THE ROUTLEDGE HANDBOOK OF CAUSALITY AND CAUSAL METHODS

Edited by Phyllis Illari and Federica Russo

First published 2025

ISBN: 978-1-032-26019-8 (hbk)
ISBN: 978-1-032-26287-1 (pbk)
ISBN: 978-1-003-52893-7 (ebk)

Chapter 36

WHEN DECISIONS MUST BE BASED ON PARTIAL CAUSAL KNOWLEDGE: ANALYZING CAUSALITY AND EVIDENCE FOR HEALTH POLICY

Fredrik Andersen, Rani Lill Anjum, and Elena Rocca

DOI: 10.4324/9781003528937-56

The funder of the Open Access of this chapter is Østfold University colleges publication fund.

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

36

WHEN DECISIONS MUST BE BASED ON PARTIAL CAUSAL KNOWLEDGE

Analyzing Causality and Evidence for Health Policy

Fredrik Andersen, Rani Lill Anjum, and Elena Rocca

Key messages:

- Public health decisions are often based on incomplete causal evidence. When evidence is incomplete, but decisions must be made, basic implicit assumptions play an important role.
- The role of basic implicit assumptions differs depending on the type of uncertainty involved. Identifying, analyzing, and reevaluating basic implicit assumptions is particularly important when making decisions based on unavailable or diverging evidence.
- To make decisions when available causal evidence diverges, evidential pyramids are usually invoked. However, these pyramids have their own inbuilt basic implicit assumption about causality that needs to be debated.
- Debating assumptions about causality increases conceptual awareness and avoids reliance on a limited set of causal theories. A wider set of causal theories is helpful in cases of health emergencies when innovative thinking is needed.

Key readings:

- Andersen, F., Anjum, R. L., & Rocca, E. (2019) "Philosophical Bias Is the One Bias That Science Cannot Avoid," *ELife* 8: e44929. https://doi.org/10.7554/eLife.44929
- Bradford Hill, A. (1965) "The Environment and Disease: Association or Causation?" *Proceeding of the Royal Society of Medicine*, *58*.
- Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996) "Evidence-Based Medicine: What It Is and What It Isn't," *British Medical Journal*, 312: 71.
- Amoretti, M. C., & Lalumera, E. (2021) "COVID-19 As The Underlying Cause Of Death: Disentangling Facts And Values," *History and Philosophy of the Life Sciences*, 43(1).

DOI: 10.4324/9781003528937-56

> All scientific work is incomplete – whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. This does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time.
>
> *(Bradford Hill 1965)*

36.1 Causality as a philosophical bias in evidence and evidence-based decision making

Important decisions ought to be based on our best knowledge, even if the knowledge is incomplete. This principle is used also beyond clinical decision making and public health policy. However, what exactly counts as the best knowledge is not always a clear-cut matter, especially when it comes to causal knowledge. For instance, what type of evidence is most relevant for making treatment choices for patients? Are results from large population studies relevant for the unique individual, or are they only applicable on the population level? And what about the question of adverse effects? Can health authorities conclude from only a few clinical instances that an undesired symptom is a side effect of the treatment? What about otherwise safe and effective treatments where there is a risk of rare but serious side effects? Should knowledge about those risks be generated mainly statistically or based also on knowledge about causal mechanisms and individual drug-patient interactions? What, then, about moral concerns? Is it, for instance, acceptable to risk that some few people are harmed by an intervention that would benefit most people, if we cannot predict beforehand who might be harmed? To which extent should population-level concerns limit an individual's access to certain treatments? Health decisions are already riddled with moral dilemmas and weighing different information and priorities. Must one also worry about causality? Absolutely, and here is why.

There are many ways to think about causality. One might think of causality in terms of perfect regularity (Beebee 2007; Psillos 2002), natural necessity (Armstrong 1983), counterfactual dependence (Collins et al. 2004; Lewis 1973b, 1973a), physical processes (Dowe 2000; Kistler 2006; Salmon 1984), manipulability (Woodward 2003), INUS conditions (Mackie 1974), intrinsic disposition (Cartwright 1989; Mumford & Anjum 2011), probability raisers (Mellor 2002; Suppes 1970), mechanisms (Glennan 2009), or a combination of these (Cartwright 2007; Hall 1994; Illari & Russo 2014; Russo & Williamson 2007). For a discussion about several of these accounts, see the initial session of this volume (Maziarz, this volume; Suarez, this volume).

Different concepts of causality motivate different methods for empirical detection, and different experts will look for different evidence to support causal claims (see also Sözüdoğru, this volume). For instance, statistical methods are better at detecting correlations and probability raisers than they are at detecting mechanisms or intrinsic dispositions. Controlled experiments can help us detect necessary conditions and difference-makers, but perhaps not suffice to detect intrinsic dispositions. Experts might use many methods, but still rank some types of evidence as stronger or weaker regarding causal conclusions, as in the classic evidence pyramid of evidence-based medicine (EBM) (Sackett et al. 1996). One might, for instance, think that causality should manifest itself in an observable difference-maker, where the effect is statistically significant when compared to a control. If so, randomized controlled trials could be the best method for establishing causality. Perhaps, though, one thinks that the ideal is to demonstrate a necessary connection between cause and effect, supported by a theory of the mechanism underlying such connection. In that case, it seems that the statistical

difference-maker could be the first step on the ladder to causal insight, while the theory of mechanism is the highest step where causality is finally established.

