Specificity of Association in Epidemiology

Thomas Blanchard, University of Cologne

Abstract

The epidemiologist Bradford Hill famously argued that in epidemiology, specificity of association (roughly, the fact that an environmental or behavioral risk factor is associated with just one or at most a few medical outcomes) is strong evidence of causation. Prominent epidemiologists have dismissed Hill's claim on the ground that it relies on a dubious `one-cause one effect' model of disease causation. The paper examines this methodological controversy, and argues that specificity considerations do have a useful role to play in causal inference in epidemiology. More precisely, I argue that specificity considerations help solve a pervasive inferential problem in contemporary epidemiology: the problem of determining whether an exposure-outcome correlation might be due to confounding by a social factor. This examination of specificity has interesting consequences for our understanding of the methodology of epidemiology. It highlights how the methodology of epidemiology relies on local tools designed to address specific inference problems peculiar to the discipline, and shows that observational causal inference in epidemiology can proceed with little prior knowledge of the causal structure of the phenomenon investigated. I also argue that specificity of association cannot (despite claims to the contrary) be entirely explained in terms of Woodward's well-known concept of "one-to-one" causal specificity. This is because specificity as understood by epidemiologists depends on whether and exposure (or outcome) is associated with a `heterogeneous' set of variables. This dimension of heterogeneity is not captured in Woodward's notion, but is crucial for understanding the evidential import of specificity of association.

In his well-known 1965 address to the Royal Society of Medicine, Bradford Hill proposed nine considerations to take into account when drawing causal conclusions from observational studies in epidemiology (Hill, 1965). One of them is *specificity of association*, which concerns the extent to which a single risk factor (or `exposure') is correlated with a single medical outcome. Hill claimed that more specific associations are more likely to be causal. However, prominent figures in epidemiology have dismissed that claim on the ground that it relies on an outdated `one cause – one effect' model of disease etiology. The goal of this paper is to examine this methodological issue, which has attracted little attention from philosophers of science so far. I will offer a defense of Hill's view, and argue that considerations of specificity of association do have a valuable role when drawing conclusions concerning the health effects of environmental and behavioral risk factors (the type of risk factors with which Hill himself was concerned).

As I will explain, my account has a number of interesting implications. It highlights the fact that epidemiology relies on what Currie (2015) calls "localized epistemic tools" custom-made to address particular inferential problems that arise within the discipline. It also has important consequences for the debate concerning the relative merits of randomized experiments and observational studies for causal discovery in epidemiology. Finally, my account also contributes to our understanding of the notion of specificity more generally. That notion is ubiquitous in biology and medicine, and has been the object of extensive philosophical discussion.[1] That discussion has been dominated by Woodward's (2010) well-known account of causal specificity. As we will see, Woodward's account captures one dimension of specificity of association, but the notion also includes a further aspect not included in Woodward's analysis, namely the extent to which the putative effects of a variable are `heterogeneous' or `disparate'. Moreover, this aspect of the notion is crucial to understanding why and how considerations of specificity of association are evidentially significant in epidemiology. Thus another important lesson of the paper is that we need a richer notion of specificity than Woodward's to capture certain scientifically interesting forms of specificity.

I. Specificity of association and its critics

The topic of Hill's 1965 address was the following question: when and how do observational studies allow us to conclude that an exposure (i.e., putative risk factor) causes a certain medical outcome? Hill was primarily concerned with the health effects of what epidemiologists now call `environment and lifestyle factors': i.e., the environmental conditions to which people are exposed in the course of their daily life (including work) and the habits and behaviors they adopt as a result of individual choice and/or socioeconomic context. (Work-related exposures were of particular interest to Hill, as he was speaking in his capacity as President of the newly-formed Section of Occupational Medicine of the Royal Society.) Suppose, then, that we observe a correlation between such a risk factor A and a certain medical outcome B. Can we conclude that A causes B? Hill put forward nine aspects to consider when answering that question, which are

---

[1] In particular, philosophers of biology have examined in detail the idea that DNA is a specific cause of protein synthesis and how this idea relates to concepts of information and coding in molecular biology. (See e.g. Weber 2006, Woodward 2010, Griffiths et al. 2015 and Stotz & Griffiths 2017.) For discussions of other forms of specificity that play important roles in biology and medicine, see Woodward (2010) and Lean (2020).

now called "Hill's criteria of causation". (I will follow that standard terminology here, though it is worth noting that Hill himself did not regard his considerations as necessary or sufficient for causation.[2]) The specificity of the association between A and B is the third of Hill's criteria. Following Weiss (2002), we can distinguish two dimensions of specificity of association, `specificity of outcome' and `specificity of exposure'. Roughly, the former concerns the extent to which A is associated with no or few other pathologies besides B, while the latter concerns the extent to which B is correlated with no or few exposures besides A. (Both notions admit of degrees. The smaller the number of other outcomes with which A is associated, the higher the specificity of outcome; likewise, *mutatis mutandis*, for specificity of exposure.) Hill's main focus is on specificity of outcome. He gives the example of the prevalence of lung and nose cancer among nickel refiners[3], and notes that this association is highly specific insofar as exposure to nickel is not associated with other types of cancer or other serious pathologies. This, he says, is strong evidence that nickel exposure is a cause of those two cancers. More generally, Hill's position is that more specific associations are more likely to be causal.

Among contemporary epidemiologists, the evidential value of specificity of association is highly controverted. On the one hand, prominent figures in the field have dismissed Hill's specificity criterion as utterly misguided. Thus, in a classic paper on causal inference in epidemiology, Mervyn Susser writes that

> [a]rguments that demand specificity are fallacious, if not absurd. There can be no logical reason why any identifiable factor (…) should not have multiple effects. (1977, 713)

Similarly, in the second edition of their influential textbook *Modern Epidemiology*, Rothman and Greenland reject the criterion as "useless and misleading" (Greenland & Rothman, 1997), and in a more recent paper on causal inference in epidemiology they write that

> the criterion is invalid as a general rule. Causes of a given effect cannot be expected to lack all other effects. In fact, everyday experience teaches us repeatedly that single events or conditions may have many effects. (Rothman & Greenland, 2005)

Both Susser as well as Rothman and Greenland mention the case of smoking in support of their view. When statistical evidence of harmful effects of smoking emerged in the 1950s, a number

---

[2] He himself called them `viewpoints'.
[3] See Doll et al. (1970).

of epidemiologists rejected that hypothesis because the relevant associations were non-specific. For example, Berkson (1958) argued against the hypothesis of a causal link between smoking and lung cancer on the ground that smoking is also associated with many other cancers, cardiovascular diseases, and various other serious pathologies. This is a striking example where specificity considerations lead to the wrong conclusion.

