

# A philosophical analysis of the Hill criteria

Lau Caspar Thygesen, Gregers Stig Andersen, Hanne Andersen

*J Epidemiol Community Health* 2005;59:512–516. doi: 10.1136/jech.2004.027524

The epidemiological literature contains an ongoing and diversified discussion of the Hill criteria. This article offers a philosophical analysis of the criteria, showing that the criteria are related to two different views of causality. The authors argue that the criteria of strength, specificity, consistency, experiment, and biological gradient are related to a probabilistic regularity view of causality, whereas the criteria of coherence, plausibility, and analogy are related to a generative view of causality. The criterion of temporality is not related to either view, but may in contrast be central in inferring direction from cause to effect. The authors illuminate the aim and limitations of the various criteria that need to be included when discussing them.

that different Hill criteria are not related to the same view of causality, which may be a reason for the prolonged discussion of the Hill criteria.

## THE REGULARITY VIEW OF CAUSALITY

On the regularity view of causality formulated by Hume, "a cause [is defined] to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second."<sup>26</sup> In the medical sciences, this view of causality is obviously too simple and Mackie's<sup>23</sup> more subtle version of the regularity view seems more suitable.<sup>27–29</sup>

On Mackie's account, a causal complex may be seen as a conjunction of factors that only jointly are sufficient for bringing about the effect. There may be several different causal complexes that are all sufficient for bringing about the effect; hence, none of them are necessary. A factor in such a causal complex is an insufficient but necessary condition in an unnecessary but sufficient causal complex, also called an INUS condition. This model has independently been applied to epidemiology as the component-cause model,<sup>28 29</sup> and in accordance with standard epidemiological terminology we shall refer to INUS conditions as causal factors.

The full cause is the conjunction of all possible causal complexes for the effect in question. Furthermore, causal statements are made against a background of assumptions, which Mackie terms the causal field. The complex picture of both the full cause seen as a disjunction of several, individually sufficient causal complexes and the causal field as a necessary background for any analysis of causality can be described by the condition that F and (A and X or Y) is a necessary and sufficient condition for bringing about the effect, where F is the conjunction of assumptions included in the causal field, A is a known individual factor, and X and Y are conjunctions of factors that may be partially or fully unknown.

Mackie's analysis still shows a deterministic view of causality based on sufficiency: whenever all causal factors in a given causal complex are present, the effect will come about. To separate the idea of causality from determinism, philosophers such as Suppes<sup>24</sup> have suggested defining a cause as a condition that merely raises the probability of the effect. Thus, a factor A is a (partial) cause of the effect E if  $P(E|A) > P(E|\text{absence of } A)$ .

This view can be used to elaborate Mackie's regularity account of causality to include cases in which only parts of the full cause are known. If, for the sake of simplicity, we consider a full cause that can be expressed as A and X where only A but not X is known, the effect E will only follow in a percentage of the cases in which the

In the epidemiological literature, the Hill criteria<sup>1</sup> have been widely accepted as guidelines for drawing causal inferences,<sup>2–11</sup> but at the same time they have been quite controversial.<sup>12–15</sup> Hill himself explicitly denied that the guidelines should be termed criteria, but rather viewpoints.<sup>1</sup> Likewise advocates of the Hill criteria state that although they are not rigid criteria that must all be fulfilled, they still give positive support to inferences about causality.<sup>6–9 16</sup> We follow this convention by terming the nine guidelines as criteria in this paper. Critics argue that the criteria are obscurely described and partly overlapping<sup>10 13 14</sup> or that none of the criteria (except temporality) apply in all circumstances.<sup>12</sup> By the same token, it has been argued that they should not be seen as rules, but rather as values that can be interpreted and weighed differently by different scientists.<sup>17 18</sup> The arguments advanced by both critics and proponents of the criteria are primarily based on examples and counter-examples. An important exception<sup>19</sup> is the discussion in the 1970s–1980s of the merits of criteria in general evaluating causality between epidemiologists supporting Popper's falsificationist model of science who reject criteria as purely rules of induction on the one side<sup>20</sup> against the more pragmatic inductive approach suggesting criteria to evaluate causes in epidemiology.<sup>2</sup>