The way one evaluates and ranks causal evidence typically comes with science education, as part of the methods training. In 'Causal thinking in science education and the challenges it holds' (this volume), Haskel-Ittah describes how causal thinking is shaped during science education in some specific ways that range from pattern seeking to mechanistic reasoning and highlights the issues linked to each of these education strategies (see also Bakker, this volume). There is, we might say, a philosophical bias in the methods for studying and establishing causality; about what causality is (ontology) and how to gain knowledge about causality (epistemology). We have defined a philosophical bias as a *b*asic *i*mplicit *a*ssumption in science concerning ontology, epistemology, or norms (Andersen et al. 2019). Philosophical biases are *basic* in the sense that they lay the ground for methodological and theoretical choices, and they are *biases* because they *implicitly* '… skew the development of hypotheses, the design of experiments, the evaluation of evidence, and the interpretation of results in specific directions' (Andersen et al. 2019). Philosophical biases, we argue, cannot be avoided. We cannot, for instance, choose to get rid of any assumption about the nature of causality. When picking out a method for studying or establishing causality, we have also adopted a philosophical concept of causality: that which the method is designed to detect. For example, as Holly Andersen points out (this volume), the choice of a particular statistical model implies the often-unaware commitment to a certain metaphysical view of the system that is modeled. Unlike other biases in science, therefore, a philosophical bias is not something that one should strive to get rid of or avoid completely. Instead, we should be made aware of our basic implicit assumptions, to make them explicit and open for critical reflection. At the same time, we should learn about the alternative assumptions, so we can compare them and make an informed choice.

For instance, one might have been trained in a research tradition where only quantitative methods are accepted for generating causal knowledge. This, however, is based on certain assumptions about the nature of causality, typically inspired by Hume's strict empiricist analysis where any theoretical and qualitative elements of causality are dismissed. From the empiricist starting point, only features that are in principle observable could be included in the scientific inquiry, such as regularities, probability raisers, difference-makers, or necessary conditions. If, on the other hand, one aims to establish intrinsic dispositions and causal interactions of these, then knowledge about theoretical and qualitative elements is indispensable for causal evidence (Douglas 2000, 2012; Rocca & Anjum 2020). We see, then, that our philosophical biases about causality and what we see as 'best' causal knowledge will reveal itself in the choice and ranking of research methods.

36.2 Three clarificatory remarks about philosophical biases in science

Before we move on to discussing the case of partial causal knowledge, we need to make three brief clarifications about biases. First, a philosophical bias can be made explicit, in which case it would no longer count as a bias. One might, for instance, as a methodological choice, use a deterministic model, while remaining fully aware of the limitations that this entails. Perhaps the model can be used to predict something, but not everything, so is not used to make predictions for indeterministic contexts. In that case, determinism is not a philosophical bias but simply adopted as a heuristic tool. One could also be explicit about having an empiricist bias, where causality must be generated from data and comparisons of these, without the aim to develop causal theories. Any mechanism must then be understood as describing observable features, such as correlation data, difference-makers, or

probability raisers, rather than relying on theoretical speculations. As long as these basic assumptions are made explicit and transparent, they are no longer implicit or *biases*, because their influence on causal evidence and interpretations of it reflects an informed choice of the researcher. It also enables the stakeholders of the research, including decision makers, to consider the evidence in light of those assumptions. Such interrelation between evidence and assumptions has a particular relevance when decisions are based on partial causal knowledge, as we will see in what follows.

Second, the idea of a philosophical bias is not new in philosophy, although that terminology is not used. Traditionally, in cases where there is no evidential agreement on a single solution, philosophers tend to focus on ethics and how normative arguments are intrinsic to scientific practice. For instance, Putnam (2002) attacked the fact/value distinction, Hempel (1965) discussed how judgments of value are presupposed in scientific thinking, and Douglas (2000) has argued that the goal of a value-free science should be replaced by an acceptance of the fact that in some cases, non-epistemic values are necessary parts of scientific reasoning. A common theme in this philosophical tradition, which is particularly relevant for our upcoming discussion, is that when we must make decisions in the face of diverging or incomplete evidence, we must debate the relevant epistemic and non-epistemic values. We support these arguments, but at the same time we want to emphasize the importance of explicating ontological and epistemological philosophical biases in addition to the ethical assumptions (Andersen et al. 2019; Longino 2002: 176–77).

Third, while we think explication of philosophical biases is helpful in controversies, we would also stress that there are possible ways to rationally discuss and compare these biases. In other words, for the purpose of decision making, we do not have to simply accept that we make basic assumptions. We can also judge which of these are more reasonable. When faced with diverging evidence and diverging interpretations of evidence, we can evaluate the weight of evidence according to external coherence. We can look at trends in the overall field of study and see which assumptions are operative. Evidence derived from methods that cohere better with the assumptions of the overall field should arguably be given more weight (for details, see Andersen & Rocca 2020; Rocca & Andersen 2017).

