
Inductive Risk, Epistemic Risk, and Overdiagnosis of Disease

Justin B. Biddle

Georgia Institute of Technology

Recent philosophers of science have not only revived the classical argument from inductive risk but extended it. I argue that some of the purported extensions do not fit cleanly within the schema of the original argument, and I discuss the problem of overdiagnosis of disease due to expanded disease definitions in order to show that there are some risks in the research process that are important and that very clearly fall outside of the domain of inductive risk. Finally, I introduce the notion of epistemic risk in order to characterize such risks.

1. Introduction

Philosophers interested in the role of values in science have focused much attention on the argument from inductive risk. In the 1950s and 1960s, a number of authors argued that value judgments play an ineliminable role in the acceptance or rejection of hypotheses (Hempel 1965; Rudner 1953). No hypothesis is ever verified with certainty, and so a decision to accept or reject a hypothesis depends upon whether the evidence is sufficiently strong. But whether the evidence is sufficiently strong depends upon the consequences (including ethical consequences) of making a mistake in accepting or rejecting the hypothesis. Recent philosophers of science have not only revived this argument; they have also extended it. While Rudner and Hempel focused on one point in the appraisal process where there is inductive risk—namely, the decision of how much evidence is enough to accept or reject a hypothesis—more recent philosophers of science have argued that there is inductive risk at multiple points in the research

This paper was first presented at the 2013 meeting of the American Society for Bioethics and Humanities. Thanks to Rebecca Kukla, Bryce Huebner, and the other conference participants for their comments. Thanks also to the Notre Dame Institute for Advanced Study for support.

process. Heather Douglas (2000) argues that inductive risk is present in the choice of methodology (e.g., the setting of a level of statistical significance), the characterization of evidence, and the interpretation of data. Torsten Wilholt (2009) argues that there is inductive risk in the choice of model organism. The upshot of these and other extensions of the Hempelian/Rudnerian argument from inductive risk—which I will call the *classical argument from inductive risk*—is, again, that the research process is shot through with inductive risk. Indeed, it can seem that, as a result of these extensions, there is inductive risk at any point in the research process at which a *decision* must be made.

While I applaud the extensions of the classical argument from inductive risk, and while I think that they provide valuable insights into the ways in which value judgments operate in the appraisal of research, I will argue that some of the purported extensions of the classical argument do not fit cleanly within the schema of the original argument and that, for the sake of conceptual clarity, they should simply be treated as different arguments. I will discuss the growing problem of overdiagnosis of disease due to expanded disease definitions in order to show that there are some risks in the research process that are important—and that should be taken seriously by philosophers of science—that very clearly fall outside of the domain of inductive risk. Finally, I will introduce the notion of *epistemic risk* as a means of characterizing such risks.¹ This more fine-grained taxonomy of risks in the research process will help to clarify the different roles that values can play in science.

2. The Classical Argument from Inductive Risk and Its Purported Extensions

Inductive risk is the risk of wrongly accepting or rejecting a hypothesis H , given a body of evidence E that is taken to support H . Given that scientific evidence cannot provide proof that a hypothesis is true, there is always some inductive risk of accepting H (given E) when H is false or rejecting H when H is true. A number of philosophers have drawn upon the notion of inductive risk in order to argue that “non-epistemic” factors—such as ethical values—play an essential role in hypothesis choice. Hempel, in his version of the argument from inductive risk, distinguishes between rules of confirmation and rules of acceptance (Hempel 1965, p. 92). Rules of confirmation specify what kind of evidence confirms a given hypothesis and what kind disconfirms it; in some cases, such rules provide a numerical degree of support for a given hypothesis. Rules of acceptance, on the other hand “specify how strong the evidential support for a hypothesis has to be

1. The concept of epistemic risk is one that I have co-developed with Bryce Huebner and Rebecca Kukla. I am grateful to them for valuable discussions on this issue. The phrase “epistemic risk” was also used by Collins (1996) in a different context for a different purpose.

