

# The Relation of Collapsibility and Confounding to Faithfulness and Stability

Mohammad Ali Mansournia<sup>a</sup> and Sander Greenland<sup>b</sup>

**Abstract:** A probability distribution may have some properties that are stable under a structure (e.g., a causal graph) and other properties that are unstable. Stable properties are implied by the structure and thus will be shared by populations following the structure. In contrast, unstable properties correspond to special circumstances that are unlikely to be replicated across those populations. A probability distribution is faithful to the structure if all independencies in the distribution are logical consequences of the structure. We explore the distinction between confounding and noncollapsibility in relation to the concepts of faithfulness and stability. Simple collapsibility of an odds ratio over a risk factor is unstable and thus unlikely if the exposure affects the outcome, whether or not the risk factor is associated with exposure. For a binary exposure with no effect, collapsibility over a confounder also requires unfaithfulness. Nonetheless, if present, simple collapsibility of the odds ratio limits the degree of confounding by the covariate. Collapsibility of effect measures is stable if the covariate is independent of the outcome given exposure, but it is unstable if the covariate is an instrumental variable. Understanding stable and unstable properties of distributions under causal structures, and the distinction between stability and faithfulness, yields important insights into the correspondence between noncollapsibility and confounding.

(*Epidemiology* 2015;26: 466–472)

It is well known that the absence of confounding does not correspond to collapsibility of effect measures, such as odds ratios, rate ratios, and other measures that are not linear or log-linear functions of the underlying incidence propor-

tions (outcome risks) or survival times.<sup>1–6</sup> Noncollapsibility and confounding have been contrasted using both potential-outcome (counterfactual) causal models and directed acyclic graphs.<sup>3,4,6–10</sup> The distinction between them can be explained in terms of d-separation and d-connection (which we will call separation and connection) of variables in the directed acyclic graph,<sup>7</sup> and they can be distinguished and measured using a number of formulas.<sup>6</sup>

We review the distinction in relation to the concepts of faithfulness, stability, and plausibility of distributional properties relative to a given structure, such as a directed acyclic graph. We show that

- (1) Stability and faithfulness to a directed acyclic graph are distinct concepts; each can occur with or without the other.
- (2) Without artificial assumptions, collapsibility of odds ratios over a risk factor is always unstable (and hence unlikely if not impossible) when the exposure affects the outcome, whether or not the risk factor is a confounder.
- (3) When the exposure is binary with no effect, collapsibility over a confounder implies directed acyclic graph unfaithfulness (i.e., independence between connected variables).
- (4) Simple odds-ratio collapsibility sharply limits confounding of the risk ratio by the covariate.
- (5) Collapsibility is stable if the covariate is independent of the outcome given exposure.
- (6) Collapsibility is unstable if the covariate is an instrumental variable for an unmeasured confounder.

We illustrate the distinction between instability and unfaithfulness using odds-ratio collapsibility examples. The distinction illustrates limitations of causal directed acyclic graph models in representing constraints beyond those implied by missing arrows (e.g., collapsibility, homogeneity, matching, and balancing). The instability of simple odds-ratio collapsibility underscores one of the logical defects in equating odds-ratio collapsibility with no confounding. In particular, it warns against the “change in estimate” criterion for identifying confounders when using odds ratios for a common outcome, because this practice can mistakenly identify a non-confounding risk factor as a confounder. Parallel results apply to rate ratios, rate differences, and noncollapsible regression coefficients, such as those from logistic or Cox models.<sup>3</sup>

Submitted 13 February 2014; accepted 8 March 2015.

From the <sup>a</sup>Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran; and <sup>b</sup>Department of Epidemiology, UCLA School of Public Health, and Department of Statistics, UCLA College of Letters and Science, University of California, Los Angeles, CA.

**SDC** Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article ([www.epidem.com](http://www.epidem.com)). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the authors.

Correspondence: Mohammad Ali Mansournia, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, P.O. Box 14155-6446, Tehran, Iran. E-mail: [mansournia\\_ma@yahoo.com](mailto:mansournia_ma@yahoo.com).

