and Diagnostic Reasoning in Inference Systems," UCLA-ENG-CSL-83-03 (R-25), January 1983; Rev. I, May 1983; Proceedings, IJCAI-83, 190-193, 1983.

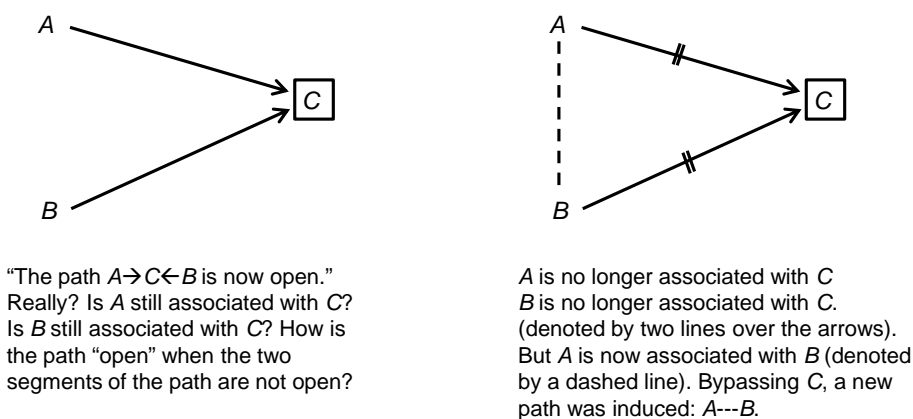
Commentary

The dashed line

If I were able to dictate only one aspect of causal diagrams to writers on this topic, it would have been the dashed line -- the line that indicates a newly formed association, or a new associational component, following conditioning on a collider. Yes, you can find the dashed line here and there – even in an early publication on causal diagrams^b – but more often than not, it does not show up in methods papers.

Many writers end the story of colliding bias by drawing a box around a collider to denote conditioning. “The path through the collider is now open (or active)”, they explain. Where exactly is it “open”? Are *A* and *B* still associated with *C* after restriction to a single value of *C*? No, they are not. So, how is the path “open at *C*” when neither of its segments ($A \rightarrow C$; $C \leftarrow B$) is “open” anymore? Draw lines over the colliding arrows to show *dissociation from C*. Then, put a dashed line between *A* and *B* (the colliding variables) and it is all clear (Figure 1): bypassing the collider, a new path was induced. Stubbornness again? Maybe. But unlike name-the-bias game (surely associated with the names behind the names), omission of the dashed line is simply a sloppy illustration of colliding bias. Moreover, the next section shows that there is something to be learned from thinking about *where* the line(s) should be drawn.

Figure 1.



^b Shown below is the dashed line in one of the figures in the first publication on causal diagram in the epidemiology literature.⁶ Unfortunately, the authors did not display anything to indicate conditioning on *C*, and dissociation of *A* and *B* from *C*.

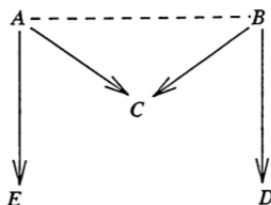


FIGURE 9.