36.3 Philosophical bias about causality in decisions based on partial causal knowledge

We now go on to analyze some scenarios for when regulatory decisions in public health are based on partial causal knowledge. We will talk about the case in which both experts and decision makers share the same view on the available evidence and thus the same knowledge gaps. In such cases, there is no clear-cut way for a decision maker to rely on the 'best available science'. Rather, the decision makers must rely on partial causal knowledge only. Causal knowledge can be partial or incomplete in different ways. As already discussed, there might be various forms of evidence available, but experts might disagree over which should count as causal evidence. This is a problem for any domain of decision making, especially when the scientific evidence points toward opposite causal conclusions. If, in contrast, all the evidence points in the same direction, and there are not ethical concerns for acting on the evidence, then the decision seems clear. Sometimes, though, there isn't much causal evidence at all, and experts might disagree over which types of evidence to prioritize when decisions are urgent. In what follows, we discuss some ways in which causal knowledge can be partial: small evidential gap (scenario 1), wide evidential gap (scenario 2), contradictory evidence (scenario 3), or the impossibility to access specific types of evidence because of health emergencies (scenario

4). For each scenario, we look at the interplay between empirical evidence, philosophical bias about causality, and ethical considerations.

36.3.1 Scenario 1: Small evidential gap and robustness of philosophical bias

We will start by considering the scenario when there are different types of available causal evidence that point to the same causal conclusion, which is nevertheless uncertain to some extent. For instance, uncertainty could come from the fact that some of the available studies are statistically underpowered. The phrase 'different types of causal evidence' has many possible meanings. It could mean that there is evidence from *different methods*, such as evidence from an observational study and from a clinical trial, or evidence from case studies combined with statistical correlations in registry studies. More relevant for our context is when there is evidence for *different causal features*, for instance if there is some evidence of a difference-maker combined with some evidence of a plausible mechanism. Following the pluralist conception of causality of Bradford Hill (1965), these features are reliable symptoms of causality.

A situation of this type can for example occur in the early, pre-marketing phase of drug development. Say for instance that a new drug has been shown in pharmacological in vitro experiments to block biological receptors that relax bladder muscles. This could constitute evidence of a causal mechanism. Imagine also that in pre-marketing clinical trials, more participants in the experimental group report the adverse effect 'urine retention' than in the control group. This could constitute a different type of causal evidence, namely evidence of the drug being a difference-maker for the effect. Both types of evidence – the mechanism and the difference-maker – support the causal hypothesis that the drug provokes stiffness of bladder muscles and therefore difficulty in evacuating the bladder.

In this type of scenario, the causal knowledge is robust, although there is some margin of uncertainty if the quality of the studies or observations is poor. If we consider causality specifically, one might call this a situation of 'philosophical bias robustness', since results from methods supported by different epistemological and ontological biases about causality converge to support the same conclusion. The need to discuss causality bias for evidence evaluation is thus comparatively smaller in this type of scenario related to the ones presented below. This is of course not to say that such a philosophical discussion would be useless. One can always use reflections over basic assumptions to improve causal thinking and clarify what type of evidence would be needed to complete the causal picture. However, we think that in this scenario the evidential gap might be bridged without a particular focus on the epistemology and ontology of causality. Purely technical and methodological discussions about the quality of the available studies might do the job. The extra-evidential discussion could instead focus on ethical assumptions: given the available causal knowledge, what action is best? What are the significance and the relevance of the causal evidence for our specific context of interest? Are the risks involved acceptable? And so on.

36.3.2 Scenario 2: Wide evidential gap and extreme philosophical bias openness

We will now move on to describing the opposite scenario, in which only limited, isolated causal evidence is available. This situation might be rarer than the first scenario. In general, evidence needs to accumulate at least to some extent before becoming visible to decision

makers, but this is not always the case (Meyboom et al. 1997). The threshold for regulators' attention is context-dependent and tends to be lower when stakes are high (Funtowicz & Ravetz 1993). In the start of a potential pandemic, for instance, regulatory evaluations are needed even if there is very limited evidence about how fast and how much the infection might spread, because the stakes are so high. In such situations, health authorities will worry about small changes that can make a big difference if not regulated (Briand et al. 2011). It is not uncommon that patients and communities are isolated before extensive evidence about an ongoing infection is gathered. Another reason for acting on an observation despite high uncertainty is that it might be particularly hazardous, new, and unexpected, and as such it might cast doubt on an otherwise dominant theory and policy (Rocca et al. 2019). A historical example of this is the observation of a few serious cases of rare fetal malformations in women who had used the drug thalidomide during pregnancy. This observation, of an unexpected increase in babies born with missing limbs, was sufficient not only to suspend the drug from the market but also to trigger further investigations as well as a series of institutional changes in the whole process of drug safety monitoring (Dally 1998). Sometimes the early evaluation of evidence does not result in immediate action, but rather in the decision to wait for further evidence. In these cases, waiting is itself a decision, and like other decisions it implies the interplay of evidence evaluation together with specific extra-evidential assumptions.