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 1044-3983/15/2604-0466

DOI: 10.1097/EDE.0000000000000291

We assume that the reader is familiar with basic causal directed acyclic graph concepts and their relation to confounding, as reviewed in many sources and illustrated in Figures 1–6.<sup>8,9,11,12</sup> We depart from that literature, however, insofar as we distinguish two concepts, faithfulness and stability, which are sometimes treated as identical.<sup>8</sup> Throughout, we discuss only superpopulation (distributional) structures and expected estimates; elsewhere<sup>13</sup> we discuss the relation of random variation to design structures.

## REVIEW OF BACKGROUND CONCEPTS

### Compatibility Versus Faithfulness

A probability distribution for a set of variables is *compatible* with a directed acyclic graph (whether causal or not) if every pair of variables that are separated on the graph are independent in the distribution. *Faithfulness*, the converse property of compatibility, means that every pair of variables independent in the distribution are separated in the directed acyclic graph.<sup>8,11,12,14</sup> Compatibility is essential when using a distribution for analysis under a given directed acyclic graph, because biases are transmitted via graphical connections that must be reflected as associations in the distribution. In contrast, under a given directed acyclic graph, validity of tests for effects is preserved by compatibility but does not require faithfulness;



FIGURE 1. Simple odds ratio collapsibility over C is unstable if C and E separately affect D.



FIGURE 2. Collapsibility over C is unstable when C affects E and D.



FIGURE 3. For binary E with no effect, collapsibility over C requires unfaithfulness when C affects E and D.



FIGURE 4. Effect measures are collapsible over C if C is independent of D given E.

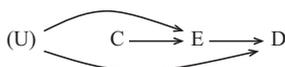


FIGURE 5. Collapsibility over C is unstable when C is an instrumental variable (U is unmeasured).



FIGURE 6. For binary E with no effect on D, collapsibility over C requires unfaithfulness when both E and D affect C.

furthermore, consistent variance estimation may require using unfaithful independencies induced by the study design.<sup>15,16</sup>

The directed acyclic graph itself summarizes a set of sharp prior independence assumptions that define distributions compatible with the structure encoded by the graph (because these assumptions represent a priori constraints, a directed acyclic graph representation of a compatible distribution is sometimes called a *Bayesian belief network* or *Bayes net*<sup>17</sup>). In particular, the absence of an arrow between two variables in a directed acyclic graph corresponds to a sharp constraint on the distributions compatible with the graph, reducing the dimensionality (degrees of freedom) of compatible distributions. Arrow absence thus represents special status (probability 1, or certainty) that the constraint holds.

An example is a simple randomized trial of the effect of a treatment E on an outcome D in the presence of a baseline risk factor C, shown in Figure 1. The arrow between C and E is absent because the randomization assumption implies the absence of any causal effect of covariates on treatment. In contrast, unfaithfulness occurs when there is a further constraint beyond those implied by the absence of certain arrows.

### Collapsibility

For binary E and D, let  $OR_{ED}$  represent the unconditional (unadjusted) odds ratio relating E to D and  $OR_{ED|c}$ , the conditional odds ratio relating E to D given C equals c. When  $OR_{ED|c}$  is assumed constant across levels of C (odds-ratio homogeneity), *simple* or *strict collapsibility* is defined as  $OR_{ED} = OR_{ED|c}$ .<sup>4,18,19</sup> When E or D have multiple levels, there will be multiple odds ratios relating E to D and this definition requires  $OR_{ED} = OR_{ED|c}$  for all of them. More generally, for a measure M of the association of E and D that is constant across C (such as a regression coefficient), one may define simple collapsibility as equality of the measure  $M_{DE}$  computed ignoring C, and the constant C-specific measures  $M_{DE|c}$ . Inverse-variance weighted averaging of estimated  $M_{DE|c}$  generally assumes that the measures are homogeneous, as in simple collapsibility.

