Epistemic Causality and Evidence-Based Medicine

Federica Russo & Jon Williamson

Draft of October 14, 2010

Abstract

Causal claims in biomedical contexts are ubiquitous albeit that they are not always made explicit. This paper addresses the question of what causal claims mean in the context of disease. It is argued that in medical contexts causality ought to be interpreted according to the *epistemic theory*. According to this approach, causal claims tell us about which inferences (e.g., diagnoses and prognoses) are appropriate, rather than about the presence of some physical causal relation analogous to distance or gravitational attraction. It is shown that the epistemic theory has important consequences for medical practice, in particular with regard to the evidence-based causal assessment.

1   Introduction

In biomedical contexts, causal claims are not always made explicit. Largely due to the influence of the mantra 'correlation does not imply causation', claims about correlations and risk factors have tended to replace explicit causal claims. But there is a price to be paid for eradicating causal claims. Indeed, one might adopt another mantra, 'correlation does not license treatment': for one to intervene to treat A in order to alleviate B, A needs to be a cause of B, for otherwise the treatment will be pointless. Thus while the presence of umbrellas is correlated with rain, one would not intervene to ban umbrellas in order to prevent rain. Arguably, therefore, research papers in the biomedical sciences that draw conclusions about correlations are only of interest to the extent that those correlations are understood as supporting corresponding causal claims.

But in recent years causal claims have begun to be reintroduced into the sciences, largely thanks to formalisms that make the relationship between correlation and causation more precise (Illari et al., 2011). This rehabilitation of causal talk demands not only formal clarity, but also philosophical clarity in at least two respects.

First, it is incumbent on us to be clear about the kind of causal claim we are making. For example, a causal claim can be categorised according to whether it is a generic claim or a single-case claim. It may assert a generic fact about disease causation: e.g., according to recent discoveries, gastric ulcers may be caused by the presence of the bacterium *Helicobacter pylori*. These generic causal claims express general medical knowledge. On the other hand a causal claim can also be used to pick out single cases of disease causation. For example: *biopsy revealed*

that Helicobacter pylori *is a cause of Mr Jones' gastric ulcer.* These single-case causal claims typically concern diagnosis or prognosis in particular patients but may also concern facts of disease causation for a whole population at a certain time and place: e.g., cholera epidemics in London in 1854 were caused by human sewage contamination.

Second, we need to know what we are talking about when we make claims of the form 'A is a cause of B'. This is the focus of the present paper. In §2 we will introduce two standard kinds of interpretation of causal claims: 'A is a cause of B' is normally either interpreted as saying that A makes a certain sort of difference to B or as saying that there a certain mechanistic connection between A and B. We will argue that causal claims in medicine cannot be given either of these standard interpretations—rather, they need to be given an *epistemic* interpretation. The epistemic theory is introduced in §3 and then in §4 we explain the reasons for adopting an epistemic theory in medicine. In §5 we explore the consequences of such an account for evidence-based medicine.

2   Difference-making and mechanistic interpretations of causal claims

The question of how to interpret causal claims in medicine is of considerable interest to those working in medicine and epidemiology (see, among others, Susser (1991); Parascandola and Weed (2001); Rothman and Greenland (2005); Frumkin (2006); Charlton (1996); Elwood (1988); Rizzi and Pedersen (2002)). A notable attempt to categorise various possible interpretations is that of Parascandola and Weed (2001), who identify five different meanings of 'cause': production, necessary cause, sufficient-component cause, probabilistic causation and counterfactual. While all these meanings are are present in the burgeoning philosophical literature on causality, in recent years this literature has tended to compartmentalise theories of causality into two camps: difference-making accounts and mechanistic accounts.

Difference-making theories.   Difference-making theories typically come in three variants: probabilistic, counterfactual, and manipulationist or interventionist theories. All three theories can be labelled "difference-making" because what characterises the causal relation is that a cause makes a difference to the occurrence or level of its effects.

In probabilistic theories there are three main ideas: (i) positive causes raise the probability of their effect(s), i.e. $P(E|C) > P(E)$; (ii) preventatives, or negative causes, lower the probability of their effects, i.e. $P(E|C) < P(E)$; and (iii) if A and B are correlated, then either A causes B, or B causes A, or there is a common cause C that accounts for the correlation between A and B. See Williamson (2009) for an introduction to probabilistic theories of causality, and Reichenbach (1956), Suppes (1970), Eells (1991) and Arntzenius (2008) for more details, including discussion of the highly controversial issue of (iii), which is a version of the *principle of common cause*.