In contrast with the example driven arguments and the philosophical discussions of the merit of criteria in general, the aim of this article is to carry out a theoretical philosophical analysis of the Hill criteria based on two philosophical accounts of causation, a probabilistic elaboration of the regularity view of causation<sup>21–24</sup> and the generative view of causation.<sup>25</sup> We shall argue

See end of article for authors' affiliations

Correspondence to:  
Mr L C Thygesen, Centre  
for Alcohol Research,  
National Institute of Public  
Health, Øster  
Farimagsgade 5, DK-1399  
Copenhagen K, Denmark;  
lct@niph.dk

Accepted for publication  
5 January 2005

factor A occurs, namely the percentage corresponding to the prevalence of X:  $P(E|A) = P(X)$ . As E never occurs without the prior occurrence of A,  $P(E|\text{absence of } A) = 0$ , and the condition  $P(E|A) > P(E|\text{absence of } A)$  is therefore fulfilled.

Likewise, considering the full cause A and X or Y, and assuming X and Y to be independent,

$$P(E|A) = P(X \text{ OR } Y) = P(X) + P(Y) - P(X)P(Y) \tag{1}$$

while

$$P(E|\text{absence of } A) = P(Y) \tag{2}$$

The probabilistic account can thus describe all aspects of the regularity account.

### THE GENERATIVE VIEW OF CAUSALITY

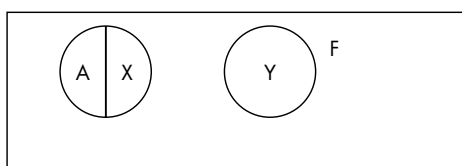
On the above account, causality is understood in terms of regularities. An alternative view is that causality can be understood as a generative relation of production.<sup>25-27</sup> On this view, the statement that “A causes E” implies that A generates or produces E by some kind of mechanism.<sup>19</sup> The generative mechanism is seen as an objective process underlying the relation between cause and effect, and evolving according to general, biological laws. Thus, as pointed out by Renton, on this view the association between A and E is a particular instance deducible from general, biological laws rather than a relation induced from repeated observations.<sup>19</sup>

According to Parascandola,<sup>30</sup> it has been popular throughout the history of epidemiology to describe epidemiological research in terms of inductive inferences from observed association and laboratory research in terms of deductive inferences from theoretical principles; a contrast that often has been seen as if laboratory research elaborating on our fundamental medical knowledge of a phenomenon can establish causal relations with certainty, while epidemiological research provides only circumstantial evidence.

At first sight, such a contrast would seem to fit the contrast between the regularity view and the generative view of causality. However, one may argue that many biological laws obtained through laboratory research are also based on observations of regularities, usually at a lower level of organisation. Thus, the underlying specification of a biological mechanism responsible for the association will often consist in the specification of a sequence of causes on a micro-level that intervene between the original macro-level cause and effect.

As noted by Parascandola and Weed<sup>27</sup> there is a tendency to prioritise knowledge at the molecular level. However, as far as these micro-level causes are also based on observations of regularities, research on biological mechanisms may not be different in kind from research on the associations between given causal factors and a given disease.

Parascandola and Weed argue that one probable reason why the molecular level has nevertheless been prioritised by many researchers is the underlying belief that associations at the macro-level can be fully reduced to causes at the molecular level that behave deterministically. However, as they



**Figure 1** Mackie’s analysis of the full cause A and X or Y in the causal field F.

also note, this kind of simple reductionism has also been widely criticised.<sup>27</sup>

Another reason why a generative view of causality may seem attractive to many researchers in the health sciences may be that the exploration of intervening causes in the causal net leads to a more and more extensive net of interrelated theories that will have a larger and larger explanatory power. On that view, the interesting difference between the regularity view and the generative view of causality is whether focus is on individual associations or on the explanatory power of a network of mutually related associations.

### THE HILL CRITERIA

Based on these analyses of causality we shall argue that the Hill criteria of strength, specificity, consistency, experiment, and biological gradient all concern observed associations and are therefore all primarily related to the probabilistic regularity view of causality, while the criteria coherence, plausibility, and analogy are in contrast all primarily concerned with the explanatory power of a network of mutually related associations and therefore are related to the generative view of causality. Epidemiologists agree to find the criterion of temporality useful,<sup>2-14 16 31-33</sup> we will briefly show that the criterion is not related to either views of causality, but by contrast may be essential to infer direction from cause to effect.