In a related but distinct concept, we call *marginal collapsibility*, the unadjusted measure  $M_{DE}$  equals the adjusted summary measure  $M_{DE|p(C)}$  derived by averaging (marginalizing or standardizing) E-specific outcomes over the total population distribution  $p(C)$  of C,<sup>7,20</sup> as in ordinary inverse-probability weighting.<sup>21</sup> Simple and marginal collapsibility are equivalent for risk ratios, survival-time ratios, mean differences, and other measures called “collapsible,” but they are not equivalent for odds ratios, rate ratios, and other “noncollapsible” measures.<sup>4</sup> Other definitions of collapsibility can be found in the statistics literature, and collapsibility can be defined for any kind of adjustment for C, but in what follows we will focus on simple and marginal collapsibility.

### Risk Factors, Confounding, and Confounders

There is enormous variation in definitions of risk factors, confounding, and confounders, with considerable mismatch across texts and authors.<sup>4–8,10</sup> For simplicity, we will use “risk factor”

as a shorthand for independent causal risk factors (covariates that affect D through a pathway not involving E, as with C in Figures 1–3) or covariates that exhibit temporal associations compatible with such factors as C in Figure 5.

Some definitions of confounding are equivalent to violation of ignorability of treatment assignment or exchangeability of treatment groups, which in a causal directed acyclic graph corresponds to confounding paths from E to D (undirected open paths from E to D that end with an arrow into D). We will take confounding of an effect measure  $M_{ED}$  by a particular covariate C as a bias in estimating  $M_{ED}$  due to indirect connections of E and D transmitted through effects of C on D. In that case, we will call C a confounder and say it confounds  $M_{ED}$ .<sup>4,9,11,12</sup> By this definition, C cannot be a confounder of any measure under Figures 1 and 4. Under Figures 2 and 3, however, C will always confound some effect measure,<sup>8</sup> and as we will discuss C will usually confound all marginal effect measures.

As with most the current causal-inference literature, we do not address random confounding,<sup>4</sup> leaving that for a companion paper.<sup>13</sup> We also do not consider monotonicity restrictions (apart from the extreme case of homogeneity), which can render impossible certain types of rule exceptions.<sup>22,23</sup>

### FAITHFULNESS VERSUS STABILITY

Unfaithful distributions satisfy additional independence constraints not needed for compatibility. The plausibility of such constraints is important for determining whether unfaithfulness should be of practical concern. To formalize plausibility ideas, we will say a property (such as a constraint) is *unstable* relative to a structure if the structure does not induce (imply) the property. Such properties extend beyond independencies to include parametric assumptions such as homogeneity of an effect measure. Unstable properties may be deemed highly implausible, unlikely, artificial, or contrived if the structure encodes all the available information on causal linkages among the variables; in eAppendix 1 (<http://links.lww.com/EDE/A903>), we provide one formalization of the concept of artificiality by identifying it with any set of dimension-reducing constraints that are not implied by the structure.

Pearl<sup>18</sup>(Sect. 2.4) initially defines the term “stability” as a synonym for “faithfulness,” whereas in our usage faithfulness is a property of a distribution, whereas stability is a property of a property (e.g., as we will show, collapsibility may be stable or unstable). A faithful distribution is thus one for which all independencies are stable. The converse does not hold, however: Cohort matching and blocked randomization induce stable independencies that are unfaithful to the basic causal directed acyclic graph for the starting cohorts.<sup>13,16</sup> Faithful distributions may also have important unstable properties such as homogeneity and collapsibility.

We think our usage better conforms to Pearl’s verbal description of the concept that underlies stability and his usage in the context of confounding.<sup>8</sup>(Sect. 6.4) A stable property

is one induced by (deducible from) the causal mechanism, (i.e., implied by the causal structure). An unstable property is then accidental and unlikely to be replicated in other populations obeying the same diagram or unlikely to be maintained in the current population over time. For these reasons, some causal-graph theorists adopt faithfulness as a core assumption for methodologic development,<sup>14</sup> although practical<sup>11</sup> and theoretical<sup>24</sup> reservations against doing so have been given (e.g., an unfaithful independency may reflect some unmodeled but real causal process, or an intentional design strategy<sup>13,16</sup>).