Commentary

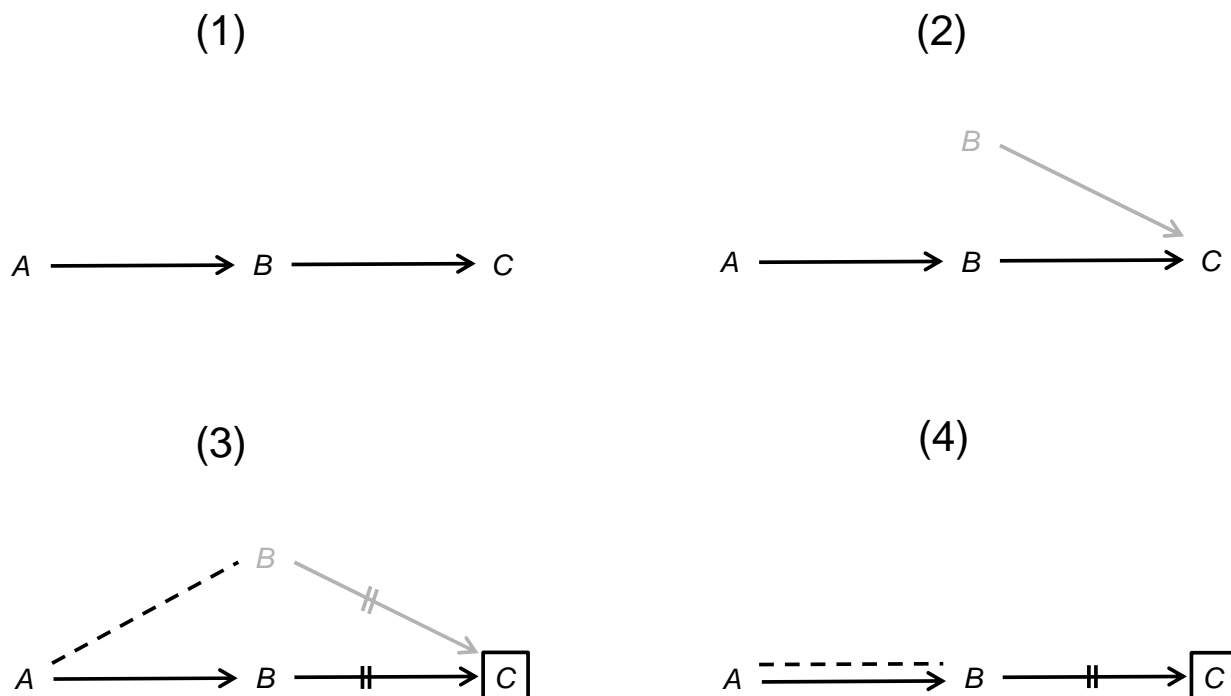
Where to draw the dashed line

Many properties of colliding bias are antithetical to those of confounding bias.¹ Here is one example: Confounding bias is always present in the presence of a single confounder ($A \leftarrow C \rightarrow B$), whereas colliding bias does not always arise after conditioning on a single collider – even in the structure $A \rightarrow C \leftarrow B$ (!) When does colliding bias arise? When should we draw a dashed line? The answer is on the long side, and the underlying theorems are yet to be stated and proved. You will see later (part 2 of the commentary) a sample of sloppy statements in lieu of theorems, and a maze of citations. Quite amazing, I think.

For the time being, let's focus on correct placement of the dashed line(s) — whenever associations are indeed created or altered following conditioning on a collider.

I will start with uni-path colliding,^{1,7,8} the basic structure, which is still widely ignored: $A \rightarrow B \rightarrow C$. Just like any two causes of C , the variables A and B collide at C ($A \rightarrow C$; $C \leftarrow B$), albeit through a shared terminal arrow. Placing a box around C should be coupled with adding a dashed line between A and B because conditioning on C adds a new component to the association between A and B .^c Figure 2 provides an intuitive explanation of uni-path colliding bias (diagrams 1 through 4).

Figure 2.



^c The association between A and B is not altered (a dashed line should not be drawn) if that association is quantified by the odds ratio. An intuitive explanation and a proof can be found on this website (a commentary on the case-control study).

Commentary

If Figure 2 is not convincing enough, you might like the following heuristic example. Suppose A is smoking (yes/no), B is cancer (yes/no), and C is vital status ($A \rightarrow B \rightarrow C$). On an imaginary planet, the following is true:

- Causation is indeterministic (probabilistic)
- Cancer is the only cause of death
- The effect of cancer on death is very strong, nearly deterministic
- There are no shared causes of smoking and cancer (no confounders)

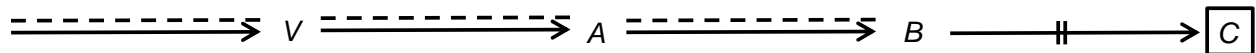
To estimate the effect of smoking on cancer, we conducted a cohort study on that planet and computed the probability ratio, $\Pr(CANCER=yes|SMOKING=yes)/\Pr(CANCER=yes|SMOKING=no)=50$: the probability of cancer in smokers is 50 times the probability of cancer in non-smokers.