#### The probabilistic regularity view of causality Strength

This criterion states that if an observed effect of A on E is strong it could implicate causality,<sup>1</sup> thereby claiming that the variation of A influence the probability of E.

A measure of the strength of the association between a causal factor A and the event E can be expressed as the ratio between the probability of E conditioned on A divided by the probability of E conditioned on the absence of A.

$$\Psi = \frac{P(E|A)}{P(E|\text{absence } A)} \tag{3}$$

Assuming that X and Y are independent  $\Psi$  can be derived from (1) and (2):

$$\Psi = 1 + \frac{(1 - P(Y)) \cdot P(X)}{P(Y)} \tag{4}$$

Table 1 shows how the strength of the association between A and E (measured by  $\Psi$ ) depends on the occurrence of X and Y.

The strength of the causal factor A is completely dependent on the prevalence of X and Y.

Basically, this conclusion is in agreement with an epidemiological understanding of causality where the effect of one factor depends on the prevalence of complementary factors.<sup>34</sup> But including competing causal complexes (Y) in the full cause, our analysis shows that even a high prevalence of the complementary causal factors (X) may be dominated by highly prevalent competing causal complexes (Y) and that A will therefore seem to be weak.

This has considerable consequences for the epidemiological understanding of causality, where the criterion of strength is

**Table 1** Strength of the association ( $\Psi$ ) between A and E depending on the occurrence of X and Y

	$P(Y) \rightarrow 0$	$P(Y) \rightarrow 1$
$P(X) \rightarrow 0$	$\Psi \rightarrow 1 + P(X)/P(Y)$	$\Psi \rightarrow 1$
$P(X) \rightarrow 1$	$\Psi \rightarrow \infty$	$\Psi \rightarrow 1$

often emphasised as especially important<sup>16 32</sup> and a high strength is argued to minimise the risk of confounding.<sup>5 6 8 9 11</sup> This analysis shows that the strength of an association is dependent on the occurrence of both complementary causal factors and competing causal complexes.

### Consistency

This criterion states that if an observed effect of A on E is shown in different situations it could indicate causality.<sup>1</sup> This criterion is by epidemiologists emphasised as being especially important, because *“it is rare that any single study is considered definitive. The power of the epidemiologic approach lies in observing similar findings consistently from a large set of diverse studies that address specific relations.”*<sup>35</sup> The underlying principle is that the potential cause A of the effect E should be constant while other factors may vary. But repeated observations of an association in different populations under different circumstances can be achieved in different ways. Firstly, the factors varying between the different studies may not be part of the full cause or the causal field. Secondly, these factors may form part of the full cause as part of either X or Y. As the strength of the effect of A on E is dependent on the prevalence of X, to achieve repeated observation of the same observation it is required that the prevalence of X remains constant in all populations. To fulfil the consistency criterion, the variation of the factors included in X should therefore be balanced in such a way that the prevalence of their conjunction remains constant. By the same token, a variation of the prevalence of the competing causal complex Y may cause variation of the effect, even if the prevalence of the examined causal factor A and its causal complement X remains constant.

Our analysis may have considerable consequences for the value of meta-analysis, which rests on the principle of homogeneity of results in different populations. In fact, consistency is only expected if factors in both X and Y are equally distributed in all populations included in the meta-analysis. If this requirement is not fulfilled, homogeneity between studies might be a consequence of other factors than A, and in contrast heterogeneity might stem from differences in prevalence of the other factors X and Y in the full cause.

Therefore, one may conclude that a lack of consistency may be the result of a causal complex A and X or Y in which X or Y contain factors that occur only under very specific circumstances. That is congruent with others, who add that consistency only can be used to discard hypotheses in which the association could be ascribed to one of the factors varying between studies.<sup>5-9 12</sup>

### Specificity

This controversial criterion, which Hill adopted from Yerushalmy and Palmer,<sup>36 37</sup> says that it is preferable if the causal association is between one specific cause A and one specific effect E.<sup>1</sup> The criterion covers two different claims.<sup>38</sup> The first claim is that there should be one specific cause only. Therefore, it is necessary that the full cause only consists of A (no inclusion of X and Y in the full cause). If we consider X as an always present condition in a given causal field where Y is absent, the cause A could be considered specific. But in this case X may as well be seen as part of the background on which the causing goes on—that is, as part of the causal field. This claim is therefore connected to the definition of the causal field.