### COLLAPSIBILITY, FAITHFULNESS, AND STABILITY

For a binary C, simple odds-ratio collapsibility over C ( $OR_{ED} = OR_{ED|c}$ ) occurs if and only if either C and E are independent given D ( $OR_{CE|d} = 1$  for all values d of D) or C and D are independent given E ( $OR_{CD|e} = 1$  for all values e of E).<sup>25</sup> If C is not binary, there are exceptions in which there is collapsibility without either independency, characterized by cancelations across levels of C.<sup>19</sup>(Tables 4, 6) These exceptions do not arise from hidden independencies between the variables illustrated on a directed acyclic graph, but rather from independencies among compound events,<sup>26,27</sup> and thus can arise from distributions faithful to the directed acyclic graph. Nonetheless, these exceptions involve strict additional constraints beyond those imposed by the graph, and thus the collapsibility they exhibit is unstable given only the directed acyclic graph.

Parallel observations apply to other measures, including those ordinarily labeled “collapsible” effect measures because they are collapsible under marginal CE independence ( $OR_{CE} = 1$ ). For example, the risk ratio can be collapsible over nonbinary C despite a marginal CE association and dependence of D on E given C.<sup>18</sup> As discussed below, however, risk-ratio collapsibility can be induced by simple design strategies, and thus can be stable, whereas odds-ratio collapsibility is always unstable when both C and E affect D.

### NONCONFOUNDING BY A RISK FACTOR ALMOST ALWAYS IMPLIES ODDS-RATIO NONCOLLAPSIBILITY OVER THE RISK FACTOR

As is well known,<sup>4–8</sup> without causal restrictions such as from a causal directed acyclic graph, noncollapsibility over C refers only to a difference in association with and without C adjustment and should not be equated with confounding. We illustrate and apply the above concepts to delineate further the distinction between noncollapsibility and confounding in terms of what we should usually expect of the underlying structure under study. eAppendix 1 (<http://links.lww.com/EDE/A903>) provides a more detailed discussion and formalization of what we mean by “usually expect” and “almost always” using mathematical concepts.

Our first point is that researchers routinely attempt to engineer unconfounded estimates using methods that we should expect to leave residual noncollapsibility. Consider

simple randomization of E. This design feature cuts off the effect of C on E by placing E under experimental control, and leads to Figure 1. On average over randomizations, C and E are unconditionally independent; thus the ED risk ratio is collapsible over C<sup>18</sup> (as is the risk difference) and C is not a confounder in the average sense used in definitions that equate confounding with violations of ignorability (sometimes expressed by saying the treatment-assignment mechanism is unconfounded<sup>28</sup>).

Nonetheless, under Figure 1, C will be associated with E given D; this can be seen by noting that D is a collider on the only path from C to E, and conditioning on it opens that path.<sup>7</sup> Under Figure 1, C also remains associated with D given E. Thus, the two conditional associations necessary for noncollapsibility of the ED odds ratio are present; these conditions also preclude simple odds-ratio noncollapsibility when C is binary.<sup>25</sup> This noncollapsibility in the absence of confounding was demonstrated analytically long ago and shown to be toward the null,<sup>1</sup> but it was often interpreted as a bias in OR<sub>ED</sub>.<sup>29,30</sup> However, it is not a bias if the marginal causal effect of E on D is the target.<sup>1,2,4,7</sup> A numerical example is provided in Table 1, in which under standardization to the total<sup>31</sup> all the measures are marginally collapsible, yet the odds ratio is not simply collapsible; in fact under Figure 1 with binary D, this must be the case.

Table 4 in Whittemore<sup>19</sup> with F<sub>1</sub> = C, F<sub>2</sub> = E, F<sub>3</sub> = D shows that with polytomous outcomes, simple odds-ratio collapsibility is possible under Figure 1 (although when D is polytomous, multiple ED odds ratios are involved). Nonetheless, such examples are unstable given only Figure 1. Thus, we expect odds ratios to be noncollapsible whenever confounding is controlled but certain risk factors are averaged over or not controlled more than needed to remove confounding. Consider studies of clusters (such as schools or households) with individual data, in which confounders are controlled in a model for correlated outcomes. With sufficient adjustment, there will be no confounding, but the population-average odds

ratio (as produced by, say, familiar GEE model fitting) will be closer to the null than the subject-specific odds ratio (as produced by random cluster-effects models), to the extent the outcome is common and clusters predict risk conditional on the covariates.<sup>32</sup>