if the hypothesis is to be accepted into the system of scientific knowledge” (Hempel 1965, p. 92). Clearly, one aims to accept hypotheses when they are true and to reject them when they are false, but because hypotheses cannot be proven to be true or false, there is always some possibility that one accepts a false hypothesis and rejects a true one. This is where ethical values come into play. If there are serious non-epistemic consequences of wrongly accepting or rejecting a hypothesis, then one should demand a relatively high degree of evidential support before accepting or rejecting the hypothesis; on the other hand, if there are no serious non-epistemic consequences, then one need not demand such a high degree of support. To paraphrase Rudner’s well-known example, if the hypothesis in question is that a drug does not have deadly side effects, we should require a high degree of confirmation before accepting the hypothesis, whereas if the hypothesis is that a lot of belt buckles is not defective, we need not require such confidence. “How sure we need to be before we accept a hypothesis will depend upon how serious a mistake would be” (Rudner 1953, p. 2). As this discussion illustrates, the core of the classical argument from inductive risk is that the decision of how much evidence is sufficient to accept or reject a hypothesis is one that has important non-epistemic consequences; as a result, non-epistemic factors such as ethical values legitimately influence the decision. In this paper, I will simply assume that the classical argument as put forward by Rudner is successful; however, the argument of this paper does not depend upon this assumption.²

Recent philosophers of science have extended the classical argument. In her examination of research on the carcinogenic effects of dioxins on laboratory rats, Douglas (2000) argues that inductive risk is present in the choice of methodology, the characterization of evidence, and the interpretation of data. Her primary example of inductive risk in methodological choice is the choice of a level of statistical significance (Douglas 2000, p. 565). If one sets a relatively high level of statistical significance, then one is expressing a relative tolerance of false negatives, where the hypothesis in question is that dioxins cause cancer in laboratory rats. This can have important implications for policy making, as it could lead to under-regulation of dioxins, which in turn could put human beings at significant risk of harm. On the other hand, if one sets a relatively low level of statistical significance, one expresses a relative tolerance of false positives. This could lead to over-regulation, which in turn could harm businesses that might otherwise profit from the use of these chemicals. As an example of inductive risk in evidence characterization, Douglas

2. The two most well-known objections to Rudner’s version of the argument are put forward by Jeffrey 1956 and Levi 1960. For responses to Jeffrey, see Biddle and Winsberg 2010 and Biddle 2013. For responses to Levi, see Wilholt 2009.

discusses the characterization of rat liver slides. In the course of their research, scientists had to examine slides of rat livers in order to ascertain the presence or absence of tumors and, if tumors were present, whether or not they were malignant (Douglas 2000, p. 569). While some slides were unambiguous, many were not, such that different groups of scientists characterized the slides in very different ways. In some cases, there were disagreements within a group, which were resolved by majority vote. According to Douglas, there is inductive risk in the characterization of these data, because different characterizations of ambiguous slides have different implications for policy making and regulation. Finally, as an illustration of inductive risk in the interpretation of results, Douglas discusses the choice of a model for extrapolating data (Douglas 2000, p. 573). There is a controversy within the relevant toxicology community about the effects of low-dose exposure to dioxins. Some argue that there is a threshold below which there is no carcinogenic effect of exposure. Others argue that there is no threshold; some exposure, no matter how little, presents some risk of cancer. As in the other cases, Douglas argues that there is inductive risk in the choice of a model for extrapolating data, due to the different implications of each choice for policy making and regulation.

One additional extension of the classical argument from inductive risk comes from Wilholt's example of research on the health effects of exposure to bisphenol A (Wilholt 2009). While many studies have drawn an association between bisphenol A and adverse health effects, there are other studies (particularly industry-funded studies) that have found no such correlation. Investigation into these latter studies reveals that many of them used as model organisms a strain of rat that was particularly insensitive to estrogen. Given that the toxicity of bisphenol A is associated with its similarity to human estrogen, this choice of a model organism significantly lowered the probability of finding an association between bisphenol A and adverse health effects. Thus, Wilholt argues that there is inductive risk in the choice of model organism.

3. Are These *Inductive* Risks?

Some of these extensions fit nicely within the model of the classical argument from inductive risk. Douglas's example of the choice of a level of statistical significance fits very clearly within this model. When one sets a relatively high (or low) level of statistical significance, one demands a relatively high (or low) degree of confirmation before accepting a hypothesis. Thus, the choice of a level of statistical significance represents an answer to the question of how much evidence is enough to accept or reject a hypothesis.