On the other hand, some epidemiologists have offered qualified defenses of the specificity criterion, and in the last two decades the criterion has regained some popularity, thanks in particular to a paper by Noel Weiss (2002).[4] For example, in recent years, specificity considerations have been used to argue for or against a causal interpretation of, among others, the correlation between vitamin C intake and cardiovascular disease (Lawlor et al., 2004), downsizing and cardiovascular deaths (Vahtera et al., 2004), asbestos and mesothelioma of the lung (Freeman & Kohles, 2012), cannabis use and academic performance (Stiby et al., 2015), and vitamin D deficiency and multiple sclerosis (Simpson & van der Mei, 2019). (These studies rarely if ever contain an explicit rationale for their use of the criterion: often, the only justification provided is a passing reference to Weiss's paper.) But leaving those sporadic uses of the criterion aside, the view that the criterion has little value seems to remain the dominant opinion in the field. (For instance, in an overview of Hill's criteria, the EPA's guidelines on carcinogenic risk assessment state that specificity `is now considered one of the weaker guidelines for causality' (U.S. Environmental Protection Agency, 2005, 2–14).)

To examine the chief complaint lodged by the epidemiologists cited above against the specificity criterion, it is helpful to introduce Woodward's concept of `one-to-one specificity', which he proposes to explain specificity of association and a variety of other specificity concepts that play important roles in the biological and medical sciences (Woodward 2010: 308-313). On his account, the causal relationship between *X* and *Y* is one-to-one specific insofar as *X* is the only variable (within a specified range) that causes *Y*, and *Y* the only variable (within a specified range) that is caused by *X*. In effect, Susser as well as Rothman and Greenland are accusing Hill of presupposing that exposure-outcome relationships must be or at least generally are one-to-one specific. (Thus Susser attacks "arguments that demand specificity", while Rothman and Greenland target specificity conceived as a "general rule".) This view of disease causation as one-to-one specific is associated with late 19th century epidemiology, and in particular with

---

[4] I will return to Weiss's paper in section 2.

Koch's postulates.[5] It is arguably correct for the case of infectious diseases, and as Koch's work illustrates it can greatly facilitate the discovery of the causes of infections.[6] But this model of disease causation is patently inadequate for the chronic pathologies to which epidemiologists devote most of their attention nowadays. It is commonplace in contemporary epidemiology that chronic diseases are caused by a multiplicity of factors.[7] Moreover, those factors are often causally relevant to multiple diseases at once. For instance, cardiovascular disease is caused by (among others) diet, smoking, and physical inactivity, and each of those factors is a cause of many other pathologies as well. (A look at the history of the specificity criterion reinforces the suspicion that it is a relic of 19th century disease etiology, as it appears to have been first proposed by Yerushalmy and Palmer (1959) as part of an attempt to adapt Koch's postulates to the study of chronic diseases.) Note that the epidemiologists cited above are not the only ones to make these criticisms; similar ones can be found in the philosophy of science. Thus, Russo and Williamson (2007), whose influential account of causal inference in epidemiology is otherwise favorable to Hill's criteria, single out the specificity criterion as problematic on the ground that it assumes a one-to-one model of disease causation.

Now, insofar as they are directed at Hill's actual position, these criticisms have the wrong target. As Woodward (2010: 310) notes, Hill's text makes it clear that specificity of association is not a necessary condition or even a general rule. Thus Hill points out that even in his example nickel exposure is associated with not one but two distinct forms of cancer, and cites unpasteurized milk as a cause of a wide variety of diseases. Moreover, he also states explicitly that "one-to-one relationships are not frequent" and that "multicausation is more likely" (1965: 297). Instead, Hill's actual position on the evidential value of specificity is that

> *if specificity exists* we may be able to draw conclusions without hesitation; if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence. (1965, 297; my emphasis)

---

[5] See Koch (1987). Koch's postulates are often presented as embodying a `monocausal' model of disease, on which a given contagious disease has a single cause (a given pathogen). (See Broadbent 2009 and Fuller 2018.) As Ross and Woodward (2016: 41) emphasize, the postulates also assume that the relevant disease is the only pathology caused by the relevant pathogen. Thus in effect his postulates embody the assumption that causal relationships in contagious disease etiology are one-to-one specific.

[6] See Ross and Woodward (2016: 42) for discussion of that point. For instance, as they note, on the background assumption that every instance of a disease has a cause, if we find a case of disease D without presence of a pathogen C Koch's postulates let us immediately conclude that C does not cause D.

[7] See e.g. Krieger (1994), and Rothman and Greenland's `causal pie' model of disease causation (Rothman and Greenland 2005). More recently, multilevel modeling and complexity approaches to causal inference have emphasized the idea that chronic disease causation involves variables at different levels that interact in complex ways with one another (see e.g. Galea et al. 2010).

That is, Hill thinks that in cases where specificity of association is observed, we may conclude that we have causation, at least in the absence of countervailing considerations. But lack of specificity is not conclusive evidence against a causal interpretation of the correlation. (This is implied by the second part of the sentence just quoted, which suggests that if specificity is not found, other considerations can help decide for or against a causal conclusion.) In other words, his view is not that specificity of association is necessary or the norm, but only that it is an especially telling sign of causation.[8] In fact, as indicated above, Hill seems to think that specificity is rare. Still, even so, it is in his view a good idea to check if it is present, since we might thereby acquire conclusive evidence that we are dealing with a causal relationship – something that is very hard to obtain in epidemiology. So understood, Hill's criterion is perfectly compatible with the point that multicausation is the rule for chronic diseases.[9]

But even though the critics of specificity cited above miss their mark, this is not the end of the matter. For one thing, Hill's actual view is itself in need of defense. Why, after all, should a high degree of specificity be especially strong evidence that the exposure causes the outcome, at least if the exposure in question is the sort of `environment and lifestyle' factor with which Hill was concerned? In section III I will offer a defense of Hill's view that answers this question, and explains why specificity considerations are especially relevant in environmental, lifestyle and occupational epidemiology (though perhaps not in other sub-fields of epidemiology, as we will see).

A further issue is that one finds several instances in contemporary epidemiology in which the *non*-specificity of an association is used as a reason to reject a causal interpretation of the association. For instance, Petitti et al. (1986) reject the hypothesis that hormone replacement therapy (HRT) prevents cardiovascular disease on the ground that HRT is associated with a lesser risk for many other pathologies as well. Another example concerns the correlation between use of cimetidine (an antihistaminic used to treat gastric ulcers) and stomach cancer. A causal

---

[8] There are similarities here between Hill's stance and Woodward's own view of one-to-one specificity. In both cases, one-to-one specificity is seen as a property that causal relationships can (but need not) have and which makes relationships that have it significant in certain respects. One difference is that Hill focuses on the evidential significance of specificity, whereas Woodward emphasizes other reasons why one-to-one specific relationships are important, especially their practical relevance (Woodward 2010: 314-6).

[9] As we will see in section 3, a further consideration here is that the degree of specificity of an association need not be very high for it to provide evidence of causation.

interpretation of this correlation has been rejected on the ground that antacid use is also associated with higher prevalence of gastric cancer (Schumacher et al., 1990). And *that* pattern of inference, which Hill's criterion does not license, may seem to rely on a dubious `one cause – one effect' assumption, as critics of specificity allege. (Why couldn't HRT, like smoking, be involved in the etiology of many different medical outcomes?) In section IV, I will argue that it is in fact often reasonable to reject a causal interpretation of a correlation on ground of its non-specificity, as in the examples just cited.