We see, then, that scenarios can arise in which regulatory decisions need to bridge a wide evidential gap. What is the role of philosophical bias in these situations, and particularly what is the role of causality bias? In line with Douglas, Longino, and others, we maintain here that the reliance on extra-evidential assumptions grows proportionally with the uncertainty of the evidence (Douglas 2000, 2012; Flage et al. 2014; Hartley et al. 2016; Longino 1990). When the facts under evaluation are few, isolated, or even idiosyncratic, there are typically many scientifically justifiable explanations of why the evidence might look the way it does. Since the consequence of adopting one over the other explanation has an impact on the overall evaluation, and therefore on decision making, the choice must imply extra-scientific considerations, Douglas argues (Douglas 2012). As already mentioned, many of the arguments in the literature primarily focus on extra-evidential considerations of an ethical type, and they require an ethical or social commitment of scientists as well as a transparent and democratic governance (Douglas 2000; Wickson & Wynne 2012). However, we see no reason to restrict the requirement of extra-evidential considerations to the ethical realm. The choice of the best scientific explanation that can accommodate thin evidence requires extensive methodological reflections, hence, reflections also on causal biases (see Part II and Andersen & Rocca 2020). For instance, if evidence of an unexpected, rare correlation arises, the need for evidence for a plausible mechanism must be discussed.

A long list of real-life examples could be quoted here from pharmacovigilance, the field of drug safety monitoring. The use of certain medicines is sometimes correlated with serious, lethal conditions in extremely rare, idiosyncratic cases. Any evaluation of a causal hypothesis from expert committees requires conceptual discussions about the evidential value of single cases (Aronson & Hauben 2006), the relevance of different types of causal evidence (Shakir & Layton 2002), the relevance of a certain piece of evidence for different evaluating committees (Dinis-Oliveira & Vieira 2021), causal analogy, and the value of established knowledge (Gérard et al. 2021), to mention some.

When decisions need to bridge a wide evidential gap, the discussions typically tend to become open and explicit about conceptual assumptions. These situations will not only require a consideration of ethical concepts and priorities but also an explicit discussion about

methodological choices and biases about causality. We should not mainly be looking at implicit philosophical bias, then, but at explicit assumptions that could help to support one or the other argument in the face of insufficient evidence.

36.3.3 Scenario 3: Accumulated conflicting evidence and institutionalized philosophical bias

Sometimes decision making happens after different types of evidence have accumulated, but such evidence is conflicting. In those cases, the causal knowledge is partial even if the body of evidence is extensive. Uncertainty derives from the fact that different types of evidence point to diverging conclusions. For example, statins are widely used cholesterol-lowering medicines whose benefit-risk profile has been under debate (Abramson et al. 2013). Besides the rare and serious adverse events, there is the concern that statins might provoke also milder but debilitating muscle symptoms, such as pain and weakness, in a relatively large number of patients. Real-life evidence, such as routine health records and observational studies, have suggested that statins strongly increase the risks of musculoskeletal disorders (Bruckert et al. 2005; Finegold et al. 2014). On the other hand, systematic reviews of large RCTs have not shown an increased rate of musculoskeletal symptoms reported among study participants taking statins versus participants taking placebo (Blazing et al. 2022; Finegold et al. 2014). The causal interpretation of this body of evidence is clearly important to inform clinical action. Should patients who start to use statins be warned of an increased risk of muscle pain? Or should they be told that such symptoms, if experienced, are probably caused by other factors than the statins therapy?

Similar cases have been used to show how the same set of conflicting evidence is evaluated by experts with divergent philosophical biases (Andersen & Rocca 2020; Guttinger 2020; Rocca & Andersen 2017; Rocca & Anjum 2019). In the context of medicine and public health, it is common to evaluate complex bodies of evidence, but at the same time there is a wish for evidence-based actions to be homogeneous among evaluators and not dependent on assumptions of the single practitioners. Because of this, evidence hierarchies are used to guide evidence evaluation (Sackett et al. 1997). When conflicting evidence emerges, then, evaluators and decision makers typically grade the evidence according to the relevant evidence hierarchy, to provide evidence-based guidelines. Evidence that is graded higher is considered more reliable for establishing cause-effect relationships and should therefore be given more weight. For a historical overview of how public health guidelines in the UK have used evidence, see Kelly (this volume).

A crucial point is that hierarchies of evidence come with their own biases about causality, placing one concept on the top of the pyramid. Usually, this is difference-making combined with regularity. Methods at the top of the pyramid of evidence, which are said to be best suited to establish causality, would then be those that are best to establish difference-making at group level. Indeed, at the top of the EBM evidence hierarchy are controlled clinical experiments, followed by controlled observation, while case studies and other types of non-controlled observations are usually at the bottom. Consequently, when evaluators adopt the EBM hierarchy, they also adopt the philosophical bias that a cause is something that makes a difference at group level (ontological bias), or that the best way to detect a cause is to look for difference-making at group level (epistemological bias), or both. In this way, evidence hierarchies are *institutionalized philosophical biases* that determine some generally applicable criteria for evidence evaluation. Although there are several versions of the evidence pyramid, none to our knowledge adopts a competing view on causality (Murad et al. 2016). Alternative

rankings of evidence have increased focus on internal validity and place N of 1 clinical trials (where difference-making is observed in single patients) above the conventional RCTs (Howick 2011). Other rankings challenge the privileged positioning and role of systematic reviews of RCT (Murad et al. 2016), but without modifying the implicit assumption that controlled studies are more reliable for establishing causality.