Similarly, summary odds ratios derived from typical propensity-score adjustments (such as score-stratified or inverse-probability weighted estimates) will tend to be closer to the null than the covariate-conditional odds ratios derived from direct outcome regression even if both odds ratios are unconfounded.<sup>33</sup> This difference depends on the size of the outcome risk and how strongly the covariates predict risk conditional on the propensity score.<sup>34</sup>

### CONFOUNDING BY C ALMOST ALWAYS IMPLIES NONCOLLAPSIBILITY OF EFFECT MEASURES OVER C

A shared consequence of causality-based definitions of confounding by C is that for C to confound, it must be unconditionally associated with E and associated with D given E. In a directed acyclic graph with only C, E, and D, these associations require arrows from C to E and D, as in Figure 2; conversely, C will be a confounder under Figure 2, apart from exceptions with polytomous C involving special cancellations that would be considered unstable given only the diagram.

In eAppendix 2 (<http://links.lww.com/EDE/A903>), we show that if the odds ratio is simply collapsible, the unconditional (“crude”) risk ratio RR<sub>ED</sub> will fall between the minimum and maximum of the stratum-specific risk ratios RR<sub>ED|c</sub>, thus constraining confounding of the risk ratio by C. The utility of this result is, however, limited by the following observations: Like Figure 1, Figure 2 implies that C is associated with E given D, which (when combined with the association of C with D given E) implies that the E–D odds ratio will be noncollapsible (OR<sub>ED</sub> ≠ OR<sub>ED|c</sub>) apart from unstable exceptions.<sup>7</sup> Thus, under Figure 2, we would ordinarily expect odds-ratio noncollapsibility over C, as well as confounding by C.

As an example of an exception, Table 2 is compatible with Figure 2, in which C and E are connected through two paths C → E and C → D ← E. Nonetheless, Table 2 is constructed so that conditional on D, the positive association between C and E through C → E and the negative association between C and E through C → D ← E cancel each other, making C and E independent given D (180/60 = 60/20, and 20/40 = 90/180) and thus unfaithful to Figure 2. While the odds ratio is simply collapsible in Table 2, confounding by C is evident from the marginal noncollapsibility of all the measures.

A practical question is whether we should be concerned that such an example may arise in practice. We would argue no: Simple collapsibility of odds ratios over a binary C requires independence of two past variables (C and E) given a future variable (D), and thus appears highly artificial under Figures 2 and 3. To ensure independence of C and E given D in a cohort study, the study designer would have to already

**TABLE 1.** An Example of OR Noncollapsibility Without Confounding

D	C = 1		C = 0		Collapsed	
	E	E	E	E	E	E
	1	0	1	0	1	0
1	180	60	80	10	260	70
0	20	40	120	90	140	130
Total	200	100	200	100	400	200
	RR = 1.5		RR = 4		RR = 1.86 <sup>a</sup>	
	RD = 0.30		RD = 0.30		RD = 0.30 <sup>a</sup>	
	OR = 6		OR = 6		OR = 3.45 <sup>a</sup>	

<sup>a</sup>Equal to the RR, RD, OR standardized to total using C-specific risks with weights 300/600 = 1/2 for each stratum.<sup>28</sup>  
OR indicates odds ratio; RR, risk ratio; RD, risk difference.

**TABLE 2.** An Example of OR Collapsibility with Confounding

D	C = 1		C = 0		Collapsed	
	E	E	E	E	E	E
	1	0	1	0	1	0
1	180	60	60	20	240	80
0	20	40	90	180	110	220
Total	200	100	150	200	350	300
	RR = 1.5		RR = 4		RR = 2.57 <sup>a</sup>	
	RD = 0.30		RD = 0.30		RD = 0.42 <sup>a</sup>	
	OR = 6		OR = 6		OR = 6 <sup>a</sup>	

<sup>a</sup>RR = 1.91, RD = 0.30, OR = 3.46 when standardized to total using weights 300/650 and 350/650 for the stratum-specific risks.<sup>28</sup>

OR indicates odds ratio; RR, risk ratio; RD, risk difference.

know the outcome D as well as E and C, and thus would already (a priori) know or at least be able to estimate consistently the effect under study (of E on D). Thus to engineer simple collapsibility would require knowledge on the part of the designer that would curtail the motivation for the study. It seems even more unlikely that natural disease processes would lead to the conditional independence required for simple odds-ratio collapsibility.