II. Prior defenses of specificity

I will start with a brief look at earlier defenses of the specificity criterion found in the philosophical and epidemiological literatures. One is due to Alexander Bird (2011: 243). Bird starts by noting that to establish that a correlation between a risk factor A and a medical outcome B is due to A causing B (hereinafter the *causal hypothesis*), one must rule out certain alternative hypotheses: that the A-B correlation is due to a confounding factor, that it is due to B causing A, and that it is merely a statistical fluke. According to Bird, specificity of association provides evidence of causation by decreasing the probability of the fluke hypothesis. Roughly, the idea is that a fluky correlation might occur if some factor C causes B, and is by chance highly prevalent among people exposed to A. But if A is the sole exposure associated with B this possibility can be ruled out, since it would be highly implausible for C to be by chance highly prevalent among subjects exposed to A but not among other categories of the population. (This reasoning concerns specificity of exposure, but similar considerations apply for specificity of outcome.) While Bird's reasoning seems correct, it does not vindicate Hill's actual position. On Bird's view, specificity considerations do not help rule out the hypothesis that the correlation between A and B is real but due to confounding, which is often the most serious contender to the causal hypothesis in epidemiology. Accordingly, on his account specificity cannot be a conclusive reason in favor of the causal hypothesis (except in unlikely circumstances where the confounding hypothesis is already ruled out but the fluke hypothesis isn't). The defense I will present, by contrast, will show that specificity can help rule out the possibility of confounding and thereby leave the causal hypothesis as the only serious contender.

In epidemiology, Noel Weiss (2002) has offered an influential defense of the evidential value of specificity considerations. (As noted above, recent epidemiological studies that appeal

to specificity considerations often refer to that paper to justify their use of the criterion.) Weiss presents a number of actual and hypothetical cases in which specificity of association provides strong evidence of causation. One of his examples concerns the association between wearing a bike helmet and reduced risk of head injury. This correlation might be due to confounding rather than a real preventive effect of helmets: perhaps more careful riders are more likely to wear helmets, and also independently less likely to have accidents. That hypothesis, however, implies that association between helmet use and head injuries should be non-specific: helmet-wearers should be at lower risk for other injuries as well. By contrast, the causal hypothesis implies specificity of association, as any preventive effect of helmets should be restricted to head injuries. Hence, in this case specificity of association would be strong evidence of causation. Weiss's paper has been influential, and has contributed to partially rehabilitating the specificity criterion among contemporary epidemiologists. Yet while his examples are convincing, his case for specificity has important limitations. In the biking example, the inference from specificity to causation requires a large amount of background knowledge (e.g., about the mechanism by which helmets work). One may therefore be tempted to conclude, as the epidemiologist Michael Höfler does in a critical discussion of Weiss, that "the consideration of specificity appears to be useful only when a causal system is simple and the knowledge about it is largely certain" (Höfler, 2005). More importantly, Weiss's defense proceeds entirely by examples, and thus only shows that there exist cases in which specificity of association is evidence of causation. But to fully vindicate Hill's position and existing uses of the specificity criterion in the epidemiological literature, one would need to offer a principled reason to think that specificity of association is *generally* strong evidence of causation in epidemiology. In the next section I will present such a rationale.

III. Specificity and social confounding

In a nutshell, this rationale goes as follows. When seeking to establish whether a given environment and lifestyle factor A causes medical outcome B, often the main issue is to rule out the possibility that the A-B correlation is due to social confounding instead. In this context, specificity of association is strong evidence of causation because it helps exclude that possibility. This idea is not entirely new; a similar one is expressed by epidemiologists Davey Smith and Ebrahim, who write that "when exposures are related with a wide variety of outcomes it is likely

that confounding by socially patterned behavioural and environmental factors is at play" (Davey Smith & Ebrahim, 2003, 3). But they do not expand on this suggestion, nor does it appear to have had much influence on the subsequent epidemiological literature (or the philosophical literature on Hill's criteria, for that matter). But the idea, suitably elaborated, offers a principled and general rationale for Hill's position, as I will now argue.

First, a terminological point: here I use `social factor' in a broad sense that includes but is not limited to what epidemiologists and social scientists call socioeconomic status (SES). As standardly defined, SES is a composite variable that incorporates a number of dimensions such as level of income, wealth, social and cultural capital, and occupation. In addition to SES, social factors as I understand the term include factors that affect one's position in social structure (for instance, race and gender), other factors involving relationships to other people (such as level of social support, family structure, etc.), and psychological traits such as health-conscious attitudes and degree of risk aversion. (The inclusion of psychological factors might seem to stretch the term `social' beyond its usual meaning, but is in fact not inappropriate since factors such as concern for health and risk tolerance are themselves strongly influenced by socioeconomic position.)

On the view I propose, the key fact that underlies the value of specificity considerations is that social confounding is a pervasive concern in environmental, lifestyle and occupational epidemiology. Because environment and lifestyle factors (e.g. exposure to a chemical, pollution, diet, physical activity, smoking, etc.) are clearly influenced by social considerations, when one is found to be correlated with a medical outcome, there is virtually always a serious possibility that the correlation might be due to social confounding (i.e., the correlation arises not because the exposure causes the outcome, but only because both are effects of one or more social factors). As an illustration, note that virtually all examples of exposure-outcome correlations mentioned so far in the paper might conceivably be due to background social factors. Vitamin D deficiency might be correlated with multiple sclerosis because socioeconomic status affects one's risk of developing multiple sclerosis, and also independently influences diet and daylight exposure (and hence vitamin D intake). Asbestos and lung mesothelioma might be correlated merely because low SES increases the risk of both asbestos exposure and lung disease. Cannabis use might be correlated with low academic performance simply because underprivileged students tend to do poorly in school, and also have higher rates of recreational drug use; and so on. Of course, to

address the concern of social confounding, one might try to control for all potentially relevant social factors in one's statistical analysis, provided that data about those variables are available. But there are generally so many social variables that might induce confounding that it is hard to control for all of them; so that even if the correlation between exposure and outcome persists after adjusting for social factors, the possibility that residual confounding is at play often remains a serious one.

It is in this context that specificity of association shows its value: specificity, if observed, helps exclude the hypothesis of social confounding. The reason, in turn, is that social factors are known to have a causal influence on a multitude of medical outcomes, and are in that respect highly non-specific medical causes. If an exposure-outcome relationship is due to social confounding, one should thus find the correlation to be non-specific (i.e. the exposure should be correlated with many other outcomes as well). Conversely, if the association is observed to be specific, this is a good reason to conclude that it is not merely due to social confounding.