Returning to the example of statins and musculoskeletal symptoms, the EBM evidence pyramid would rank results from systematic reviews of RCTs, which suggest no causal connection, as the most reliable. This was also reflected by the authors of the review, who suggested that musculoskeletal symptoms should be removed from the list of adverse effects in the medicine label, based on their results (Blazing et al. 2022). This got considerable media amplification, suggesting that muscle pain in statin users is 'imaginary' (Graedon 2023). For an example from nutrition science, see Jukola (this volume).

There is a conceptual downside of adopting pre-established evidence guidelines for evaluating complex evidence. When institutionalized philosophical biases are tacitly accepted in the mainstream scientific community, the need for critical discussion among experts and stakeholders decreases. As discussed in Part II, there is a large variety in the ways causality can be conceptualized, and each different assumption results in different methodological reflections. Nevertheless, within the current EBM paradigm, evidence of causality is consistently taken to mean evidence of difference-making, a bias that remains undiscussed and ultimately unscrutinized by the mainstream community (Clarke et al. 2013). As a result, evidence for causality tends to revolve around the methods that are best for detecting difference-making. For instance, Ilardo and Reiss (this volume) analyze how difference-making accounts are at the foundation of the WHO guidelines for vaccine safety. Such an almost exclusive 'investment' in the difference-making bias of causality can become a disadvantage for preparedness and resilience in case of emergencies, as the next scenario suggests.

36.3.4 Scenario 4: Inaccessible evidence and the need for conceptual preparedness

In this last scenario, we look at the situation in which decisions must be made although the established standards of evidence necessary to inform science-based policy are difficult or even impossible to obtain. Examples are natural disasters, pandemics, or other 'emergency' scenarios. These are situations that cannot be handled by established structures and which require a modification and adaptation of measures that are otherwise common. The COVID-19 pandemic, which spread globally within a few months, was a health emergency that posed some considerable challenges for evidence-based policy. Social and medical measures were initiated without the backing of clinical experiments or appropriately designed observational studies. Here, the role of real-time and real-world evidence became key, for instance, to understand how the virus spreads, how deadly it is, the pathophysiology of the infection, and the effectiveness and safety of possible therapies. In addition, knowledge had to progress exceptionally fast, and the scientific community had to be responsive to observations made outside experimental designs that are traditionally acknowledged as most reliable. The failure of traditional standards of reference increased public uncertainty and fueled controversies. For example, Ioannidis claimed that social distancing as a measure against the spread of the infection risked becoming a 'once-in-a-century evidence fiasco' because it was not supported by good evidence or reliable data (Ioannidis 2020). Notice, however, that here the understanding of 'reliable data' and 'good evidence' has one implicit but precise meaning among many possible meanings. Ioannidis implicitly emphasizes EBM's institutionalized epistemological bias

of internal validity and experimental control (Jukola & Canali 2021), meaning that he was referring simply to the lack of evidence from randomized controlled trials.

The exclusive reliance on the difference-making concept of causality, however, falls short in this type of scenario. To illustrate this, we can look at the debates surrounding COVID-19 death counts. This is crucial information for both policy and public perception. However, counting the number of COVID-19 victims turned out to be difficult when the suggested difference-making approach could not deliver reliable results. The method was to compare total deaths in a population before and during COVID-19. This could give a rough idea of how deadly the virus was. There are two core issues with this approach. First, it requires a longer time-period to calculate meaningful averages. This is of course not acceptable since we must then wait for high levels of mortality before we decide how dangerous the virus is. The second issue is that the numbers of deaths before and during COVID-19 are not necessarily comparable. As soon as physical distancing and similar measures were implemented, society and the underlying conditions that might generate death rate changed. Granted, it is possible to control for some of these issues by statistical methods, but changes as radical as we witnessed during the COVID-19 pandemic cannot all be accounted for statistically. Without comparable situations, we cannot apply the difference-making concept. For instance, one would have to conclude that COVID-19 was deadly in England but health promoting in Norway, which had fewer than average deaths in 2020 (Sønstebø 2021).

One alternative was to rely on the WHO fine-grained method of determining the cause of death of a person. Importantly, the WHO system requires that one registers a single cause of death, which is defined as the one cause that starts a chain of events. In order to be registered as a COVID-19 death, it is not sufficient to have the infection, but the fatal symptoms must be relatable through a causal chain back to COVID-19. Additional contributors or underlying pathologies are instead registered as comorbidities. However, it is not always clear how to distinguish between underlying pathologies and the cause of death. If a terminally ill patient with a weak immune-system that is expected to die soon gets COVID-19 and dies, is the single cause of death in that case COVID or the terminal illness? We are in a case of co-determination, which in philosophy is called the problem of *overdetermination*. If more than one factor is sufficient to produce the effect, then the effect counterfactually depends on none. It's not true that 'if it hadn't been for COVID-19, the person wouldn't have died', nor is it true that 'if it hadn't been for the terminal illness, the person wouldn't have died'. Therefore, one cannot establish that the person died because of either.