The second claim is that a given causal factor A can only be a causal factor for one effect E. This part of the criterion is closely related to the definition of the effect E. If the effect is defined broadly, for example, as the development of disease in general, it could be claimed that the full cause A only enters in the full cause of one effect. However, such a broad definition will rarely be of much use. Similarly, the effect

may often be defined very narrowly, including the causal factor in the definition of the effect (for example, including a specific agent in the definition of an infectious disease), but in this case the causal argument becomes circular as the presence of the causal factor is already included in the definition of the effect.

Several authors argue that causes cannot be expected to have only one effect on any logical grounds and sees the criterion as misleading and useless.<sup>5-7 9 10 12</sup> It seems reasonable to conclude that monocausality seldom arises as described above. Hill was aware of this, but wondered if it was attributable to the lack of knowledge of aetiology; if we knew all aspects of a certain disease it would be possible to point out one single cause.<sup>1</sup> Instead, we shall emphasise that the definitions of both cause and effect are central for the application of the criterion.

### Biological gradient

The preference for associations showing a dose-response relation entails different interpretations: a higher exposure to factor A may result in a higher incidence of the disease E, or a higher exposure to factor A may result in a higher biological effect. Additionally, a higher exposure of A could both mean a higher intensity of A, or a longer duration of the exposure of A.

If higher exposure of A is interpreted as a longer duration, we would usually accept that the longer the exposure lasts, the larger is the probability of obtaining a causal complex for the effect in question. This argument rests on the assumption that X and Y are time varying.

If higher exposure of A is interpreted as a higher intensity, so far as the factor A can be understood as additive entities or processes (such as, for example, molecules, radiation, etc) intensity can be understood as the simultaneous presence of more (or fewer) As. Again, the more As present, the larger is the probability of obtaining a causal complex for the effect in question. Similar to the case of duration, this argument rests on the assumption that X and Y are not universally present.

In both cases, increasing the probability of obtaining a causal complex points to a higher incidence of the disease E. However, an increased E may instead be interpreted as a higher biological effect, similar to the interpretation of A in terms of intensity. So far as the effect E can be understood as additive processes, if the effect is interpreted as a higher biological effect, this can be seen as the presence of several causal complexes, AX, thereby initiating several Es.

In all cases, it must be noted that the prevalence of the other factors (X and Y) in the full cause has pronounced influence on the effect of A on E. Thus, if the prevalence of Y is high for high levels of A and the prevalence of X is low, it is possible that the dose response follows the association between Y and E and not A and E.

Several authors state that repeated observations of dose response is a strong argument for a causal inference,<sup>5-11 16 31 32</sup> while other argue that there are examples of causal relations in which the prevalence of E does not follow the prevalence of A in a simple manner.<sup>12</sup> This argument is congruent with our analysis where E is dependent not only on A but also on X and Y.

### Experiment

This criterion, which Hill calls the strongest support for a causation hypothesis, is closely related to Hume's second definition of causation: *“where, if the first object had not been, the second never had existed.”*<sup>26</sup> The criterion is occasionally interpreted as concerning laboratory experiments, but in his original article Hill is concerned with population based preventive interventions.<sup>1</sup> In studies of preventive interventions, the criterion examines whether the effect E occurs when the factor A is present and absent, respectively, while



### What this paper adds

- This article has two main findings. Firstly it is shown that the Hill criteria are related to two different views of causality. The first is the probabilistic regularity view of causality based on the observed association between the effect and a factor A. It is shown that the effect of A is influenced by the prevalence of other factors in the same causal complex, X, and other causal complexes, Y. We have argued that the criteria of strength, specificity, consistency, experiment, and biological gradient are related to a probabilistic regularity view of causality. The second view of causality is the generative view, which focuses on the explanatory power of networks of mutually related associations. We have argued that the criteria of coherence, plausibility, and analogy are related to this view of causality. The criterion of temporality is not related to either view, but may in contrast be central in inferring direction from cause to effect.
- Secondly, this study shows that the criteria related to the probabilistic regularity view of causality hold under certain specific conditions, included in our analysis as the causal factors X and Y, which are partly or fully unknown. The criteria related to the generative view of causality may be attractive because of their focus on the explanatory power of a net of mutually related associations. We have thereby illuminated the differences in focus as well as the limitations of the Hill criteria that need to be included when discussing them.

no other factors vary. However, it must be emphasised even if only the factor A does vary, the other factors X and Y in the full cause are important for the occurrence of E. In the case Y is present or X is absent the prevalence of E is independent of A. The experiment will only support the causal claim of A on E where X is present and Y is absent.