For polytomous C, both simple and marginal odds-ratio collapsibility can occur despite presence of the dependencies necessary for noncollapsibility.<sup>19</sup>(Table 6) The same is true of other effect measures, such as risk ratios and mean differences; that is, the dependencies together are not sufficient for noncollapsibility, and thus collapsibility can occur despite faithfulness. Nonetheless, under distributions compatible with and faithful to Figure 2, such collapsibility requires contrived cancelations of changes in the measure as levels of C are combined (collapsed), which makes the collapsibility unstable despite the underlying faithfulness of the distribution.

In contrast, suppose Figure 3 holds (in which there is no effect of E on D but C affects both E and D); under common conditions collapsibility implies unfaithfulness. Because E has no effect,  $M_{ED|C}$  is null (1 for ratios, 0 for differences) at all levels of C; hence, collapsibility requires the unadjusted measure  $M_{ED}$  to be null as well, despite the fact that E and D are connected through C. Table 3 provides an example that is compatible with Figure 3 but unfaithful (and unstable) because E and D are unconditionally independent despite this connection. Thus, if  $M_{ED}$  being null implies independence of E and D (as when E and D are binary, or under models in which the E effect on D is captured entirely by a single coefficient), collapsibility will require unfaithfulness, in addition to being unstable.

Although technical in form, the preceding results provide a rationale for common intuitions about confounding and its relation to noncollapsibility: Given the instability of collapsibility when C affects both E and D (Figures 2 and 3), we should usually expect marginal and simple noncollapsibility

**TABLE 3.** An Example of OR, RR, and RD Collapsibility When C and D Are Associated Given E, and C and E Are Associated Given D

D	C = 2		C = 1		C = 0		Collapsed	
	E	E	E	E	E	E	E	
	1	0	1	0	1	0	1	0
1	60	40	30	20	20	60	110	120
0	30	20	60	40	20	60	110	120
Total	90	60	90	60	40	120	220	240
	RR = 1		RR = 1		RR = 1		RR = 1	
	RD = 0		RD = 0		RD = 0		RD = 0	
	OR = 1		OR = 1		OR = 1		OR = 1	

OR indicates odds ratio; RR, risk ratio; RD, risk difference.

of all effect measures over confounders. For the odds ratio, we would also expect simple noncollapsibility whenever both C and E affect D (Figures 1 and 2). Thus, simple odds-ratio collapsibility is not something to expect or rely on in most settings involving covariate adjustment (in which the covariate affects the outcome but the effect of exposure on the outcome is unknown).

Nonetheless, the growing popularity of methods based on exposure modeling (such as propensity scoring) has led to increased risk of harmful adjustment involving covariates whose only effect on the outcome is through exposure, as in Figures 4 and 5. Adjustment for such covariates is inadvisable because it can inflate the variance<sup>35</sup> and amplify bias<sup>36</sup> of exposure-effect estimates. Under Figure 4, C will be independent of D given E, making all effect measures collapsible over C, and this collapsibility will be stable. In contrast, under Figure 5 with U uncontrolled, C is an instrumental variable, and collapsibility over C will typically be unstable (because C is connected to D conditional on E via the path  $C \rightarrow E \leftarrow U \rightarrow D$  where E is a collider), as under Figure 2; the difference is that under Figure 5, the expected noncollapsibility will represent increased bias in the C-adjusted estimate instead of confounding of the unadjusted estimate by C.<sup>36</sup>

Strong collapsibility has been defined as no change in the measure under any degree of collapsing (coarsening or combining) of levels of C.<sup>18,27</sup> Unlike marginal and simple collapsibility, strong collapsibility requires either independence of C and D given E, or independence of C and E (unconditionally for risk ratios,<sup>18</sup> conditional on D for odds ratios<sup>27</sup>). Consequently, strong collapsibility cannot occur under distributions compatible with and faithful to Figures 2 and 3, or 5; it also cannot occur for odds ratios under distributions compatible with and faithful to Figure 1. In this sense, strong collapsibility of effect measures is more intuitive in behavior than simple collapsibility. It is however so strong that it is equivalent to assuming C is independent of E or D or both.<sup>18,27</sup>