The WHO's pragmatic approach to this issue is that in cases of co-determination, death should always be attributed to COVID-19 (World Health Organization 2020). This approach could be motivated by a philosophical bias toward the manipulability concept of causality (Amoretti & Lalumera 2021). This theory states, very roughly, that the cause of an effect is the element that could be manipulated to change (or *make a difference to*) the effect. For example, if I slipper on the icy street and fall, this theory will emphasize the ice over gravity as the cause for the fall, since the first could be removed or prevented, but not the latter. The WHO explicitly states that they use this approach to improve possible public health policies: 'whether a causal relationship is considered acceptable for mortality coding is founded not only on a medical assessment but also in epidemiological and public health considerations' (World Health Organization 2020). It is, of course, also an admittance that it is often almost impossible to determine a single cause of death and that choices must be made based on the preferred socio-political value (Amoretti & Lalumera 2021).

The WHO's choice of applying a manipulability theory, although not explicitly, is rationally valid. Moreover, prioritizing public health is the WHO mandate. Finally, the merging of

scientific evidence and values is supported by a long and increasing trend in philosophy of science. It is therefore rationally defensible as 'appropriate science'. Nevertheless, this strategy was also the source of much controversy, especially from politicians and the public (Aschwanden 2020). Why? One answer, which is the point we are promoting here, is that because of institutionalized philosophical bias about evidence evaluation, scientists, regulators, and laypeople think about causality in terms of difference-making and counterfactuals. When one finds that the WHO applies a different theory, a natural response is to claim that they are not establishing causality. The central problem here, in our view, is not that the WHO applies a manipulability theory nor that the public's expectation is a different one. The problem is rather that the theories remain implicitly assumed from both parts. The WHO describes their procedure but does not specify that this is a manipulability theory of causality (Amoretti & Lalumera 2021). This can be problematic when, because of common practice, by the term 'causality' many stakeholders of science mean 'difference-maker'.

If basic assumptions about causality were made explicit, there would be no actual controversy concerning whether we are correctly determining the cause of death or not, at least in this case. Instead, we could have an open and reflected debate over which causal concept is more appropriate for our purpose. We could also discuss the limitations and advantages of different methods for establishing causality under that definition. For instance, one could discuss what makes something an underlying cause versus a trigger for the effect, and whether it makes sense to operate with single causes in this case. Greenhalg, Engebretsen and Sandset (this volume) describe further debates within the COVID-19 pandemic that would have benefited from a transparent analysis and communication of basic assumptions about causality, for instance the debate about the cause of the pandemic, or the evaluations of evidence about the way the virus spreads. When the debate is shaped the way we envision here, there is less room for confusion about what the causal evidence amounts to, and there would be little fuel for conspiracy theories. This vision is in line with other proposals that promote pluralistic strategies to causality, in particular with the 'causal mosaic approach' developed by Illari and Russo, which emphasizes the selection of an appropriate notion for appropriate contexts (Illari & Russo 2014 and general introduction to this volume).

36.4 Promoting conceptual awareness and diversity of causal evidence

We have seen that philosophical biases are unavoidable in scientific methods and practice and that conceptual assumptions become increasingly important for decision makers as the stakes increase and there is evidential divergence or uncertainty. In order to maintain science-based decision making of the highest quality, it is vital to detect, analyze, and openly discuss philosophical biases. However, bias relating to causality is particularly important because this concept is intrinsically connected to our choice of methods, and how we evaluate the evidence generated by these.

If we want to prepare ourselves for future high stakes events, both at individual and population level, we could look to the mechanisms behind evolutionary survival. Evolutionary biology teaches us that whenever nature poses a specific challenge for species survival, diversity provides a clear benefit. This is because high levels of diversity increase the probability of survivors, where species uniformity would risk losing the whole population. The same logic can be applied to causal methods and their underlying *conceptual* assumptions in the context of evidence-based decision making. As we have seen, different methods are applicable to different causal concepts. Furthermore, different methods are suitable at detecting causal relations in different situations. Since we cannot predict the future, we cannot predict

which method will be most suitable to future situations. Therefore, our best contingency plan is to maintain a large diversity of methods at all times. If we are trained to recognize our own assumptions about causality, and at the same time acknowledge that there is a plurality of other possible assumptions, we increase resilience for situations where causal knowledge is lacking, diverging, or uncertain. We also increase resilience for situations in which a more radical renewal of foundational thinking is required. However, if we adopt a strict hierarchy of methods and train new scientists to adhere to this hierarchy without conceptual or critical reflection, we risk narrowing our scope too much and miss out on causally valuable knowledge. In such a scenario, we risk educating a whole generation of scientists and decision makers, specialized at detecting one specific aspect of causality, for example difference-making and regularities, while being unprepared to think in terms of dispositions, mechanism, manipulability, or probability raising, when difference-makers or regularities cannot be applied. Methodological diversity should be the preferable pragmatic choice and explicit reflection about the concept of causality ought to be promoted among scientists, decision makers, and other stakeholders.

References

Abramson, J. D., Rosenberg, H. G., Jewell, N. & Wright, J. M. (2013) "Should People at Low Risk of Cardiovascular Disease Take a Statin?" *BMJ (Online)*, 347: 6123. https://10.1136/bmj.f6123

Amoretti, M. C. & Lalumera, E. (2021) "COVID-19 As the Underlying Cause of Death: Disentangling Facts and Values," *History and Philosophy of the Life Sciences*, 43(1): 4.