Also, it could be argued that support from experimental studies of animals is necessary in making inferences about causality from an observational association. However, contradictory experimental evidence does not immediately refute an observational study because of obvious differences between humans and animals. More important, animal experiments will often be used to examine an alleged biological mechanism, which we shall discuss below.

### The generative view of causality Plausibility and coherence

According to the criterion of “plausibility” the association ought to be biologically plausible. According to the criterion “coherence”, the association should not conflict with the generally known facts of the biology of the disease.

Epidemiological authors have questioned the difference between the two criteria, as both imply that an association should be in congruence with the biological knowledge of the day.<sup>5 10 12 33</sup> Both criteria rely on the existence of a biological pathway from cause to effect—that is, there is some effect from the causal factor A, which could be expressed as a kind of mechanism or generative relation, that leads to the development of the disease E. With this background, the difference between the two criteria can be seen as the difference between admitting that inferences from current biological theory vindicate the observed association (plausibility), and denying that inferences from biological theory contradict the observed association (coherence).

### Policy implications

This study shows that all criteria hold under certain specific conditions, included in our analysis as the causal factors X and Y, which are partly or fully unknown. The Hill criteria may therefore to some seem useful and sound because they in many conditions do hold, while others may find them useless and misleading, because under other conditions they do not hold.

It is obvious that coherence and plausibility support pre-existing theory<sup>33</sup>—a qualification also noted by Hill—and several authors have stated their reservations with regard to this implied conservatism.<sup>7–9</sup> For example, conflicting information may on a first look falsify a causal hypothesis, but this information may later be proved wrong.<sup>12</sup> Therefore, it is necessary to be careful before dismissing an association that is not vindicated by (or even contradicted by) current biological theories.

Other epidemiologists advance that biological evidence is collected from animal models, in vitro cell systems, and human metabolic and clinical studies, where the relevance of each type of evidence is controversial when evaluating causality. The incorporation of genetic and other biological markers as exposures (and sometimes as end points) in epidemiological studies suggests that biological plausibility will become more important to causal inference in the future.<sup>39</sup>

As described above, there has been a tendency to discuss biological theory and laboratory research in terms of deduction and certainty and epidemiology in terms of induction and uncertainty. This may be the reason why these criteria, concerned with inferences from biological theory, have been described as “the most important consideration”<sup>40</sup> or an “ultimate” criterion for causality.<sup>33</sup> However, as likewise described above, causes at the molecular level do not necessarily behave deterministically,<sup>27</sup> neither is research on biological mechanisms necessarily different in kind from research on macro-level associations. However, plausibility and coherence may be attractive because of their focus on increasing the extension and explanatory power of a net of interrelated associations, laws, and theories.

### Analogy

According to this criterion, we should be ready to accept a slighter but similar evidence with a similar risk factor and a similar affect.<sup>1</sup> However, the crucial point is what counts as similar evidence, as any two things will be similar in some respect.<sup>41</sup> Thus, if “similar factors” are to be assumed to have “similar effects”, they should not be similar in any arbitrary sense, but similar in a way that may plausibly be involved in the causal mechanism. The criterion of analogy is therefore closely linked to the criteria of plausibility and coherence.

The requirement that the relation of analogy between conjectured causal factors has to be specified with respect to how the two factors may play similar parts in a causal mechanism qualifies the criticism that analogy is dependent of the imagination of the scientist.<sup>12</sup> On the other hand, if mechanisms are specified, which explain the similar action of different risk factors, this contributes to the extension of a coherent set of interrelated associations, laws, and theories and thereby an increase of its explanatory power.