## EXTENSIONS TO INTERMEDIATES AND COLLIDERS

Collapsibility conditions are purely associational, and thus our results for Figures 1–5 hold after reversing arrows directions if those reversals leave the associational structure unchanged (even if they change the temporal or causal structure). The associational structure will remain unchanged if the reversals neither create nor destroy a  $v$ -structure (i.e., two converging arrows whose tails are not connected by an arrow).<sup>8(p. 19)</sup> For example, the results from Figures 2–4 apply with the C–E arrow reversed (i.e.,  $E \rightarrow C$ ) making C an intermediate between E and D in Figures 2 and 3 and a proxy for E in Figure 4. This is not the case for Figure 5, however, where the C–E arrow reversal destroys the  $v$ -structure  $C \rightarrow E \leftarrow U$ .

Although reversing both arrows directions in Figure 3 introduces the  $v$ -structure  $E \rightarrow C \leftarrow D$  (Figure 6), our result still apply: For a binary exposure with no effect, collapsibility over a variable affected by both exposure and outcome requires the variable to be polytomous and implies unfaithfulness. Because E has no effect,  $M_{ED}$  is null; hence, collapsibility requires the adjusted measure  $M_{ED|C}$  to be null as well, despite the fact that E and D are connected conditional on C. Table 3 provides an example compatible with Figure 6 but unfaithful (and unstable) because E and D are independent given C.

## DISCUSSION

Simple collapsibility of odds ratios and logistic coefficients over risk factors is a particularly artificial condition when the exposure affects disease, thus reinforcing advice to avoid evaluation of confounding using odds ratios or logistic coefficients when the outcome is common. Furthermore, this simple collapsibility requires satisfaction of a set of constraints that involve the covariate and outcome<sup>19,37</sup> and hence is not induced neither by any ordinary design nor by a natural causal process of which we are aware. Thus, we would never expect simple odds-ratio collapsibility over a risk factor, regardless of whether it is a confounder. Parallel comments apply to rate ratios and to proportional hazards and Poisson-regression coefficients, although the discrepancy between confounding and noncollapsibility is smaller for these measures.<sup>5</sup> Although one could engineer selection probabilities dependent on both exposure and outcome so that the data exhibited simple odds-ratio collapsibility, estimation would require use of the selection probabilities to remove the selection bias and would produce noncollapsible odds ratios.

Beyond the practical objections to inducing or assuming collapsibility over a risk factor C, we could ask: why would we want simple odds-ratio collapsibility over C if we already had no confounding by C or had adjusted for confounding by C? We know of no practical reason. This is especially so if we abandon odds ratios and rate ratios as our target causal parameter, and regard them instead as only approximations to a risk ratio by virtue of outcome rarity, as in classical case-control studies. For rare outcomes, any noncollapsibility apart from

confounding will be too small to be of concern, given that the risk ratio being approximated is itself collapsible.

For common outcomes, there is usually nothing to recommend a case-control design, and we may instead focus on cohort data and direct estimation of risks or survival times and their contrasts. Although these estimates may be produced by standardizing (averaging or marginalizing) risks from models with noncollapsible coefficients (such as logistic or Cox regression),<sup>38</sup> the resulting risk or time ratio estimates will be collapsible over C when C is independent of E (as in Table 1 and Figure 1).