Andersen, F., Anjum, R. L. & Rocca, E. (2019) "Philosophical Bias Is the One Bias That Science Cannot Avoid," *ELife*, 8: e44929. https://doi.org/10.7554/eLife.44929

Andersen, F. & Rocca, E. (2020) "Underdetermination and Evidence-Based Policy," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 84: 101335. https://10.1016/j.shpsc.2020.101335

Armstrong, D. M. (1983) *What Is a Law of Nature?* Cambridge: Cambridge University Press.

Aronson, J. K. & Hauben, M. (2006) "Anecdotes That Provide Definitive Evidence," *British Medical Journal*, 333(7581): 1267–1269.

Aschwanden, C. (2020) "Debunking the False Claim That COVID Death Counts Are Inflated," *Scientific American*. https://www.scientificamerican.com/article/debunking-the-false-claim-that-covid-death-counts-are-inflated1/

Beebee, H. (2007) "Hume on Causation: the Projectivist Interpretation" in H. Price & R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press, 224–249.

Blazing, M., Braunwald, E., de Lemos, J., Murphy, S., Pedersen, T., Pfeffer, M., White, H., Wiviott, S., Clearfield, M., Downs, J. R., Gotto, A., Weis, S., Fellström, B., Holdaas, H., Jardine, A., Gordon, D., Davis, B., Furberg, C., Grimm, R., … Byington, R. (2022) "Effect of Statin Therapy on Muscle Symptoms: An Individual Participant Data Meta-Analysis of Large-Scale, Randomised, Double-Blind Trials," *The Lancet*, 400(10355): 832–845.

Bradford Hill, A. (1965) "The Environment and Disease: Association or Causation?" *Proceeding of the Royal Society of Medicine*, 58(5): 295–300.

Briand, S., Mounts, A. & Chamberland, M. (2011) "Challenges of Global Surveillance during an Influenza Pandemic," *Public Health*, 125(5): 247–256.

Bruckert, E., Hayem, G., Dejager, S., Yau, C. & Bégaud, B. (2005) "Mild to Moderate Muscular Symptoms with High-Dosage Statin Therapy in Hyperlipidemic Patients – The PRIMO Study," *Cardiovascular Drugs and Therapy*, 19(6): 403–414.

Cartwright, N. (1989) *Nature's Capacities and Their Measurements*. Oxford: Oxford University Press.

Cartwright, N. (2007) *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press.

Clarke B, Gillies D, Illari P, Russo F, Williamson J. (2013) "The Evidence That Evidence-Based Medicine Omits," *Preventive Medicine*, 57(6): 745–747.

Collins, J. D., Hall, E. J. & Paul, L. A. (Eds.). (2004) *Causation and Counterfactuals*. Cambridge, MA: MIT Press.
Dally, A. (1998) "Thalidomide : Was the Tragedy Preventable?" *Lancet*, 351(9110): 1197–1199.
Dinis-Oliveira, R. J. & Vieira, D. N. (2021) "Acute Liver Failure Requiring Transplantation: A Possible Link to Ulipristal Acetate Treatment?" *Basic and Clinical Pharmacology and Toxicology*, 129(3): 278–282.
Douglas, H. (2000) "Inductive Risk and Values in Science," *Philosophy of Science*, 67(4): 559–579.
Douglas, H. (2012) "Weighing Complex Evidence in a Democratic Society," *Kennedy Institute of Ethics Journal*, 22(2): 139–162.
Dowe, P. (2000) *Physical Causation*. Cambridge: Cambridge University Press.
Finegold, J. A., Manisty, C. H., Goldacre, B., Barron, A. J. & Francis, D. P. (2014) "What Proportion of Symptomatic Side Effects in Patients Taking Statins Are Genuinely Caused by the Drug? Systematic Review of Randomized Placebo-Controlled Trials to Aid Individual Patient Choice," *European Journal of Preventive Cardiology*, 21(4): 464–474.
Flage, R., Aven, T., Zio, E. & Baraldi, P. (2014) "Concerns, Challenges, and Directions of Development for the Issue of Representing Uncertainty in Risk Assessment," *Risk Analysis*, 34(7): 1196–1207.
Funtowicz, S. O. & Ravetz, J. R. (1993) "Science for the Post-Normal Age," *Futures*, 25(7): 739–755.
Gérard, A. O., Laurain, A., Fresse, A., Parassol, N., Muzzone, M., Rocher, F., Esnault, V. L. M. & Drici, M. D. (2021) "Remdesivir and Acute Renal Failure: A Potential Safety Signal from Disproportionality Analysis of the WHO Safety Database," *Clinical Pharmacology and Therapeutics*, 109(4): 1021–1024.
Glennan, S. (2009) "Mechanisms," in H. Beebee, C. Hitchcock & P. Menzies (Eds.), *Oxford Handbook of Causation*. Oxford: Oxford University Press, 315–325.
Graedon, J. (2023) "Statin Side Effects Are All in Your Head! Really?" *The People's Pharmacy*. https://www.peoplespharmacy.com/articles/statin-side-effects-are-all-in-your-head-really
Guttinger, S. (2020) "Frankenswine, Genome Editing, and the Question of Sustainable Pig Farming," *Interdisciplinarity, Sustainability and Expert Disagreement*. Available at: https://interdisciplinarityandexpertdisagreement.wordpress.com/frankenswine-genome-editing-and-the-question-of-sustainable-pig-farming/
Hall, N. (1994) "Two Concepts of Causation," in J. Collins, N. Hall, & L. Paul (Eds.), *Causation and Counterfactuals*. Cambridge, MA: MIT Press, 225–276.
Hartley, S., Gillund, F., van Hove, L. & Wickson, F. (2016) "Essential Features of Responsible Governance of Agricultural Biotechnology," *PLoS Biology*, 14(5): e1002453. https://doi.org/10.1371/journal.pbio.1002453
Hempel, C. G. (1965) "Science and Human Values," in *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press, 81–96.
Howick, J. (2011) *The Philosophy Of Evidence-Based Medicine*. Chichester: Wiley-Blackwell, BMJ Books.
Illari, P. & Russo, F. (2014) *Causality: Philosophical Theory Meets Scientific Practice*. Oxford: Oxford University Press.
Ioannidis, J. P. A. (2020) "First Opinion: A Fiasco in the Making? As the Coronavirus Pandemic Takes Hold, We Are Making Decisions without Reliable Data," 17 March 2020, *STAT News*. https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/
Jukola, S. & Canali, S. (2021) "On Evidence Fiascos and Judgments in COVID-19 Policy," *History and Philosophy of the Life Sciences*, 43(2): 61.
Kistler, M. (2006) *Causation and Laws of Nature*. London: Routledge.
Lewis, D. (1973a) "Causation," in *Philosophical Papers v.2* (1986th ed.). Oxford: Oxford University Press, 159–213.
Lewis, D. (1973b) *Counterfactuals*. London: Blackwell.
Longino, H. (2002) *The Fate of Knowledge*. Princeton: Princeton University Press.
Longino, H. E. (1990) *Science as Social Knowledge: Values and Objectivity in Scientific Tradition and Change*. Princeton: Princeton University Press.
Mackie, J. L. (1974) *The Cement of the Universe*. Oxford: Oxford University Press.
Mellor, D. H. (1995) *The Facts of Causation*. London: Routledge.
Meyboom, R. H., Hekster, Y. A., Egberts, A. C., Gribnau, F. W. & Edwards, I. R. (1997) "Causal or Casual? The Role of Causality Assessment in Pharmacovigilance," *Drug Safety*, 17(6): 374–389.