The best evidence for the fact that social factors are non-specific medical causes comes from research on the "social determinants of health" conducted in the last 30 years[10], which provides considerable evidence that low SES and social disadvantage more generally is associated with higher risks of cancer, diabetes, cardiovascular disease, mental illness, ulcers, physical injuries, and many other pathologies. That research also provides evidence that these correlations reflect a causal influence of social factors by identifying and documenting several mechanisms linking social factors to health. Specifically, there is extensive evidence that social factors affect health outcomes via at least four different pathways: by affecting access to health care, by influencing the adoption of health-conscious behaviors, by influencing environmental exposure to toxins and pollution, and through allostatic load (the "wear-and-tear" of the body caused by chronic psychosocial stress). These pathways are all clearly non-specific insofar as they do not affect one type of pathology in particular, but should be expected to have generic effects on all aspects of health.[11]

While contemporary research in social epidemiology provides excellent evidence that social factors are causes of many pathologies, it should be noted that knowledge of this fact long

---

[10] See Galea (2007), Berkman et al. (2014) and Braveman & Gottlieb (2014) for recent overviews.
[11] This is true also of allostatic load, which is thought to involve physiological changes in multiple bodily systems and hence to have a cascade of effects on health.

predates that research. That poverty, social inequalities and so on are associated with many illnesses has been known since the study of population health emerged in the 17th century, and has been a dominant theme in epidemiology and public health advocacy ever since (see e.g. Rose et al. 2008). Likewise, the idea that certain psychological attitudes have a global effect on health is fairly commonsensical and can be regarded as part of folk psychology. We all know, for instance, that more careful and health-conscious people tend to be at lesser risk of many diseases (or injuries, as in Weiss's example of the helmet wearers). In that respect, the fact that social factors tend to cause many different medical conditions is part of epidemiological (and to an extent, folk) background knowledge about disease etiology.

Now, in itself the fact that specificity of association helps rule out the possibility of social confounding does not mean that specificity is sufficient to establish that the exposure causes the outcome. Reverse causation (the outcome causing the exposure) might also be a possibility[12], or there may conceivably be other potential sources of confounding besides social factors that could explain the correlation. But often enough, the hypothesis that social confounding is at play is the only serious alternative to the hypothesis that the exposure causes the outcome. This is plausibly so in actual examples where epidemiologists have actually appealed to specificity considerations, e.g. nickel and nose and lung cancer, or vitamin D deficiency and multiple sclerosis.[13] And when social confounding is the only alternative, specificity of association is a conclusive reason in favor of the causal hypothesis, at least in the absence of other countervailing considerations. In that sense, Hill's claim about the evidential value of specificity in environmental, lifestyle and occupational epidemiology is generally correct.

This rationale for Hill's view reinforces the point that the specificity criterion does not presuppose that disease causation must be one-to-one specific. On the contrary, its justification relies explicitly on the fact that certain causes of medical outcomes (namely social factors) are non-specific, i.e. affect multiple health outcomes at once. In fact, on my proposal the value of the criterion does not rest on any general theory of disease causation, but on our background knowledge concerning the way in which social factors causally influence health. In that respect, my proposal fits well with Worrall's (2011) interpretation of Hill's criteria. On Worrall's

---

[12] One example where reverse causation is a serious possibility is the correlation between drug use and mental illness.

[13] Note in particular that in those cases reverse causation can be rejected out of hand. Certainly lung and nose cancer could not cause exposure to nickel, nor could vitamin D deficiency be caused by subsequent multiple sclerosis.

account, the function of Hill's criteria is to point to features of epidemiological background knowledge that can help address the pervasive problem of underdetermination of causal facts by observational data. Worrall himself does not discuss specificity of association, but my account shows that his claims apply to that criterion as well. Note also that on the story I proposed, the degree of specificity of an association need not be very high for it to provide evidence of causation. All that is needed is that the exposure-outcome association be *more* specific than we should expect on the hypothesis that it is driven by socioeconomic confounding. That condition may well be satisfied even if the exposure is associated with a number of other outcomes besides the one under consideration. This is further evidence that Hill's criterion is fully compatible with the fact that many-to-many causation is the rule in epidemiology.

My account of the specificity criterion also sheds light on the question of its scope. It explains why specificity considerations are particularly relevant in environmental, lifestyle and occupational epidemiology, which were the focus of Hill's address. This is because ruling out social confounding is an especially salient concern in those fields, as the factors they investigate are especially susceptible to social influence. This is not to say that specificity considerations have no application in other sub-fields of epidemiology. For instance, the specificity criterion has been profitably used in clinical epidemiology to evaluate whether post-treatment surveillance effectively reduces breast cancer mortality (Lash *et al*. 2005). Here appealing to specificity considerations makes sense, as social confounding is a salient possibility (higher SES patients both have more access to surveillance and receive better treatments). But my rationale also implies that specificity considerations may be of little use in sub-fields of epidemiology that deal with exposures that are less susceptible to social confounding. As an example, consider genetic epidemiology of mental illness. It is a priori highly unlikely that an association between a gene and a mental illness is due to social confounding, so that even if we found the association to be specific this may not provide special evidence of causation – or if it does, it would likely be for very different reasons than in the case of environment and lifestyle factors. In addition, whereas some environment and lifestyle factors *are* relatively specific, it may well be that virtually all genes involved in mental illness causation are causally relevant to many different conditions, as evidence from genome-wide association studies seems to indicate (see e.g. Uher & Zwicker 2017). So another reason why Hill's specificity criterion may not be relevant in genetic epidemiology is that specific associations may be entirely absent in that domain. Social

epidemiology is another sub-field in which specificity considerations are likely to be of little relevance, and for the same reason: since social factors such as SES, family structure and so on are known to causally affect multiple health outcomes at once, we have no reason to expect specific associations between those factors and medical outcomes, so that there is little point in checking for specificity. These remarks show that proper application of the specificity criterion requires good judgment on the investigator's part, and a good sense of whether the criterion is likely to be of help in the causal inference problem at hand. (This is true of all of Hill's criteria, as he himself insisted.) They also serve as a reminder that the methodology of causal inference in epidemiology is best seen as pluralistic, i.e. as relying on the judicious use of diverse tools that need not work in all circumstances.[14]

To be clear, although my account of the specificity criterion implies that specificity considerations may not be particularly useful in social epidemiology, it does not imply that social factors are not proper causes of health outcomes.[15] As noted above, my account presupposes that social factors do play an important role in disease causation: it is precisely because social factors play a pervasive role in disease causation that it is necessary to control for them when assessing whether a given environment and lifestyle factor has a certain medical effect.[16] Nor does my account of the specificity criterion imply that social factors fall outside of the proper province of epidemiology: again, it merely implies that the criterion might not be of much use for studying the health effects of social factors. (This is one area where it is crucial to interpret the criterion in the right way: if we misread it as saying that specificity is necessary for disease causation, the criterion would indeed mistakenly exclude social factors from the scope of epidemiology.)

I will close this section by noting several implications of my account for the methodology of epidemiology. First, my reconstruction of Hill's criterion highlights the fact that methods of causal discovery in epidemiology rely in part on `localized epistemic tools' – a concept I borrow from Currie's (2015, 2018) illuminating discussion of methodological strategies in history. As Currie uses the phrase, a localized epistemic tool is a procedure or method that is `custom-built' to solve a quite specific epistemic predicament faced by a scientific discipline, and which does so

---

[14] On methodological pluralism, see Vandenbroucke et al. (2016).

[15] Thanks to a reviewer for pressing me to clarify this point.

[16] Indeed, properly understood, the specificity criterion is compatible with views that ascribe a central status to social factors in disease etiology, e.g. the view that social factors are "fundamental causes" of health outcomes (see e.g. Link & Phelan, 1995).

by exploiting particular features of the epistemic problem in question. If my account is correct the specificity criterion fits that description. Starting with the study of smoking in British doctors and the Framingham heart study (which have been paradigmatic for the field), contemporary epidemiology has devoted and still devotes considerable attention to the role of environment and lifestyle factors in chronic diseases. As a consequence, the problem of ruling out social confounding has become an important epistemic issue in the discipline.[17] The specificity criterion helps solve this highly specific predicament by seizing on a particular distinguishing feature of social causes of medical outcomes – their lack of specificity. My discussion thus suggests that localized epistemic tools have an important role to play in the methodology contemporary epidemiology (or at least a large portion thereof), just like they do in the historical sciences according to Currie.