With a counterfactual theory—see, e.g., Lewis (1973, 2004)—to say that an event C is a cause of E is to say that C and E occurred but, had C not occurred, E would not have occurred either. This last counterfactual conditional

(a conditional with a false antecedent) helps to pick out a single-case causal relation by comparing what happens in the actual world with what happens in other possible worlds.

Manipulationist, interventionist and agency theories cash out causation in terms of the notion of intervention or experimentation. They claim that a variable C is a cause of variable E if, were we to change C, E would change accordingly, and the relation between C and E would remain invariant. See, for instance Woodward (2003); Hausman and Woodward (2004); Hausman (1997); Woodward and Hitchcock (2003).

Mechanistic theories   Mechanistic theories of causality say that C is a cause of E if there is a mechanism of the appropriate sort that links C to E (Williamson, 2011). Mechanistic theories are usually divided into process theories and complex-systems mechanisms theories.

Process theories take a mechanism to be a low-level physical process: a process which involves the transfer of a mark from C to E or the transfer from C to E of a conserved physical quantity such as angular momentum or electric charge (Reichenbach, 1956; Salmon, 1998; Dowe, 2000). An example of such a process is the signal from a remote control to open a garage door.

Complex-systems theories take mechanisms to consist of (typically high-level) entities and activities organised in a particular way so as to produce some phenomenon of interest (Machamer et al., 2000; Glennan, 2002, 2010; Craver, 2007; Bechtel and Abrahamsen, 2005). Stock examples of complex-systems mechanisms discussed in the literature include the mechanism for photosynthesis in a plant and the flushing mechanism of a toilet. Note, however, that complex-systems mechanisms typically used as toy examples should not be confused with complex systems as investigated in the sciences, e.g., in systems biology. Complex-systems mechanisms are complex in the sense of having multiple components, and need not be complex in the sense of being unpredictable (on account of shere scale, non-linearity, chaotic behaviour etc.). Thus the toilet is a complex-systems mechanism but arguably not a complex system.

Unfortunately there are well-known examples of causal relationships that are not accompanied by difference-making, as well as examples of causal relationships that are not accompanied by mechanisms (Hall, 2004; Williamson, 2009, 2011). In the former case, a cause may make no difference to its effect if the effect is overdetermined—i.e., if the effect would have happened anyway. In the latter case, a cause cannot be linked to its effect by a physical mechanism if the cause or effect are non-entities such as absences—but it is a commonplace to say, e.g., that a failure to take the medicine was a cause of the lack of a heartbeat. While some attempts continue to be made to circumvent such counterexamples, it is more commonly held that the counterexamples are decisive and that neither a difference-making interpretation nor a mechanistic interpretation is entirely adequate.

If none of the standard interpretations of causality are satisfactory, the question remains as to how best to interpret causal claims in medicine. In the next section an *epistemic* theory of causality is presented that, we will argue, offers a viable interpretation of such claims.

3 Epistemic causality

It is uncontroversial to say that by making causal claims in medicine we can draw a wide range of useful inferences. Our causal claims allow us to diagnose and prognose for example, as well as to make treatment decisions; they also allow us to construct explanations. To the extent that the causal claims are correct, the corresponding inferences will be successful.

Standard accounts of causality—the difference-making and mechanistic accounts of §2—interpret such causal claims as charting difference-making relationships or physical mechanisms respectively. One can then use the interpretation to explain why causal inferences are successful: in particular, if causal claims chart difference-making relationships then one can predict effect from cause and vice versa, enabling diagnosis, prognosis and treatment decisions; on the other hand, if causal claims chart physical mechanisms then one can see how they can be explanatory, for to say that E occurred because of C is to say that E occurred because there is a physical mechanism involving C that is responsible for E. Unfortunately, the standard accounts are prone to counterexamples.

An alternative, *epistemic* account of causality interprets causal claims as directly charting successful inferences (predictions and explanations). Thus our web of causal claims is used to draw the sorts of inferences alluded to above and can be thought of as a map of the inferences that it licenses. To say that C is a cause of E (in the context of a set of other granted causal claims) is just shorthand for a rather complicated list of inferences that one can make on the basis of such a claim. The explanation of why such inferences are successful when they are successful may also be rather complicated, involving a mixture of difference-making and mechanistic considerations (see Russo (2011); Russo and Williamson (2007, 2011)). Given the failure of the standard accounts, it is implausible to suggest that a single standard account can explain the success of all successful causal inferences: if it could, then that account could be used to provide an analysis of the causal relation itself.