Mumford, S. & Anjum, R. L. (2011) *Getting Causes from Powers*. Oxford: Oxford University Press.
Murad, H., Asi, N., Alsawas, M. & Alahdab, F. (2016) "New Evidence Pyramid," *Evidence Based Medicine*, 21(4): 125–127.
Psillos, S. (2002) *Causation and Explanation*. Montreal: McGill-Queen's University Press.
Putnam, H. (2002) *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, MA: Harvard University Press.
Rocca, E. & Andersen, F. (2017) "How Biological Background Assumptions Influence Scientific Risk Evaluation of Stacked Genetically Modified Plants: An Analysis of Research Hypotheses and Argumentations," *Life Sciences, Society and Policy*, 13(11). https://doi.org/10.1186/s40504-017-0057-7
Rocca, E. & Anjum, R. L. (2019) "Why Causal Evidencing of Risk Fails. An Example from Oil Contamination," *Ethics, Policy and Environment*, 22(2): 197–213.
Rocca, E. & Anjum, R. L. (2020) "Causal Evidence and Dispositions in Medicine and Public Health," *International Journal of Environmental Research and Public Health*, 17(6): 1813.
Rocca, E., Copeland, S. & Edwards, I. R. (2019) "Pharmacovigilance as Scientific Discovery. An Argument for Trans-Disciplinarity," *Drug Safety*, 42(10): 1115–1124.
Russo, F. & Williamson, J. (2007) "Interpreting Causality in the Health Sciences," *International Studies in the Philosophy of Science*, 21(2): 157–170.
Sackett, D. L., Richardson, W. S., Rosenberg, W. & Haynes, B. (1997) *Evidence-Based Medicine: How to Practice and Teach EBM*. London: Churchill Livingstone, Elsevier.
Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B. & Richardson, W. S. (1996) "Evidence-Based Medicine: What It Is and What It Isn't," *British Medical Journal*, 312(7023): 71-72.
Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
Shakir, S. A. W. & Layton, D. (2002) "Causal Association in Pharmacovigilance and Pharmacoepidemiology," *Drug Safety*, 25(6): 467–471.
Sønstebø, A. (2021) *Ingen Overdødelighet i 2020*. 11 March 2021, Statistics Norway. https://www.ssb.no/befolkning/fodte-og-dode/statistikk/dode/artikler/ingen-overdodelighet-i-2020
Suppes, P. (1970) *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.
Wickson, F. & Wynne, B. (2012) "Ethics of Science for Policy in the Environmental Governance of Biotechnology: MON810 Maize in Europe," *Ethics, Policy and Environment*, 15(3): 321–340.
Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
World Health Organization. (2020) *International Guidelines for Certification and Classification (Coding) Of COVID-19 as Cause of Death*. 20 April 2020. https://www.who.int/publications/m/item/international-guidelines-for-certification-and-classification-(coding)-of-covid-19-as-cause-of-death