In addition, my defense of the specificity criterion has interesting implications for the question of the relative merits of randomized controlled trials (RCTs) and observational studies in epidemiological causal inference.[18] A well-known argument in favor of RCTs is that they control for all potential confounders, both known and unknown. (Randomization makes it highly probable that all such confounders are distributed equally in the treatment and control groups.) RCTs thus have high reliability even in cases where causal structure is largely unknown. By contrast (the argument goes), in observational studies one can only control for those factors that are already known or suspected to be potential confounders. Thus, observational inference is reliable only in cases where we already have ample knowledge about the causal structure of the phenomenon investigated. Worrall (2002) has argued that the deconfounding powers of RCTs are not in fact as great as this argument claims, and that background causal knowledge is still required to draw reliable causal conclusions from randomized experiments. If my account is correct it also turns out that, contrary to what this argument presupposes, observational causal inference in epidemiology can in fact proceed with fairly limited background knowledge, at least in certain cases. If specificity of association is observed, one can confidently rule out the

---

[17] A concern for confounding, including by social factors, was also present in the study of infectious diseases that were the primary focus of late 19th and early 20th epidemiology. But one key difference is that the kinds of causal hypotheses studied by epidemiology at the time, which pertained to a potential causal link between a certain microorganism and a certain disease, were much more amenable to experimental study than the causal hypotheses examined in contemporary epidemiology. (The key experimental technique was animal inoculation of the relevant microorganisms.) This meant that the concern that the correlation is due to confounding was much easier to address and rule out.

[18] See e.g. Worrall 2010, Bird 2011 and Vandenbroucke et al. 2016 for discussion of that issue.

possibility that social confounding is at play, even without having much of an idea about which exact social factor(s) - education, income, or what have you - might plausibly induce confounding in the case at hand. (In that respect, Höfler (2005) is wrong to claim that specificity considerations are useful only in cases where the structure of the system is largely certain.) Admittedly, as noted above, this establishes causation only if social factors are known to be the only plausible source of confounding. So unsurprisingly, some background knowledge about the structure is required – but nevertheless much less than one might have supposed.

Finally, on my account the specificity criterion nicely provides a nice illustration of the value of "high-level" causal principles for observational causal inference. As a number of authors have noted (e.g. Griffiths & Tenenbaum 2007 and Henderson et al. 2010), high-level causal principles – for instance, principles stipulating which kinds of variables can cause other kinds of variables - have a useful and perhaps indispensable role to play in causal inference: by putting substantial constraints on the range of possible lower-level causal theories, they can considerably speed up causal learning from observation. The assumption that socioeconomic factors are non-specific causes of pathologies can be conceived as one such highly general causal principle that facilitates inference in certain cases by ruling out a large number of possible confounding causal structures in one fell swoop.[19]

IV. The evidential import of non-specificity

According to Hill, specificity is strong evidence of causation, but lack of specificity by itself does not tell us much. Yet as mentioned earlier, epidemiologists sometimes use the *lack* of specificity of a correlation as a reason to conclude that it is non-causal. Examples include the correlation between HRT and cardiovascular disease (Pettiti et al. 1986), and between cimetidine prescription and stomach cancer (Schumacher et al. 1990). And that particular use of specificity considerations may seem to rely on the outdated assumption that disease causation must be one-to-one specific, as critics of specificity allege. Moreover, this pattern of reasoning has a bad track-record: it is the very one that led a number of epidemiologists to reject a causal

---

[19] As Ross and Woodward (2016: 42) note, the specificity assumptions encapsulated in Koch's postulates play a similar evidential role, by greatly constraining the range of hypotheses about the causes of contagious diseases that are compatible with the data.

interpretation of the correlation between smoking and lung cancer when evidence for it emerged in the 1950s (see e.g. Berkson, 1958). Can anything be said in its favor?

One point to note is that in some of the relevant examples, it turns out on further inspection that non-specificity *per se* is not the reason why the causal hypothesis is rejected. For instance, in the case of HRT and cardiovascular disease, the key argument is not that HRT is associated with many other medical outcomes as well, but that some of those outcomes are ones that HRT could not possibly cause, e.g. lower risks of violent death. The only reasonable explanation of the association between HRT and that outcome is that there is a confounder – presumably socioeconomic status. Since socioeconomic status also plausibly affects the risk of cardiovascular disease, the simplest explanation is that the correlation between HRT and cardiovascular disease is due to this confounder as well. This reasoning (which is perfectly reasonable) is an instance of the method of "negative controls" (Lipsitch et al. 2010). In negative control, one checks whether a correlation between an exposure A and an outcome B may be due to a confounder C by examining whether A is also associated with an outcome B' that could reasonably be caused by C but not by A. (Alternatively, one can also check if B is associated with some other exposure A' that could not possibly cause B, but is known to be caused by C.) If such associations are found, the simplest and therefore most likely hypothesis is that the A-B correlation itself is due to confounding. As Lipsitch et al. (2010) observe, negative control (which finds its origins in experimental biology) and the idea of specificity of association are closely related to one another. Nevertheless, in negative control, it is not the lack of specificity of the correlation *per se* that drives the rejection of the causal hypothesis. Instead, it is the fact that the exposure of interest could not possibly cause the other outcomes with which it is related.

However, not all the studies that use (or seem to use) lack of specificity as a reason to reject the causal hypothesis can be plausibly regarded as instances of the method of negative control. For instance, Schumacher et al. (1990) argue that the correlation between cimetidine use and stomach cancer is likely to be non-causal because patients who take antacids (which are also commonly prescribed for gastric ulcers) have a higher risk of stomach cancer too. By contrast to the correlation between (e.g.) HRT and violent death, the possibility that antacids cause stomach cancer cannot be dismissed out of hand. So it is not the fact that stomach cancer is also associated with other exposures that could not possibly cause it that drives the reasoning.

16

Likewise, Stiby et al. (2015) argue that the correlation between adolescent cannabis use and poor educational outcomes is unlikely to be causal on the ground that tobacco use is also negatively associated with academic performance. Here also it is hard to see this inference as an instance of negative control, since the hypothesis that smoking negatively affects school performance is not in itself completely implausible. In those examples, it does seem to be lack of specificity *per se* that drives the rejection of the causal hypothesis.

But we can see that this pattern of inference is reasonable once we take into account two facts. First, in all the examples under consideration the causal hypothesis is rejected in favor of the hypothesis that the relevant correlation is due to a confounding factor, often (but not always) a social one. That is, the inference is contrastive in nature. In the cannabis and academic performance example, that confounder might be any of the many social factors that plausibly affect both drug use and school performance. In the case of cimetidine, the likely confounder is early (undetected) stomach cancer causing patients to complain of stomach problems and hence being prescribed anti-ulcer medication. (For smoking and lung cancer, some of the factors regarded as potential confounders included a nervous disposition, genetic profile, and health-conscious behavior.)