In a sense then, mechanistic and difference-making relationships do make causal claims true, but only indirectly, in virtue of the fact that they make certain inferences successful. Moreover, they do not make causal claims true in a way that is simple enough to yield a mechanistic-cum-difference-making theory of causality: the inferential aspect is ineliminable. To say that causal claims are claims about inferences and not directly about more worldly relationships such as mechanistic and difference-making relationships, is not to say that causality is subjective—there is still a fact of the matter about what causes what. Nor, of course, does it imply that diseases (or the factors that cause disease) are unreal: a mechanism for a particular disease, along with the fact that the parts of the mechanism make a certain difference to the disease, will validate inferences about the disease and it is these inferences that constitute the causal relation, so it is the disease that helps determine the causal relation, not vice versa (on this point see also Russo (2009)).

The view that causal claims should be interpreted directly in terms of the inferences they license can be found in the writings of David Hume, Immanuel Kant, Ernst Mach and Frank Ramsey, among others. What is particular to the epistemic theory, as put forward in (Williamson, 2005, Chapter 9) and defended in Russo and Williamson (2007, 2011), is an account of the way in which we can come to learn causal relationships, i.e., an account of *causal assessment*.

According to the epistemic theory, causal claims need to be made on the basis of evidence of *both* difference-making and mechanisms, as well as evidence such as temporal information and information about the nature of the events in question.[1] (A formal framework for integrating evidence as multifarious as this can be found in Williamson (2011, Appendix A).) If one of the standard accounts were correct, then good evidence of difference-making—or good evidence of the relevant physical mechanisms—would be sufficient to establish a causal claim. But, as argued in Russo and Williamson (2007), the biomedical sciences contain compelling cases in which it is clear that neither good evidence of difference-making nor good evidence of mechanisms is enough to establish the causal claim: typically, one needs both. Indeed, as we shall see next, an IARC panel will be better off deciding about carcinogenicity by adopting the epistemic theory rather than other theories, because this theory is the only one that treats mechanistic and difference-making evidence in an egalitarian way.

4 Why the epistemic theory? To underpin causal assessment

So far we have presented a range of contenders for the meaning of causal claims: difference-making and mechanistic theories on the one hand and the epistemic theory on the other hand. The question arises as to why one should endorse an epistemic interpretation of causality in medicine. The short answer is that the epistemic theory, unlike difference-making or mechanistic theories, best fits the practice of causal assessment in medicine.

Consider, for instance, the very common situation of a general practitioner examining a patient, say Mr Jones, who is suffering from a severe flu. The GP will prescribe antibiotics to Mr Jones if she believes that the flu has bacterial complications such as pneumonia. Thus, the decision concerning the best treatment to prescribe to Mr Jones depends on what the GP comes to infer about Mr Jones' illness. Consider now a slightly different situation, where decisions and actions do not concern a particular patient but rather the whole population. In the UK, the National Health Service is performing screening programmes to test for chlamydia infections in young women. Such screening programmes are put in place because the medical community inferred that untreated chlamydia infection may lead to serious long-term health effects and even to infertility, whence the importance to control and prevent this kind of sexually transmitted infection.

These two examples show that it is of utmost importance to have sound causal assessment procedures and a conceptual framework that explains the effectiveness of those procedures. As we will argue in the remainder of this section, whilst difference-making and mechanistic theories cannot offer such a framework, the epistemic theory can.

Causal claims that form the basis of medical decisions and actions should:

(i) be based on evidence;

---

[1] We leave open the question as to whether there are other kinds of evidence for or against causal claims. Temporal evidence and evidence of the nature of the events in question can be used as evidence for or against causal claims; however, one might view these as mechanistic considerations.

(ii) be objective in the sense of non-arbitrary, that is, not the result of personal idiosyncrasies.

The first requirement doesn't need much argument: scientists and philosophers supporting other theories of causality will readily agree that causal assessment must be based on evidence. Disagreement arises as to what counts as evidence and as strong evidence. As discussed earlier, the epistemic theory takes a stance on this issue: evidence needed for causal assessment involves evidence of mechanisms and evidence of difference making.

The second requirement needs some clarification. Note that in many cases causal claims are not simply made by a single individual (e.g., the GP prescribing antibiotics to Mr Jones) but by a scientific community—think of the committees of the International Agency for the Research on Cancer (IARC) or the committees of European Commission preparing documents on regulations of health and food. It is important that the decisions and actions coming out of the deliberations of these committees be the result of intersubjective agreement and that such agreement be non-idiosyncratic. Now, interpreting causality as an inferential map avoids arbitrariness in causal assessment. The reason is that some inferences are better than others and hence some causal maps are better than others; as to which map is best is not a question of personal taste (Williamson, 2005, §9.7).