Alternatively, we studied the same effect (smoking \rightarrow cancer) much more quickly using only death certificates from that planet. In other words, we conditioned on vital status (C) by restricting the study to $C=dead$. We try again to compute the effect of smoking on cancer. What do we find? Since cancer is the only cause of death on that planet, and its effect is very strong, almost all members of our sample (who are dead) must have had cancer. As a result, the probability of cancer in our sample is nearly 1 – for both smokers and non-smokers – which means that the probability ratio (smokers versus non-smokers) ≈ 1 . The strong association between smoking and cancer was almost nullified. Conditioning on vital status changed the association between smoking and cancer such that the causal component (smoking \rightarrow cancer) plus some new component (smoking \dashrightarrow cancer) added up to a nearly null association. Under less stringent conditions, the “sum” may be very different from null.

To sum up: If we wish to estimate the effect of A on B , and $A \rightarrow B \rightarrow C$, conditioning on C creates uni-path colliding bias (Figure 2). The dashed line between A and B , along with the arrow, indicate that the association between A and B contains two components: the causal path $A \rightarrow B$ (which we try to estimate) and an induced path $A \dashrightarrow B$ (colliding bias).^d

Figure 3 shows an extension to a causal path that is antecedent to A . The dashed line appears between every pair along the path.

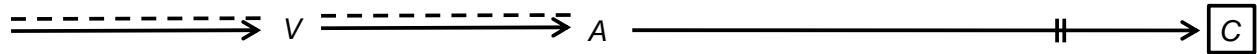
Figure 3.



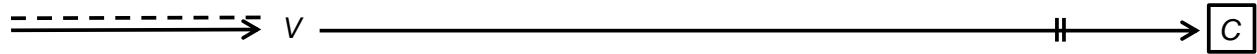
^d Uni-path colliding bias does not arise under the precise null (for any measure of effect). If the effect $B \rightarrow C$, or the effect $A \rightarrow C$ (through B), is precisely null, the variables A and B do not collide at C through causal arrows. Stated differently, when the arrows correspond to null effects, the dashed line corresponds to null contribution to the association...

Commentary

Intuition? If we didn't know about B , we would have drawn a dashed line between V and A , as shown below. Of course, $V \dashrightarrow A$ does not substitute for $A \dashrightarrow B$. Both associations are altered following conditioning on C .



And if we didn't know about *both* A and B , we would have drawn...



We have already seen the next structure (Figure 4), which depicts colliding through separate paths (bi-path colliding¹). If conditioning on C indeed creates an association between A and B , there is no question where to draw the dashed line.

Figure 4.

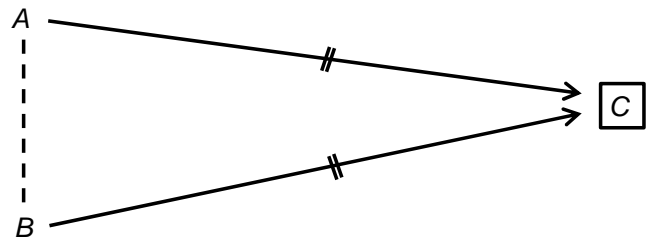
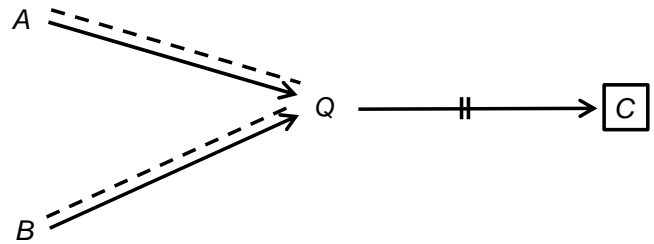


Figure 5 is more challenging. If uni-path colliding is present, a dashed line is added twice: between A and Q ($A \rightarrow Q \rightarrow C$) and between B and Q ($B \rightarrow Q \rightarrow C$). Should we also draw a dashed line between A and B ? No. The two lines do the job, depicting the induced path $A \dashrightarrow Q \dashrightarrow B$ which shows that A and B are now associated. Two dashed lines parsimoniously describe the three associations that were altered or created (between A and Q ; between B and Q ; and between A and B). Another line between A and B is redundant.

Figure 5.