## REFERENCES

1. Samuels ML. Matching and design efficiency in epidemiological studies. *Biometrika*. 1981;68:577–588.
2. Miettinen OS, Cook EF. Confounding: essence and detection. *Am J Epidemiol*. 1981;114:593–603.
3. Greenland S. Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. *Epidemiology*. 1996;7:498–501.
4. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci*. 1999;14:29–46.
5. Greenland S, Rothman KJ, Lash TL. Measures of effect and measures of association. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:51–70.
6. Pang M, Kaufman JS, Platt RW. Studying noncollapsibility of the odds ratio with marginal structural and logistic regression models. *Stat Methods Med Res*. 2013 Oct 9. [Epub ahead of print].
7. Greenland S, Pearl J. Adjustments and their consequences - collapsibility analysis using graphical models. *Int Stat Rev*. 2011;79:401–26.
8. Pearl J. *Causality*. 2nd ed. New York, NY: Cambridge University Press; 2009.
9. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37–48.
10. Geng Z, Li G. Conditions for non-confounding and collapsibility without knowledge of completely constructed causal diagrams. *Scand J Stat*. 2002;29:169–181.
11. Glymour MM, Greenland S. Causal diagrams. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:183–209.
12. Greenland S, Pearl J. Causal diagrams. In: Lovric M, ed. *International Encyclopedia of Statistical Science*. Vol 3. New York, NY: Springer; 2011:208–216.
13. 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 Feb 17. [Epub ahead of print].
14. Spirtes P, Glymour C, Scheines R. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press; 2001.
15. Weinberg CR. On pooling across strata when frequency matching has been followed in a cohort study. *Biometrics*. 1985;41:103–116.
16. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol*. 2013;42:860–869.
17. Pearl J. *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufmann; 1988.
18. Geng Z. Collapsibility of relative risk in contingency tables with a response variable. *J Roy Stat Soc Ser B (Methodological)*. 1992;54:585–593.
19. Whittemore AS. Collapsibility of multidimensional contingency tables. *J Roy Stat Soc Ser B (Methodological)*. 1978;40:328–340.
20. Greenland S, Maldonado G. The interpretation of multiplicative-model parameters as standardized parameters. *Stat Med*. 1994;13:989–999.
21. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.
22. Weinberg CR, Umbach DM, Greenland S. When will nondifferential misclassification preserve the direction of a trend? *Am J Epidemiol*. 1994;140:565–571.
23. Ogburn EL, VanderWeele TJ. On the nondifferential misclassification of a binary confounder. *Epidemiology*. 2012;23:433–439.

24. Robins JM, Scheines R, Spirtes P, Wasserman L. Uniform consistency in causal inference. *Biometrika*. 2013;90:491–515.
25. Agresti A. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: Wiley; 2013.
26. Wermuth N. Parametric collapsibility and the lack of moderating effects in contingency tables with a dichotomous response variable. *J R Statist Soc B*. 1987;49:353–364.
27. Ducharme GR, Lepage Y. Testing collapsibility in contingency tables. *J Roy Stat Soc Ser B (Methodological)*. 1986;48:197–205.
28. Rubin DB. Practical implications of modes of statistical inference for causal effects and the critical role of the assignment mechanism. *Biometrics*. 1991;47:1213–1234.
29. Gail MH. Adjusting for covariates that have the same distribution in exposed and unexposed cohorts. In: Moolgavkar SH, Prentice RL, eds. *Modern Statistical Methods in Chronic Disease Epidemiology*. New York, NY: John Wiley & Sons; 1986:3–18.
30. Gail MH, Wieand S, Piantadosi S. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika*. 1984;71:431–444.
31. Greenland S, Rothman KJ. Introduction to Stratified Analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:258–282.
32. Neuhaus JM, Kalbfleisch JD, Hauck WW. A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *Int Stat Rev*. 1991;59:25–35.
33. Austin PC, Grootendorst P, Lise-Normand ST, Anderson, GM. Conditioning on the propensity score can result in biased estimation of common measures of treatment effect. *Stat Med*. 2007;26:754–768.
34. Martens EP, Pestman WR, Klungel OH. Re: Conditioning on the propensity score can result in biased estimation of common measures of treatment effect: a Monte Carlo study. *Stat Med*. 2007;26:3208–3210.
35. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163:1149–1156.
36. Pearl J. On a class of bias-amplifying variables that endanger effect estimates. In: Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence. Corvallis, WA: AUAI Press; 2010:417–427.
37. Guo J, Geng Z. Collapsibility of logistic regression coefficients. *J Roy Stat Soc Ser B (Methodological)*. 1995;57:263–267.
38. Greenland S. Introduction to regression modeling. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:418–455.