To further see why the epistemic theory best fits causal assessment in medicine, consider Bradford Hill's guidelines (Hill, 1965), which are widely accepted as a comprehensive inventory for causal assessment in medicine. Here is how Frumkin (2006, p. 1) summarises Hill's original presentation.

> 1: Strength of Association. The stronger the relationship between the independent variable and the dependent variable, the less likely it is that the relationship is due to an extraneous variable.
> 2: Temporality. It is logically necessary for a cause to precede an effect in time.
> 3: Consistency. Multiple observations, of an association, with different people under different circumstances and with different measurement instruments increase the credibility of a finding.
> 4: Theoretical Plausibility. It is easier to accept an association as causal when there is a rational and theoretical basis for such a conclusion.
> 5: Coherence. A cause-and-effect interpretation for an association is clearest when it does not conflict with what is known about the variables under study and when there are no plausible competing theories or rival hypotheses. In other words, the association must be coherent with other knowledge.
> 6: Specificity in the causes. In the ideal situation, the effect has only one cause. In other words, showing that an outcome is best predicted by one primary factor adds credibility to a causal claim.
> 7: Dose Response Relationship. There should be a direct relationship between the risk factor (i.e., the independent variable)

and people's status on the disease variable (i.e., the dependent variable).

8: Experimental Evidence. Any related research that is based on experiments will make a causal inference more plausible.

9: Analogy. Sometimes a commonly accepted phenomenon in one area can be applied to another area.

In the scientific literature there has been considerable debate concerning whether or not it is a good thing to have criteria à la Bradford Hill to assess causal relations. For instance, among the partisans of the use of criteria we find names as eminent as Susser (1977, 1991) and Weed (1997); among the detractors we find also important epidemiologists such as Rothman and Greenland (2005). Recently, Phillips and Goodman (2004) re-read Hill and vehemently argued against an interpretation of Hill's points as a checklist. In fact, Hill never intended the above as *criteria* but rather as guidelines or 'aspects' (Hill, 1965, p. 295):

Disregarding then any such problem in semantics we have this situation. Our observations reveal an association between two variables, perfectly clear-cut and beyond what we would care to attribute to the play of chance. What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?

Of course, even if the 'checklist' interpretation is rejected, it is contentious as to what the status of the nine aspects is intended to be. The question arises as to what can be left out and under what circumstances. But Hill (1965, p. 299) himself gave a clear answer:

Here then are nine different viewpoints from all of which we should study association before we cry causation. What I do not believe— and this has been suggested—is that we can usefully lay down some hard-and-fast rules of evidence that *must* be obeyed before we accept cause and effect. None of my nine viewpoints can bring indisputable evidence for or against the cause and effect hypothesis and none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question—is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect? [Emphases in the original.]

Now, Hill put forward the above-mentioned 'aspects' or 'viewpoints' for causal assessment independently of the precise interpretation of causality. But we argue that an epistemic interpretation of causality is required to underpin Hill's causal assessment.

Actually, Hill's guidelines appeal to both difference-making and mechanistic evidence. Roughly, items 2, 4, 5, 8, 9 concern evidence of mechanisms, while items 1, 3, 7, 8 concern evidence of difference-making—see also §5. As should be clear from §3, standard accounts of causality struggle to explain the need for multifarious evidence. A difference-making account of causality cannot explain why, should strong difference-making evidence be available, mechanistic

evidence remains important. Conversely, a mechanistic account cannot explain the importance of difference-making evidence when the relevant mechanisms are known. The epistemic theory, on the other hand, does not tie itself to a single kind of indicator of causality—it alone can account for the need for both sorts of evidence.

Therefore, the epistemic theory of causality provides a conceptual framework to underpin causal assessment as advised by Bradford Hill.

5 Consequences of adopting the epistemic theory. Reframing the evidence hierarchy

In the previous section we argued that the epistemic theory, unlike other rival interpretations of causality, is able to account for causal assessment in medicine, according to Bradford Hill's guidelines. While Hill's guidelines are well-entrenched, in recent years controversial 'evidence hierarchies' have been put forward as an alternative protocol for causal assessment in medicine. In this section we argue that the epistemic theory of causality provides a conceptual framework for improving the evidence hierarchy in evidence-based medicine and in evidence-based public health. When modified accordingly, the evidence hierarchy becomes compatible with Hill's guidelines for causal assessment.