The second important fact concerns how epidemiologists evaluate specificity (and lack thereof). So far, I have assumed that specificity is only a matter of the number of other putative effects (or causes) with which a variable is associated. (This is in line with Woodward's (2010) notion of one-to-one causal specificity, on which the specificity of a putative causal relationship between *X* and *Y* is determined by the number of other putative effects of *X*, and other putative causes of *Y*.) Yet epidemiologists also take into account a further dimension in their assessments of specificity: namely, the extent to which the putative effects or causes of a variable are "heterogeneous" or "disparate". The more they are, the less specific the relationship. Thus Berkson claimed that the relationship between smoking and lung cancer is non-specific because smoking is associated with "*so wide* a variety of categories of disease" (1958: 32; my emphasis). Likewise, Hill says that unpasteurized milk is a non-specific cause of disease insofar as it "can produce *such a disparate galaxy* as scarlet fever, diphtheria, tuberculosis, undulant fever, sore throat, dysentery and typhoid fever" (1965: 297; my emphasis). And Davey Smith and Ebrahim describe non-specificity (of outcome) as the fact that "exposures are related with a *wide variety*

of outcomes" (2003: 4; my emphasis). Heterogeneity considerations are also implicitly at play in the cannabis and cimetidine studies. In both cases the relevant correlation is judged to be non-specific on the ground that the outcome is also correlated with another exposure (smoking in one case, antacid intake in the other). But if number of variables was the only consideration that played a role in assessments of specificity, surely the presence of just one other correlate would not be enough to make the relationship non-specific. (Remember that Hill deemed nickel exposure to be a highly specific cause even though it was associated with *two* distinct types of cancer.) Presumably, in those examples the claims of non-specificity are based on the fact that the relevant exposures are very different from one another in salient respects.

Such heterogeneity assessments appear to rely on a variety of considerations. For heterogeneity of outcomes, these considerations include variety in sites of occurrence, differences in symptoms and severity, and acute vs. chronic nature. At any rate this is what Hill's list of disparate effects of contaminated milk suggests. (Scarlet fever and tuberculosis affect different sites of the body; diphtheria and dysentery have very different symptoms; sore threat is less severe than dysentery; undulant fever is often chronic whereas scarlet fever is acute.) Turning to exposures, note that the active substances of cannabis and tobacco are fairly chemically dissimilar to one another: for instance, nicotine is an alkaloid while THC isn't. And cimetidine and antacids have different modes of action on stomach ulcer: cimetidine prevents the secretion of stomach acid by blocking the release of its trigger histamine, whereas antacids work by suppressing stomach acid already present in the stomach. Likewise, THC and nicotine are well-known to have different proximal side effects. So in the case of exposure heterogeneity relevant factors presumably include heterogeneity in chemical structure, as well as the extent to which the known causal profiles of the exposures already differ from one another. These lists of considerations are not meant to be exhaustive, and it is likely that other factors also come into play.

In sum, then, my proposal is that given how epidemiologists assess heterogeneity, the fact that A is correlated not only with B but with other heterogeneous variables as well indicates that the hypothesis that A causes B does a poor job at predicting A's correlations with those other variables. Provided that there is some confounding hypothesis that does a better job at predicting those correlations, and that this hypothesis is not entirely implausible to begin with, then it is

rational to reject the causal hypothesis. In that respect, the non-specificity of an association can be a good reason to regard a causal interpretation of the correlation as highly unlikely, at least in the absence of countervailing considerations.

If this argument is correct, the pattern of inference at work in those studies does not presuppose a one cause – one effect model of disease causation. Indeed, the confounding hypothesis in favor of which the causal hypothesis is rejected explicitly posits a cause of many effects (e.g. in the cannabis study the relevant socioeconomic confounder must cause both cannabis and tobacco use, as well as poor school outcomes). This raises a puzzle. Why should heterogeneity considerations be a mark against the causal hypothesis, but not the confounding hypothesis? We can glean an answer by further examining the case of cannabis use and academic performance. The puzzle in that specific example is why the fact that cannabis and tobacco are heterogeneous substances should be a reason to reject the hypotheses that they each affect school performance, but not a reason to reject the hypothesis that they are both effects of the same social factor. The answer, I propose, has to do with mechanistic knowledge (or lack thereof). On the one hand, the mechanisms by which cannabis and tobacco might produce cognitive and motivational effects are not very well understood, so that mechanistic knowledge is not a helpful guide to evaluate how plausibly they might affect school performance. Hence we have to rely on something else than mechanistic information to make that assessment, and this is where heterogeneity considerations become relevant. By contrast, we already have ample knowledge about the mechanisms by which difficult social conditions can lead to recreational substance use. (Some of that knowledge comes from through our ordinary experience and folk theory of the social world. And some of it comes from research on the social determinants of health, which has provided detailed evidence concerning the mechanisms by which social factors affect substance use.) Thus, because we already have ample mechanistic knowledge to establish the hypothesis of a social common cause of cannabis and tobacco use, the fact that these effects are in certain aspects heterogeneous does little to decrease the credibility of that hypothesis. Indeed, because we know that cannabis and tobacco use are caused by the same social mechanisms, we may be inclined to regard them as fairly homogeneous factors after all, at least from a social if not from a biochemical standpoint. This is reflected in the fact that from a sociologist's point of view it may well make most sense to lump those two factors into a single category of recreational substance

use – further distinctions within that category may appear relatively invidious as far as social science is concerned.

More generally, then, I suggest that heterogeneity considerations are most relevant when we have little mechanistic knowledge by which to evaluate the plausibility of a postulated causal link between an environmental or behavioral risk factor A and an outcome B. In those circumstances, looking at whether the exposure or outcome are associated with a wide variety of other variables is a helpful heuristic to assess the plausibility of a causal link between A and B. (Thus it is no surprise that such considerations were prominent in early discussions of smoking, when little was known about the mechanisms by which smoking might affect health.) But for certain factors, especially socioeconomic ones, we already are very familiar with the mechanisms by which they produce their effects (because of our folk knowledge, and/or because the existence of those mechanisms has already been documented by prior epidemiological research). That knowledge of mechanisms will often enough to satisfactorily evaluate the plausibility of a causal link involving a socioeconomic factor, so that heterogeneity considerations now become evidentially irrelevant. In fact, as the example of cannabis and tobacco shows, if we know that a number of variables are effects of the same social mechanism, we may be inclined to regard them as fairly homogeneous after all, at least from a social science standpoint. This implies that heterogeneity assessments may be relative to a discipline and the kind of mechanism on which it focuses: a social scientist may not evaluate a set of variables as heterogeneous in the same way that a biologist or medical scientist would. It also suggests that heterogeneity assessments themselves are influenced by the amount of mechanistic knowledge one has, so that variables known to be the product of the same mechanism are thereby judged to be more similar to one another. (As a further case in point, consider various forms of tuberculosis such as phthisis, scrofula and Pott's disease. We now regard those conditions as instances of the same disease on the ground that they all result from infection by *Mycobacterium tuberculosis*, as Koch discovered. But before Koch's work one might well have judged that set of conditions to be highly disparate, in particular because they occur at distinct bodily sites – lungs, neck and spine respectively.[20])

---

[20] See Lange (2007: 268).