### Temporality

Hill emphasises the criterion of temporality as necessary, because in order for A to cause E, A must precede E in time. But although epidemiologists agree that the criterion is useful,<sup>2–14 16 31–33</sup> a philosophical analysis shows that it is difficult to argue about why a cause should precede its effect

in time. It is possible to imagine cause and effect operating simultaneously and in physics causation may not even follow temporality.<sup>42</sup>

It is normally assumed that causation is asymmetric, for example, if A causes E, then, typically, E will not also cause A. This may pose a problem for the probabilistic regularity view of causality. By applying INUS conditions it could be shown that if A is a INUS condition of E, then E is also a INUS condition of A.<sup>43</sup> Therefore, by applying Mackie's theory the relation between cause and effect is symmetric. One way of enforcing asymmetry between cause and effect, originally adopted by Hume,<sup>26</sup> is to stipulate that cause precedes its effect in time.

The generative view of causality does not immediately have the problem of symmetry, because the coherence of a network of mutually related associations may have an encapsulated explanation about cause and effect. But, as introduced above, the generative view of causality although introducing a more coherent view of mutually related associations is not different from the regularity view of causality in kind, and the difficulties of inferring asymmetry for Mackie's theory may therefore also apply for the generative view of causality.

Therefore, this analysis shows that the criterion of temporality is not related to either views of causality, but in contrast may attribute the views of causality by inferring direction from cause to effect.

## CONCLUSION

This article has two main findings. Firstly, it is shown that the criteria are related to two different views of causality. The first is the probabilistic regularity view of causality based on the observed association between the effect and a factor A. It is shown that the effect of A is influenced by the prevalence of other factors in the same causal complex, X, and other causal complexes, Y. We have argued that the criteria of strength, specificity, consistency, experiment, and biological gradient are related to a probabilistic regularity view of causality. The second view of causality is the generative view, which focuses on the explanatory power of networks of mutually related associations. We have argued that the criteria of coherence, plausibility, and analogy are related to this view of causality. The criterion of temporality is not related to either view, but may by contrast be central in inferring direction from cause to effect.

Secondly, this study shows that the criteria related to the probabilistic regularity view of causality hold under certain specific conditions, included in our analysis as the causal factors X and Y, which are partly or fully unknown. The criteria related to the generative view of causality may be attractive because of their focus on the explanatory power of a net of mutually related associations. We have thereby illuminated the differences in focus as well as the limitations of the Hill criteria that need to be included when discussing them.

## ACKNOWLEDGEMENTS

We thank Peter Barker and Morten Gronbaek for valuable comments on a draft of this paper as well as Susanne Dahl for linguistic corrections.

## Authors' affiliations

**L C Thygesen, G S Andersen**, Centre for Alcohol Research, National Institute of Public Health, Copenhagen, Denmark  
**H Andersen**, Department of Medical Philosophy and Clinical Theory, University of Copenhagen, Copenhagen, Denmark

Funding: none.

Conflicts of interest: none declared.