The so-called 'evidence-based' movement started around the late Seventies–early Eighties with the goal of assessing the strength of evidence in order to make optimal decisions in a variety of biomedical contexts. Such decisions—ranging from deciding about a treatment for a particular patient to deciding about a public health action—are based on a *hierarchy of evidence*. Roughly, the highest form of evidence is given by randomised controlled trials (RCTs), which are thought to be more reliable than observational studies and expert opinion (see, e.g., Evidence-Based Medicine Working Group (1992); Rosenberg and Donald (1995); Sackett et al. (1996); Straus and McAlister (2000); Eddy (2005); Montori and Guyatt (2008); Djulbegovic et al. (2009)). In order of decreasing strength of evidence:

Ia Evidence from systematic reviews or meta analyses of randomised controlled trials

Ib Evidence from at least one randomised controlled trial

IIa Evidence from at least one controlled study without randomisation

IIb Evidence from at least one other type of quasi experimental study

III Evidence from non experimental descriptive studies, such as comparative studies, correlation studies and case control studies

IV Evidence from expert committee reports or opinions and/or clinical experience of respected authorities

From this hierarchy follows the well-known claim that randomisation is the gold standard of causal inference. Two remarks are immediately in order.

First, the evidence hierarchy is not in fact a hierarchy of evidence but is rather a hierarchy of the means by which evidence is *generated*: randomised

controlled trials are means for generating evidence and they are preferred to non-randomised controlled trials, which are preferred to other experiments, which are preferred to observational studies, which are preferred to opinion based on medical experience. The same or similar evidence may in fact be generated by any of these sources.

In fact, it may be that different questions in science and policy require different *modes* of evaluation of evidence or different evidence altogether, and therefore the 'received' hierarchy may change drastically—see e.g., Petticrew and Roberts (2003). A paradigmatic case of the inadequacy of the 'received' evidence hierarchy are the studies of rare diseases, for which only few individuals can be examined. Here, neither RCTs nor statistical studies (whether experimental or observational) are possible, simply because only a few cases to study exist. For this sort of case, mechanistic evidence and single-case difference-making evidence—which do not appear at all in the 'received' hierarchy—are crucial.

Second, evidence in the evidence hierarchy, as it is formulated and typically interpreted, is single-faceted: At least in I–III, the evidence under consideration is evidence of *difference making*, gathered by statistical methods. However, arguably the evidence hierarchy should handle mechanistic evidence as well—the full argument will be given below.

Those two remarks lay the ground for our main argument: since the evidence hierarchy only really includes difference-making evidence, it loses the generality of Bradford Hill's guidelines for causal assessment, which appeal to mechanistic as well as difference-making evidence. This is of course an undesirable situation and if we want to reclaim this generality, neither a difference-making nor a mechanistic theory of causality is appropriate, as each theory sees no need for the other kind of evidence. Rather, in order to reclaim the generality of Hill's guidelines we need the epistemic theory.

It is worth noting that, as Hill had already pointed out, any difference-making evidence, such as that appealed to by the evidence hierarchy, is fallible. If you know for sure that there is no possible physical mechanism linking A and B (where A and B correspond to physical entities rather than non-entities such as absences) then that is very good reason not to posit a causal claim on the basis of evidence that A makes a difference to B—even evidence generated by a RCT or a systematic review of RCTs. Such mechanistic evidence either indicates that the difference-making evidence is spurious (of course any—necessarily finite—trial can yield a sample correlation where none exists in the population, or, however well-randomised, can fail to eliminate bias with respect to confounding factors) or that A and B are not the sort of things to count as cause and effect (perhaps they do not represent distinct events, so that any correlation is due to overlap rather than causal connection).

So mechanisms are important, after all: mechanistic evidence can trump any evidence in the hierarchy. This arguably happened in the well-known and much discussed case of Semmelweis, who had good difference-making evidence for a causal claim that was rightly rejected at the time in the absence of any knowledge of the germ mechanism of disease. To give another example, those who grant that there is no mechanism that could explain patient recovery in terms of retroactive prayer should not take seriously evidence that retroactive prayer makes a difference to recovery, wherever that evidence lies in the evidence hierarchy (Leibovici, 2001). This is of course an extreme example. But

less extreme examples abound—e.g., there is moderate type Ia evidence that homeopathy makes a difference to headache (Owen and Green, 2004), but arguably our good evidence that there is no explanatory mechanism should trump this difference-making evidence and lead us to resist a causal claim.

This is all to say that, contra the evidence hierarchy, good RCTs don't always provide the best evidence for causal assessment. Simply put, RCTs aim to establish whether a given drug is effective or not comparing statistics of recovery of individuals in a control and a test group. EBM partisans usually claim that the force of RCTs exactly lies in the fact that causal generalisations thereby established do not require any description of the functioning of the active mechanism of the drug, but simply the reliable, objective, and content-neutral methods of statistics. It is this last claim that we challenge, however: RCTs are fallible and mechanistic evidence can help isolate cases of failure.