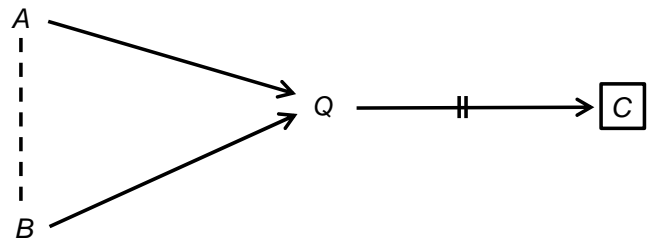
Suppose, however, we didn't know about Q , a common intermediary through which A and B affect C . Then what? Then we would have drawn Figure 4, and no harm would have been done. The dashed line between A and B

substitutes for the two dashed lines in Figure 5. As far as A and B are concerned, the structure $A \dashrightarrow B$ is equivalent to the structure $A \dashrightarrow Q \dashrightarrow B$. Both show an open induced path.

Commentary

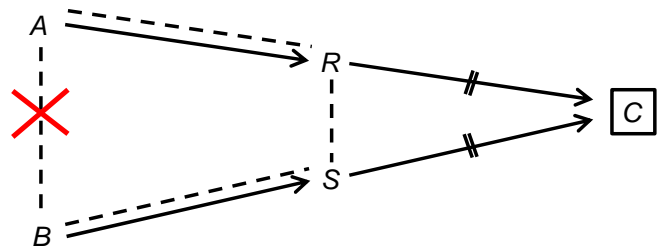
Another possibility is shown in Figure 6. A and B still affect C through Q , but uni-path colliding bias is absent (say, associations are quantified by the odds ratio). If conditioning on C still induces an association between A and B , we draw a dashed line, $A\text{---}B$ (Figure 6), just as we did in Figure 4.

Figure 6.



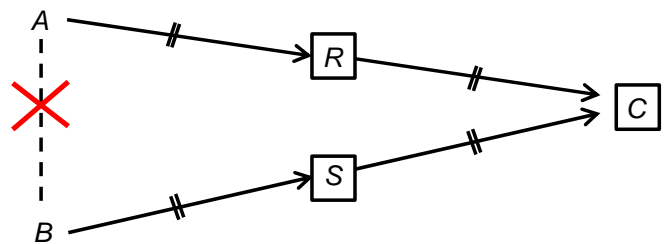
Lastly, Figure 7 is an extension of Figure 4, adding two intermediaries (R, S). Three dashed lines are self-explanatory: $R\text{---}S$ (bi-path colliding); $A\text{---}R$ (uni-path colliding); and $B\text{---}S$ (uni-path colliding). Should we also draw a line between A and B ? No, we should not. Is it redundant? No, it is wrong. The line $A\text{---}B$ does not exist at all, as illustrated by the red X.

Figure 7.



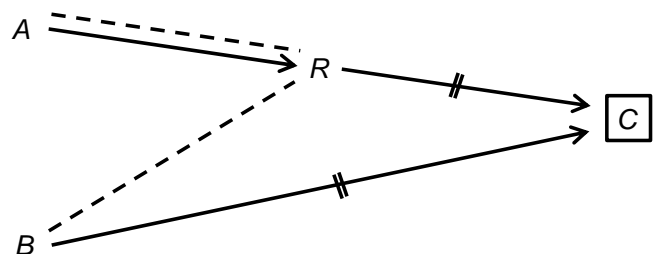
How do we know that $A\text{---}B$ does not exist? How do we know that A and B are associated *only* through the induced path $A\text{---}R\text{---}S\text{---}B$? Here is the reasoning: If $A\text{---}B$ is true, then A and B should still be associated *after blocking the path* $A\text{---}R\text{---}S\text{---}B$. But that's not true. A and B are no longer associated after conditioning on R and S (for instance), because they are now dissociated from the collider C (Figure 8). Effects that do not translate to associations ($A \rightarrow R \rightarrow C$; $B \rightarrow S \rightarrow C$) do not contribute to any association.

Figure 8.



Notice that if we didn't know about R and S , we are back in Figure 4. And if we didn't know about one of them, say S , we would have drawn Figure 9: $A\text{---}R$ (uni-path colliding); $R\text{---}B$ (bi-path colliding). Most important, the paths $A\text{---}R\text{---}B$ (Figure 9); $A\text{---}R\text{---}S\text{---}B$ (Figure 7), $A\text{---}Q\text{---}B$ (Figure 5); and $A\text{---}B$ (Figures 4 and 6) convey the same information about A and B . Each of them reveals that conditioning on C adds colliding bias – if we wish to estimate the effect of A on B (whether null or otherwise).