## REFERENCES

- Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;**58**:295–300.
- Susser M. What is a cause and how do we know one? A grammar for pragmatic epidemiology. *Am J Epidemiol* 1991;**133**:635–48.
- Gregg MB. *Field epidemiology*. Oxford: Oxford University Press, 2002.
- Woodward M. *Epidemiology: study design and data analysis*. Boca Raton: Chapman and Hall/CRC, 1999:18.
- Bakketeig LS, Magnus P. Hva er en årsaksfaktor? In: Bakketeig LS, Magnus P, eds. *Epidemiologi*. (In Norwegian). Oslo: Gyldendal Akademisk, 2003:18–21.
- Gerstman BB. From association to causation. In: Gerstman BB, ed. *Epidemiology kept simple—an introduction to classic and modern epidemiology*. New York: Wiley, 1998:201–213.
- Daly J, Kellehear A, Gliksmann M. Epidemiological surveys. In: Daly J, Kellehear A, Gliksmann M, eds. *The public health researcher: a methodological guide*. Melbourne: Oxford University Press, 1997:33–48.
- Oleckno WA. Association and Causation. In: *Essential epidemiology: principles and applications*. Prospect Heights, IL: Waveland Press, 2002:123–35.
- Olsen J, Overvad K, Juul S. *Analytisk epidemiologi—en introduktion*. (In Danish). Copenhagen: Munksgaard, 1998:58–60.
- Young TK. Assessing health risk in populations. In: *Population health: concepts and methods*. New York: Oxford University Press, 1998:143–49.
- Breslow NE, Day NE. *Statistical methods in cancer research*. Vol 1. *The analysis of case-control studies*. Lyon: International Agency for Research on Cancer, 1980:84–90.
- Rothman KJ, Greenland S. *Modern epidemiology*. 2nd ed. Philadelphia: Lippincott-Raven, 1998:7–28.
- MacMahon B, Trichopoulos D. *Epidemiology: principles and methods*. 2nd ed. Boston: Little, Brown, 1996:19–29.
- Bhopal RS. Cause and effect: the epidemiological approach. In: *Concepts of epidemiology—an integrated introduction to the ideas, theories, principles and methods of epidemiology*. Oxford: Oxford University Press, 2002:98–132.
- Weed DL. Methods in epidemiology and public health: does practice match theory? *J Epidemiol Community Health* 2001;**55**:104–10.
- Doll R. Proof of causality: deduction from epidemiological observation. *Perspect Biol Med* 2002;**45**:499–515.
- Poole C. Causal values. *Epidemiology* 2001;**12**:139–41.
- Kaufman JS, Poole C. Looking back on "causal thinking in the health sciences". *Annu Rev Public Health* 2000;**21**:101–19.
- Renton A. Epidemiology and causation: a realist view. *J Epidemiol Community Health* 1994;**48**:79–85.
- Weed DL. On the logic of causal inference. *Am J Epidemiol* 1986;**123**:965–79.
- Mackie JL. Causes and conditions. *American Philosophical Quarterly* 1965;**2**:245–55.
- Mackie JL. *The cement of the universe*. Oxford: Clarendon Press, 1974.
- Mackie JL. *The cement of the universe*. New York: Oxford University Press, 1980:59–87.
- Suppes P. *A probabilistic theory of causation*. Amsterdam: North-Holland, 1970.
- Wulff H, Pedersen SA, Rosenberg R. *Philosophy of medicine*. Oxford: Blackwell, 1986.
- Hume D. *An Enquiry concerning human understanding*. Oxford: Oxford University Press, 1999: section VII.
- Parascandola M, Weed DL. Causation in epidemiology. *J Epidemiol Community Health* 2001;**55**:905–12.
- Olsen J. What characterises a useful concept of causation in epidemiology? *J Epidemiol Community Health* 2003;**57**:86–8.
- Tam CC, Lopman BA. Determinism versus stochasticism: in support of long coffee breaks. *J Epidemiol Community Health* 2003;**57**:477–8.
- Parascandola M. Epidemiology: second-rate science? *Public Health Rep* 1998;**113**:312–20.
- Gordis L. From association to causation: deriving inferences from epidemiologic studies. In: *Epidemiology*. Philadelphia: W B Saunders, 2000:184–203.
- Vetter N, Matthews I. Causation. In: *Epidemiology and public health medicine*. Edinburgh: Churchill Livingstone, 1999:23–30.
- Susser M. *Causal thinking in the health sciences. Concepts and strategies of epidemiology*. New York: Oxford University Press, 1973.
- Rothman KJ. Causes. *Am J Epidemiol* 1976;**104**:587–92.
- Byers T, Lyle B. The role of epidemiology in determining when evidence is sufficient to support nutrition recommendations. Summary statement. *Am J Clin Nutr* 1999;**69**:1365–1367S.
- Yerushalmy J, Palmer CE. On the methodology of investigations of etiologic factors in chronic diseases. *J Chronic Dis* 1959;**10**:27–40.
- Parascandola M. Two approaches to etiology: the debate over smoking and lung cancer in the 1950s. *Endeavour* 2004;**28**:81–6.
- Weiss NS. Can the "specificity" of an association be rehabilitated as a basis for supporting a causal hypothesis? *Epidemiology* 2002;**13**:6–8.
- Weed DL, Hursting SD. Biologic plausibility in causal inference: current method and practice. *Am J Epidemiol* 1998;**147**:415–25.
- Lilienfeld AM. Epidemiological methods and inferences in studies of non-infectious diseases. *Public Health Rep* 1957;**72**:51–60.
- Goodman N. Seven strictures on similarity. In: *Problems and projects*. Indianapolis: Bobbs-Merrill, 1972:437–46.
- Weed DL. Causal criteria and Popperian refutation. In: Rothman KJ, ed. *Causal inference*. Chestnut Hill, MA: ERI, 1988:17.
- Mackie JL. *The cement of the universe*. New York: Oxford University Press, 1980:160–1.