A few further remarks are in order. First, to say that mechanistic evidence can trump all the kinds of evidence of the evidence hierarchy is obviously not to say that it will always trump such evidence. On the contrary, evidence that there is no difference-making relationship between two events in the same mechanism can be reason enough to reject any claim that there is a direct causal connection between the events. Second, it is often not true that *only* statistical evidence is involved in a RCT. Mechanistic evidence may implicitly be used in setting up and in evaluating the trial (e.g., in deciding to control for mechanistically plausible confounders). However, difference-making evidence is what is *produced* by the RCT, and such difference-making evidence is only indirectly evidence that there is an underlying mechanism that explains this difference-making.

In sum, mechanistic evidence should be considered alongside difference-making evidence in EBM. This obviously complicates the picture and may make a strict *hierarchy* of evidence implausible. But, we suggest, the question of how precisely to modify the evidence hierarchy is an important research programme with so many practical consequences as to make it extremely pressing.

It should be clear by now that only under the epistemic theory is it possible, let alone plausible, to include mechanistic evidence alongside difference-making evidence in a completely egalitarian way. The reason is that the epistemic theory is the only account (i) that does not reduce the meaning of causation to a single evidential component (i.e., to mechanisms or to difference-making) and (ii) that treats these two kinds of evidence on a par in causal assessment.

We now turn to a concern that has been raised about the thesis that typically both difference-making and mechanistic evidence are required to establish causal claims. Broadbent (2010) argues that the reading of the Semmelweis case, as well as lesson one should draw from it, hinge on whether the thesis is interpreted in a normative or a descriptive way.

Here are the salient facts of the case. In 19th century Austria, the risk of puerperal fever after childbirth was very high. It was in 1833 that statistics on births and deaths due to puerperal fever were collected at the Vienna Maternity Hospital, which was divided in two clinics. It turned out that one of the two clinics had a significantly higher death rate. Ignaz Semmelweis, an assistant physician at the hospital at the time, tried to find out why that clinic had a higher death rate. Semmelweis, having considered a number of possible explanations, came up with the idea that puerperal fever was caused by contamination

of cadaverical particles. Apparently, after having carried out autopsies, doctors and medical students systematically failed to wash their hands well enough to eliminate any possible source of contamination; thus they subsequently infected their patients during examination.

The controversy arises because, in spite of the excellent difference-making evidence coming from the extensive statistics collected, Semmelweis' claim about cadaverical contamination and puerperal fever was not accepted by the medical community at that time. The reason was that his explanation—i.e., the mechanism he was invoking—was not plausible, let alone confirmed, given the medical knowledge established at that time. His hypothesis was accepted only after the germ theory of disease was developed and *Vibrio Cholerae* had been isolated by Robert Koch. Germ theory, in fact, made Semmelweis' hypothesis viable. For a thorough discussion, including important historical details, see Gillies (2005); Kundi (2006); Thagard (1998).

Let us go back to Broadbent's criticism. On the one hand, if the thesis is given a descriptive interpretation, then it is apparently true that Semmelweis failed to establish his hypothesis because of the lack of a plausible mechanism. But notwithstanding this descriptive reading, one may argue, as Broadbent does, that Semmelweis' causal hypothesis *ought* to have been considered established *without* mechanistic evidence and that the community at the time was wrong in not taking him seriously. On the other hand, if the thesis is given a normative reading, then the thesis says that Semmelweis' hypothesis was rightly rejected and, consequently, we would be rational today too in dismissing Semmelweis' hypothesis, had the germ theory not been developed. Moreover, we would be rational in dismissing the theory even in the light of the evidence coming from very effective precautionary measures such as washing hands.

We certainly agree with Broadbent that drawing lessons from historical cases is a very delicate matter indeed. Our answer to Broadbent's worry is as follows. Our argument was: in the health sciences, it is a matter of fact that mechanistic and difference-making evidence are treated in an egalitarian way when establishing causal claims (this is a descriptive claim and one that is supported by the Semmelweis case as well as many others); so, if one wants to understand causality as it is used in the health sciences, one had better appeal to an interpretation of causality, such as the epistemic interpretation, that can account for this fact (a normative claim, but a claim about how the concept of cause as it is used ought to be interpreted, not a claim about whether the usage is correct). Semmelweis' case, rather than being a test case or an 'experimentum crucis' for the suggestion that Semmelweis' contemporaries were right to reject his causal claim, simply aims to illustrate that different types of evidence are at stake in scientific practice. Indeed, because different people have different intuitions about whether the community were right to reject Semmelweis' claim, the example can hardly be decisive in that regard, and our analysis didn't engage with the historical context deeply enough to pretend to draw such conclusions from the historical data.