Figure 9.



Commentary

Concluding Example

Figure 10 shows a causal structure. Suppose we condition on C and D , as shown. What are the consequences for estimating the effect of A on B (or the effect of any other variable on another) by the probability ratio?

If Figure 10 is all that you draw, you will state “the paths are now open (or active) at C and D ”, and then start thinking what the statement implies.

If you add dashed lines – and don’t forget about uni-path colliding! – you will draw Figure 11.

Now it is all clear. You can trace all the paths between A and B , or between any two variables. Some paths are composed of arrows alone (natural paths^{7,8}); others are composed of dashed lines, or of dashed lines in some segments and arrows in other segments (induced paths^{7,8}). Some paths between A and B fully or partially overlap. Whenever an arrow and a dashed line overlap, each provides a path of its own (like two different roads between two towns). For example, $A \rightarrow B$ and $A \dashrightarrow B$ overlap, and so do the paths $A \leftarrow E \dashrightarrow F \dashrightarrow G \rightarrow B$ and $A \dashrightarrow E \dashrightarrow F \dashrightarrow G \dashrightarrow B$. Some paths are open (e.g., $A \rightarrow H \dashrightarrow F \dashrightarrow G \dashrightarrow B$), which means “contribute to an association between the two ends”; others are blocked (e.g., $A \dashrightarrow H \rightarrow D \leftarrow B$). For any two variables, it is now clear which associations are created or altered.

Figure 10.

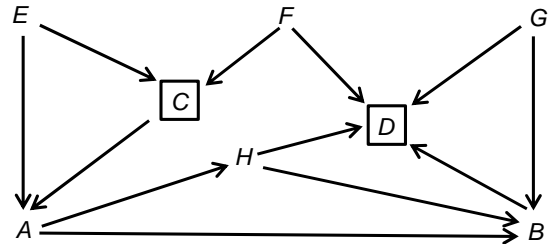
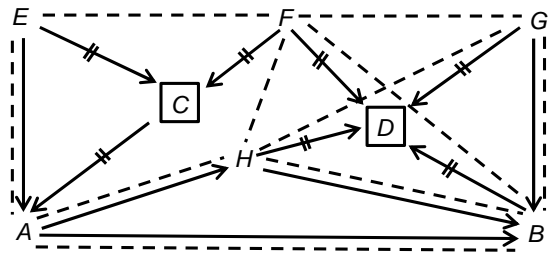


Figure 11.



Any open, induced path between the cause-and-effect of interest is synonymous with colliding bias. Try to count at least 10 such paths between A and B . (Hint: there are three overlapping, induced paths when you travel the road $A-H-B$ alone.)

Commentary

References:

1. Shahar E, Shahar DJ. Causal diagrams and three pairs of biases. In: *Epidemiology –Current Perspectives on Research and Practice* (Lunet N, Editor). www.intechopen.com/books/epidemiology-current-perspectives-on-research-and-practice, 2012: pp. 31-62
2. Elwert F. Graphical causal models. In: *Handbook of Causal Analysis for Social Research* (Morgan SL, editor) Springer 2013
3. Elwert F, Winship C. Endogenous selection bias: the problem of conditioning on a collider variable. *Annual Review of Sociology* 2014;40:31-53
4. Kim JH, Pearl J. A computational model for causal and diagnostic reasoning in inference systems. In: *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, West Germany, pp. 190-193, Technical Report R-25, May 1983
5. Del Junco DJ, Bulger EM, Fox EE, Holcomb JB, Brasel KJ, Hoyt DB, Grady JJ, Duran S, Klotz P, Dubick MA, Wade CE; ROC Investigators. Collider bias in trauma comparative effectiveness research: the stratification blues for systematic reviews. *Injury* 2015;46:775-80
6. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37-48
7. Shahar E, Shahar DJ. Causal diagrams and the logic of matched case-control studies. *Clinical Epidemiology* 2012;4:137–44. *Clinical Epidemiology* 2014;6:59 [Corrigendum]
8. Shahar E, Shahar DJ. Causal diagrams and the cross-sectional study. *Clinical Epidemiology* 2013;5:57–65

Colliding bias (part 2): missing theorems and missing proofs (forthcoming).