Let me close this section with a number of remarks. First, an obvious worry with my argument is that it seems to license the patently wrong inference by which some epidemiologists rejected the hypothesis of harmful effects of smoking on ground of non-specificity. One response is that this may not be such a bad consequence, provided we keep in mind that the inference from non-specificity to confounding is reasonable only in the absence of countervailing considerations. So while my argument might warrant initial rejection of the hypothesis of harmful effects of smoking in the 1950s, it also entails that this stance quickly became unreasonable once further evidence of tobacco carcinogenicity accumulated and further research failed to confirm the existence of a confounding factor that could explain the relevant correlations. But there is also a stronger response available, which is that considerations of non-specificity properly understood did not even warrant initial rejection of the hypothesis. The claim that smoking is an instance of an exposure that is non-specifically associated with a wide variety of outcomes presupposes that smoking should be regarded as a single exposure. But as was quickly noted by several participants to the debate about smoking and lung cancer in the 1950s and 1960s, tobacco smoke is known to contain a large number of different chemical substances, so that it is misleading to treat tobacco as a single exposure. And if smoking should in fact be regarded as a multiplicity of exposures at once, specificity considerations in themselves do not let us conclude anything about the likelihood of a causal influence of tobacco smoke on lung cancer and other medical outcomes. Interestingly, this was exactly the position of the authors of the 1964 Surgeon General's report on smoking and health, who endorsed the principle that non-specific associations are less likely to be causal, but also argued that this principle did not apply to the case at hand since smoking represented a number of different exposures at once. Hence, as the report puts it, even taking specificity considerations into account "it would not be surprising to find that the diverse substances in tobacco smoke could produce more than a single disease" (U.S. Department of Education, Health and Welfare, 1964: 185). Arguably, then, it was only careless application of specificity considerations that led some epidemiologists to reject a causal link between smoking and cancer in the 1950s.

The second remark concerns Woodward's concept of one-to-one specificity. Whereas specificity of association has been presented by Woodward (2010) and others (e.g. Bourrat, 2018) as a variety of one-to-one specificity, my argument shows that specificity of association includes a dimension not incorporated in Woodward's notion. One-to-one specificity has to do

only with the number of other variables with which a variable is causally associated, whereas when evaluating the specificity of a putative cause epidemiologists also take into account the extent to which these other variables are similar to one another in certain respects. Moreover, that dimension of specificity of association is crucial to understand the epistemic role that the notion plays in the methodology of epidemiology. Thus an important lesson here is that we need a concept of specificity richer than Woodward to explicate specificity of association and its methodological import.[21] But another lesson of my discussion is that the two dimensions of specificity are not entirely separate from each other. Obviously, the number of a putative effects (or causes) that a variable has depends on how we count effects, i.e. on our choice of variables.[22] When a single variable amalgamates a number of highly heterogeneous factors, there is theoretical pressure to split that variable into separate ones (as in the example of smoking), with the consequence that a causal structure that initially appeared highly non-specific might not do so anymore once this conceptual split has taken place.

This last point highlights an important fact. While I have focused chiefly on showing why and how non-specificity can be evidence against causation, an observed lack of specificity can be evidentially significant in other ways as well. (I owe this remark and the subsequent remarks to a reviewer.) In the case of smoking, non-specificity was relevant because it drew attention to the fact that a given variable in fact covers a wide variety of heterogeneous causal factors. The reverse case is also conceivable: that is, if an exposure is found to be associated with a number of outcomes, this may indicate that those outcomes are not as different as initially thought. (That reaction might be especially appropriate if we have strong independent (e.g. mechanistic) evidence that some of these observed associations are indeed causal.) And there may still be other ways in which lack of specificity can be evidentially relevant. For instance, as the reviewer also noted, in certain contexts a lack of specificity may serve as a guide for further causal inquiry, by indicating that further factors besides the exposure are involved in producing the outcomes and await to be discovered. Examination of these additional evidential roles of specificity considerations is beyond the scope of this paper, but clearly the issue is ripe for further investigation.

---

[21] In fact, specificity of association is not the only variety of specificity that integrates such a dimension of similarity. Other varieties of biological specificity do as well: see (Blanchard, 2022).
[22] See (Woodward, 2016) for a general discussion of variable choice in science and causal modeling.

V. Conclusion

Hill's specificity criterion has received unfairly bad press in epidemiology and philosophy of science, and he was right to claim that specificity considerations do have a valuable role to play in observational causal inference – one that is perhaps not sufficiently appreciated in contemporary epidemiology. On the account I presented, the criterion does not rely on any dubious `one-cause one-effect' assumption, as critics allege. In fact, it does not derive its legitimacy from any general account of disease causation; instead, the criterion is best regarded as a localized tool that helps address a peculiar epistemic problem that pervades contemporary epidemiology – the problem of establishing whether a correlation between an environmental or behavioral risk factor and a medical outcome might be due to social confounding. My account of the evidential value of specificity of association also shows that observational causal inference in epidemiology can proceed even with relatively limited background knowledge about the causal structure under consideration, by exploiting high-level causal principles that can rule out a great number of causal hypotheses at once. Another important lesson is that despite claims to the contrary specificity of association cannot be entirely explained in terms of Woodward's one-to-one specificity concept, as it concerns not only the number of variables with which a putative cause or effect is associated, but also the extent to which those variables form a `heterogeneous' or `disparate' set. This `heterogeneity' dimension, as we saw, is crucial to understanding the evidential import of specificity considerations in epidemiology.

References

Berkman, L., Kawachi, I., & Glymour, M. (2014). *Social Epidemiology* (Second Edition). New York: Oxford University Press.

Berkson, J. (1958). Smoking and Lung Cancer: Some Observations on Two Recent Reports. *Journal of the American Statistical Association*, *53*, 28–38.

Bird, A. (2011). The Epistemological Function of Hill's Criteria. *Preventive Medicine*, *53*, 242–245.

Blanchard, T. (2022). Host Specificity in Biological Control. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/721088

Braveman, P., & Gottlieb, L. (2014). The Social Determinants of Health: It's Time to Consider the Causes of the Causes. *Public Health Reports*, *129*, 19–31.

Broadbent, A. (2009). Causation and Models of Disease in Epidemiology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *40*, 302–311.

Currie, A. (2015). Marsupial Lions and Methodological Omnivory: Function, Success and Reconstruction in Paleobiology. *Biology & Philosophy*, *30*, 187–209.

Currie, A. (2018). *Rock, Bone, and Ruin: An Optimist's Guide to Historical Sciences*. Cambridge, MA: The MIT Press.

Davey Smith, G., & Ebrahim, S. (2003). 'Mendelian Randomization': Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease? *International Journal of Epidemiology*, *32*, 1–22.

Doll, R., Morgan, L. G., & Speizer, F. E. (1970). Cancers of the Lung and Nasal Sinuses in Nickel Workers. *British Journal of Cancer*, *24*, 623–632.

Freeman, M. D., & Kohles, S. S. (2012). Assessing Specific Causation of Mesothelioma following Exposure to Chrysotile Asbestos-Containing Brake Dust. *International Journal of Occupational and Environmental Health*, *18*, 329–336.