Having said that, we do also advocate the strong normative thesis that, when establishing causal claims, one *ought* to treat mechanistic and difference-making evidence in an egalitarian way. This strong thesis is supported by our thesis that the health sciences appeal to an egalitarian concept of cause, together with the observation that the causal claims of the health sciences have been enormously fruitful, which is compelling evidence that health scientists are going

about things the right way. It is on the basis of this strong thesis that we take Semmelweis' contemporaries' rejection of his causal claim as prima facie sensible. It is also on the basis of this strong thesis that we view as misguided the recent trend for evidence hierarchies that take mechanistic evidence to be strictly inferior to difference-making evidence.

6 Conclusion

Causal claims in the biomedical sciences are ubiquitous and yet their meaning is a matter of debate among philosophers and scientists alike. In this paper we defend the view that causality in medicine ought to be interpreted according to the *epistemic* theory. Simply put, causality is neither directly reducible to difference-making relations nor to physical mechanisms. While those relations are the evidential components that inform causal assessment, the causal relation should be interpreted as an inferential map.

A failure to clearly distinguish between the *concept* of cause and its *evidential components* results in a loss of the generality captured by Bradford Hill's guidelines. In fact, Hill's guidelines invoke evaluation of two types of evidence—difference-making and mechanistic—and an account of causality in terms of just one or the other makes it mysterious why the other type of evidence is also needed.

The epistemic theory is also able to provide a conceptual framework in which to improve the debate around the so-called 'evidence hierarchy'. In fact, the 'received' hierarchy is not really a hierarchy of evidence but rather of methods of generating evidence, in particular methods of generating difference-making evidence. One reason why the evidence hierarchy neglects mechanistic evidence is that considerations about mechanisms are more difficult to evaluate objectively than considerations arising from the mere use of statistics. This is certainly true, but, we urge, that is exactly the challenge that needs to be taken up in order to achieve a better account of evidence-based medicine.

The epistemic theory of causality is not confined to medicine—it is also applicable to many other scientific domains. Its main advantage is that it remains true to the facts of causal assessment: if some evidence needs to be taken into account in order to establish a causal claim then the epistemic theory can take that evidence into account, while for simpler reductive theories that is not always possible.

References

Arntzenius, F. (2008). Reichenbach's common cause principle. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Winter 2008 edition.

Bechtel, W. and Abrahamsen, A. (2005). Explanation: a mechanist alternative. *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 36:421–441.

Broadbent, A. (2010). Inferring causation in epidemiology: mechanisms, black boxes, and contrasts. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the sciences*. Oxford University Press.

Charlton, B. G. (1996). Attribution of causation in epidemiology: Chain or mosaic? *Journal of Clinical Epidemiology*, 49(1):105–107.

Craver, C. F. (2007). *Explaining the brain. Mechanisms and the mosaic of neuroscience*. Oxford University Press, Oxford.

Djulbegovic, B., Guyatt, G., and Ashcroft, R. (2009). Epistemologic inquiries in evidence-based medicine. *Cancer Control*, 16:158–168.

Dowe, P. (2000). *Physical causation*. Cambridge University Press, Cambridge.

Eddy, D. (2005). Evidence-based medicine: A unified approach. *Health Affairs*, 24:9–17.

Eells, E. (1991). *Probabilistic Causality*. Cambridge University Press, Cambridge.

Elwood, M. (1988). *Causal Relations in Medicine. A Practical System for Critical Appraisal*. Oxford University Press, New York.

Evidence-Based Medicine Working Group (1992). Evidence based medicine: A new approach to teaching the practice of medicine. *Journal of Americal Medical Association*, 268:2420–25.

Frumkin, H. (2006). Causation in medicine. Rollins School of Public Health, Emory University, http://www.aoec.org/CEEM/methods/emory2.html.

Gillies, D. (2005). Hempelian and Kuhnian approaches in the philosophy of medicine: the Semmelweis case. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36:159–181.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science. Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part II: Symposia Papers (Sep., 2002)*, 69(3):S342–S353.

Glennan, S. (2010). Mechanisms, causes, and the layered model of the world. *Philosophy and Phenomenological Research*.

Hall, N. (2004). Two concepts of causation. In Collins, J., Hall, N., and Paul, L., editors, *Causation and counterfactuals*, pages 225–276. MIT Press, Cambridge MA and London.