Here is a preview of the first section.

Commentary

Colliding bias (part 2): missing theorems and missing proofs [preview]

The heuristic argument

In part 1, we saw a heuristic example for uni-path colliding. Let's see the heuristic argument for bi-path colliding. To simplify, I will use deterministic jargon.

Suppose there are only two causes, A and B , of some outcome, C , and all three variables are binary. Assume that A and B are independent – that is, knowing the value of one tells us nothing about the value of the other. This state of affairs is perfectly encoded by the diagram $A \rightarrow C \leftarrow B$, in which C is a collider. According to this diagram, A and B are not associated because neither is a cause of the other, and they don't share any cause.

Assume we code the values of the three variables such that a value of "1" for A or B (the causes) will result in a value of 1 for C as well (the effect). For example: if A is trauma status, B is cancer status, and C is vital status, then $A=1$ denotes "trauma", $B=1$ denotes "cancer", and $C=1$ denotes "dead". Obviously, if $A=1$ (trauma) or $B=1$ (cancer), then $C=1$ (dead). The table below shows all permutations of the values of A , B , and C , which may be observed under these terms.

Values of the causes	Value of the effect
$A=0$	$C=1$
$B=1$	
$A=1$	$C=1$
$B=0$	
$A=1$	$C=1$
$B=1$	
$A=0$	$C=0$
$B=0$	

Recall that in our hypothetical story, A and B are independent (not associated). What will happen, however, after explicit conditioning on C ? What happens to that independence, if we know, for example, that $C=1$ (restriction)? Well, things change. Now, knowing the value of one cause of $C=1$ might inform us about the value of the other cause. For instance, if we know that $A=0$, then $B=1$. Otherwise, what has caused C to take the value $C=1$? (If both $A=0$ and $B=0$, C must take the value $C=0$). Reverting to associational language, we see that a null association between A and B before conditioning, $\Pr(B=1|A=0)/\Pr(B=1|A=1)=1$, turned into a non-null association after conditioning on C : $\Pr(B=1|A=0, C=1)=1$ whereas $\Pr(B=1|A=1, C=1) \neq 1$ (implying: probability ratio $\neq 1$). To sum up, explicit conditioning on C (that is, knowing the value of C) results in dependence between A and B . The two causes of C are associated – conditional on C .

Commentary

I used the triad “trauma-cancer-dead” as an example for $A=1$, $B=1$, and $C=1$, but anyone can make up other examples, provided they satisfy the diagram $A \rightarrow C \leftarrow B$. The next table shows several other triads that you can find in the literature on colliding bias. Pick some of them and tell yourself the story that I told you in the previous paragraph. Convince yourself that you can make some guesses about the value of B from the value of A and vice versa, provided you know the value of C . It is fairly simple.

Causes (A, B)	Effect ($C=1$)	Effect ($C=0$)
<i>Sprinkler</i>	<i>Wet lawn</i>	<i>Dry lawn</i>
<i>Rain</i>		
<i>Burglary</i>	<i>Alarm sounds</i>	<i>Alarm doesn't sound</i>
<i>Earthquake</i>		
<i>Battery</i>	<i>Car starts</i>	<i>Car doesn't start</i>
<i>Fuel</i>		
<i>Aspirin</i>	<i>Alive</i>	<i>Dead</i>
<i>Statin</i>		
<i>Talent</i>	<i>Success in Hollywood</i>	<i>Failure in Hollywood</i>
<i>Beauty</i>		
<i>Grades</i>	<i>Admitted to graduate school</i>	<i>Not admitted to graduate school</i>
<i>Musical talent</i>		

In search of proofs (or: the unsolvable citation maze)

I have a long list of questions, which might have crossed your mind, too: Heuristic arguments aside, where is the formal proof that conditioning on a collider creates or alter an association between the colliding variables? Are there any necessary or sufficient conditions for that statistical phenomenon? What happens if C takes three values, rather than two, or if C is a continuous variable? Does it matter which measure of association is used? What is the relevance of the heuristic argument above to the consequences, for example, of adding a collider to a linear regression model? In short, are there theorems on colliding bias that have been stated and proved?

You are about to be surprised by the answers.