Fuller, J. (2018). Universal Etiology, Multifactorial Diseases and the Constitutive Model of Disease Classification. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *67*, 8–15.

Galea, S. (Ed.). (2007). *Macrosocial determinants of population health*. New York, NY: Springer.

Galea, S., Riddle, M., & Kaplan, G. A. (2010). Causal thinking and Complex System

>Approaches in Epidemiology. *International Journal of Epidemiology*, *39*, 97–106.

Greenland, S., & Rothman, K. (1997). *Modern Epidemiology* (Second Edition). Philadelphia:

>Lippincott Williams & Wilkins.

Griffiths, P. E., Pocheville, A., Calcott, B., Stotz, K., Kim, H., & Knight, R. (2015). Measuring

>Causal Specificity. *Philosophy of Science*, *82*, 529–555.

Griffiths, T. L., & Tenenbaum, J. B. (2007). Two Proposals for Causal Grammars. In A. Gopnik

>& L. Shulz (Eds.), *Causal Learning*. New York: Oxford University Press.

Henderson, L., Goodman, N. D., Tenenbaum, J. B., & Woodward, J. F. (2010). The Structure

>and Dynamics of Scientific Theories: A Hierarchical Bayesian Perspective. *Philosophy of

>Science*, *77*, 172–200.

Hill, A. B. (1965). The Environment and Disease: Association or Causation? *Proceedings of the

>Royal Society of Medicine*, *58*, 295–300.

Höfler, M. (2005). The Bradford Hill Considerations on Causality: A Counterfactual Perspective.

>*Emerging Themes in Epidemiology*, *2*, 11.

Koch, R. (1987). *Essays of Robert Koch*. New York: Greenwood Press.

Krieger, N. (1994). Epidemiology and the Web of Causation: Has Anyone Seen the Spider?

>*Social Science & Medicine*, *39*, 887–903.

Lange, M. (2007). The End of Diseases. *Philosophical Topics*, *35*, 265–292.

Lash, T. L., Clough-Gorr, K., & Silliman, R. A. (2005). Reduced Rates of Cancer-Related

>Worries and Mortality Associated with Guideline Surveillance after Breast Cancer

>Therapy. *Breast Cancer Research and Treatment*, *89*, 61.

Lawlor, D. A., Davey Smith, G., Kundu, D., Bruckdorfer, K. R., & Ebrahim, S. (2004). Those
Confounded Vitamins: What Can we Learn from the Differences between Observational
versus Randomised Trial Evidence? *The Lancet*, *363*, 1724–1727.

Lean, O. M. (2020). Binding Specificity and Causal Selection in Drug Design. *Philosophy of
Science*, *87*, 70–90.

Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of
Health and Social Behavior*, *Spec No*, 80–94.

Lipsitch, M., Tchetgen Tchetgen, E., & Cohen, T. (2010). Negative Controls: A Tool for
Detecting Confounding and Bias in Observational Studies. *Epidemiology*, *21*, 383–388.

Petitti, D. B., Perlman, J. A., Sidney, S., Pike, M. C., Ross, R. K., Henderson, B. E., … Bailar, J.
C. (1986). Postmenopausal Estrogen Use and Heart Disease. *New England Journal of
Medicine*, *315*, 131–136.

Rose, G., Khaw, K.-T., & Marmot, M. (2008). *Rose's Strategy of Preventive Medicine*. Oxford:
Oxford University Press.

Rothman, K. J., & Greenland, S. (2005). Causation and Causal Inference in Epidemiology.
American Journal of Public Health, *95*, S144-150.

Russo, F., & Williamson, J. (2007). Interpreting Causality in the Health Sciences. *International
Studies in the Philosophy of Science*, *21*, 157–170.

Schumacher, M. C., Jick, S. S., Jick, H., & Feld, A. D. (1990). Cimetidine Use and Gastric
Cancer. *Epidemiology*, *1*, 251–254.

Simpson, S., & van der Mei, I. (2019). Vitamin D Deficiency is an Etiological Factor for MS -
Commentary. *Multiple Sclerosis*, *25*, 641–643.

Smith, G. D. (n.d.). The biopsychosocial approach: A note of caution. In *Biopsychosocial Medicine*. Oxford University Press.

Stiby, A. I., Hickman, M., Munafò, M. R., Heron, J., Yip, V. L., & Macleod, J. (2015). Adolescent Cannabis and Tobacco Use and Educational Outcomes at Age 16: Birth Cohort Study. *Addiction*, *110*, 658–668.

Stotz, K., & Griffiths, P. E. (2017). Biological Information, Causality, and Specificity: An Intimate Relationship. In G. F. R. Ellis, P. C. W. Davies, & S. I. Walker (Eds.), *From Matter to Life: Information and Causality* (pp. 366–390). Cambridge: Cambridge University Press.

Susser, M. (1977). Judgment and Causal Inference: Criteria in Epidemiologic Studies. *American Journal of Epidemiology*, *105*, 1–15.

U. S. Department of Education, Health and Welfare. (1964). *Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service*. Washington D. C.: United States Public Health Service.

Uher, R., & Zwicker, A. (2017). Etiology in Psychiatry: Embracing the Reality of Poly-Gene-Environmental Causation of Mental Illness. *World Psychiatry*, *16*, 121–129.

U.S. Environmental Protection Agency. (2005). *Guidelines for Carcinogen Risk Assessment*. Environmental Protection Agency.

Vahtera, J., Kivimäki, M., Pentti, J., Linna, A., Virtanen, M., Virtanen, P., & Ferrie, J. E. (2004). Organisational Downsizing, Sickness Absence, and Mortality: 10-Town Prospective Cohort Study. *British Medical Journal*, *328*, 555.

Vandenbroucke, J. P., Broadbent, A., & Pearce, N. (2016). Causality and Causal Inference in

    Epidemiology: The Need for a Pluralistic Approach. *International Journal of*

    Epidemiology, *45*, 1776–1786.

Weber, M. (2006). The Central Dogma as a Thesis of Causal Specificity. *History and Philosophy*

    of the Life Sciences, *28*, 595–609.

Weiss, N. S. (2002). Can the "Specificity" of an Association be Rehabilitated as a Basis for

    Supporting a Causal Hypothesis? *Epidemiology*, *13*, 6–8.

Woodward, J. (2010). Causation in Biology: Stability, Specificity, and the Choice of Levels of

    Explanation. *Biology & Philosophy*, *25*, 287–318.

Woodward, J. (2016). The Problem of Variable Choice. *Synthese*, *193*, 1047–1072.

Worrall, J. (2002). What Evidence in Evidence-Based Medicine? *Philosophy of Science*, *69*,

    S316–S330.

Worrall, J. (2010). Evidence: Philosophy of Science Meets Medicine. *Journal of Evaluation in*

    Clinical Practice, *16*, 356–362.

Worrall, J. (2011). Causality in Medicine: Getting Back to the Hill Top. *Preventive Medicine*, *53*,

    235–238.

Yerushalmy, J., & Palmer, C. E. (1959). On the Methodology of Investigations of Etiologic

    Factors in Chronic Diseases. *Journal of Chronic Diseases*, *10*, 27–40.