Hausman, D. (1997). Causation, agency, and independence. *Philosophy of Science*, 64(4):S15–S25. Supplement.

Hausman, D. and Woodward, J. (2004). Manipulation and the causal markov condition. *Philosophy of Science*, 71(5):846–856.

Hill, B. (1965). The environment of disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58:295–300.

Illari, P. M., Russo, F., and Williamson, J., editors (2011). *Causality in the sciences*. Oxford University Press, Oxford.

Kundi, M. (2006). Causality and the interpretation of epidemiological evidence. *Environmental Health Perspectives*, 114:969–974.

Leibovici, L. (2001). Effects of remote, retroactive intercessory prayer on outcomes in patients with bloodstream infection: randomised controlled trial. *British Medical Journal*, 323:1450–1451.

Lewis, D. (2004). *Causation and Counterfactuals*, chapter Causation as Influence, pages 75–106. MIT Press.

Lewis, D. K. (1973). Causation. In *Philosophical papers*, volume 2, pages 159–213. Oxford University Press (1986), Oxford.

Machamer, P., Darden, L., and Craver, C. (2000). Thinking about mechanisms.

Philosophy of Science, 67:1–25.

Montori, V. and Guyatt, G. (2008). Progress in evidence-based medicine. *Journal of Americal Medical Association*, 300:1814–6.

Owen, J. M. and Green, B. N. (2004). Homeopathic treatment of headaches: A systematic review of the literature. *Journal of Chiropractic Medicine*, 3:45–52.

Parascandola, M. and Weed, D. (2001). Causation in epidemiology. *Journal of Epidemiology and Community Health*, 55:905–912.

Petticrew, M. and Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, 57:527–529.

Phillips, C. V. and Goodman, K. J. (2004). The missed lessons of sir austin bradford hill. *Epidemiological Perspectives and Innovations*, 1(3).

Reichenbach, H. (1956). *The Direction of Time*. University of California Press.

Rizzi, D. A. and Pedersen, S. A. (2002). Causation in medicine: Towards a theory and terminology. *Theoretical Medicine*, 13:233–254.

Rosenberg, W. and Donald, A. (1995). Evidence based medicine: an approach to clinical problem solving. *British Medical Journal*, 310:1122–6.

Rothman, K. J. and Greenland, S. (2005). Causation and causal inference in epidemiology. *American Journal of Public Health*, 95(S1):S144–S150.

Russo, F. (2009). Variational causal claims in epidemiology. *Perspectives in Biology and Medicine*, 52(4):540–554.

Russo, F. (2011). Causal webs in epidemiology. *Paradigmi*.

Russo, F. and Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in Philosophy of Science*, 21(2):157–170.

Russo, F. and Williamson, J. (2011). Generic vs. single-case causal knowledge. the case of autopsy. *European Journal for Philosophy of Science*.

Sackett, D., Rosenberg, W., Gray, J., Muir, H., Brian, R., and Richardson, W. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal*, 312:71–2.

Salmon, W. C. (1998). *Causality and explanation*. Oxford University Press, Oxford.

Straus, S. and McAlister, F. (2000). Evidence-based medicine: A commentary on common criticisms. *Canadian Medical Association Journal*, 163:837–841.

Suppes, P. (1970). *A probabilistic theory of causality*. North Holland Publishing Company, Amsterdam.

Susser, M. (1977). Judgment and causal inference: criteria in epidemiologic studies. *American Journal of Epidemiology*, 105(1):1–15.

Susser, M. (1991). What is a cause and how do we know one? a grammar for pragmatic epidemiology. *American Journal of Epidemiology*, 133:635–647.

Thagard, P. (1998). Explaining disease: correlations, causes, and mechanisms. *Minds and Machines*, 8:61–78.

Weed, D. L. (1997). On the use of causal criteria. *International Journal of Epidemiology*, 26:1137–1141.

Williamson, J. (2005). *Bayesian nets and causality: philosophical and computational foundations*. Oxford University Press, Oxford.

Williamson, J. (2009). Probabilistic theories. In Beebee, H., Hitchcock, C., and Menzies, P., editors, *The Oxford Handbook of Causation*, pages 185–212. Oxford University Press, Oxford.

Williamson, J. (2011). Mechanistic theories of causality. *Philosophy Compass*, to appear.

Woodward, J. (2003). *Making things happen: a theory of causal explanation.* Oxford University Press, Oxford.

Woodward, J. and Hitchcock, C. (2003). Explanatory generalizations, part I: A counterfactual account. *Noûs*, 37(1):1–24.