The other extensions, however, might not fit so cleanly within the model of the classical argument. In the case of the characterization of evidence, one might argue that what is at issue is not whether there is enough evidence to

accept or reject a hypothesis, but what the evidence itself is. On this account, evidence characterization does not fit within the model of the classical argument at all. One might respond that, if one shifts one's focus, this case can be made to fit the model; if one treats the slides themselves as the evidence and statements such as "this slide indicates the presence of a malignant tumor" as the hypotheses, then this case might fit within the original model. It must be said that, even in this case, evidence characterization might not fit perfectly within this model, as one could argue that it makes no sense in this case to ask how much evidence is enough to accept the hypothesis. In this case, we have a fixed amount of evidence—one slide—and the question is whether this fixed amount of evidence is sufficient to accept or reject the hypothesis in question. Putting this worry aside, not all theories of evidence conceive of objects such as slides as evidence. On the account of evidence put forward by logical empiricists such as Hempel, only particular types of statements count as evidence. If one considers the relationship between theory and evidence to be necessarily relationship between statements, then evidence characterization as described by Douglas cannot fit within the model of the classical argument. I do not wish to present an argument for a particular account of evidence here—in fact, I find it perfectly reasonable to consider objects such as slides to be evidence—but only to point out that whether or not evidence characterization fits within the model of the classical argument depends upon a particular theory of evidence.

The choice of a model for extrapolating data is the case that most clearly does not fit within the original model. Douglas, in my view, is right to argue that ethical considerations should play a role in the choice of a model for extrapolating data, but this choice is not best seen as a decision of how much evidence is sufficient to accept or reject a hypothesis. The view that there is a threshold below which there is no carcinogenic effect is most plausibly seen as an assumption (or as relying directly upon an assumption). According to Douglas, there are some in the toxicology community who hold "as a basic assumption" that "the dose determines the poison" (Douglas 2000, p. 574). "Under this view, it is *generally assumed* that every poison has some threshold for its toxic effects; nothing is biologically potent at every dose" (2000, p. 574; emphasis added). If this is true, then the view that there is a threshold is not a hypothesis accepted on the basis of confirmatory evidence; it is simply an assumption. Douglas continues by noting that some scientists involved in dioxin research provide an argument for the threshold model. They argue that dioxins do not cause liver cancer directly but cause acute toxicity in livers, and this in turn causes liver cancer. Coupled with the assumption that acute toxicity in livers has a threshold, it follows that there is also a threshold for dioxins. But this

argument relies directly on the background assumption that there is a threshold for liver toxicity. Thus, the view that there is a threshold either is an assumption or a conclusion that relies directly upon a related assumption.

The other view, that there is no threshold, is also most plausibly seen as an assumption. If dioxin were a mutagen—that is, a chemical that can damage DNA such that cancer growth results—then the no-threshold model would be appropriate, as any exposure to a mutagenic chemical, no matter how low, is thought to have some probability of causing tumor growth. But dioxin appears not to be a mutagen, but rather a promoter; that is, it promotes cancer growth given the presence of a mutagen, but it does not damage DNA in such a way as to cause cancer growth. It is unclear whether or not promoters have thresholds. Thus, the question of whether or not dioxins have thresholds comes down to, in Douglas's words, "opposing intuitions" (2000, p. 574) or "competing general assumptions" (2000, p. 575). Because the issue in this instance is not whether there is sufficient evidence to accept or reject a hypothesis, but rather which competing assumptions to accept, the choice of a model for extrapolating data does not fit within the model of classical argument from inductive risk.

The issues surrounding the choice of a model organism are complex, but I would argue that, at the very least, this choice does not fit neatly within the model of the classical argument. What is primarily at issue in this case is not whether there is sufficient evidence to accept or reject a hypothesis, but rather whether this hypothesis, and the body of evidence that supports it, is relevant to some other hypothesis. Consider a body of evidence E_{eir} (exposure of estrogen-insensitive rats to bisphenol A) that stands in a relation of confirmation to the hypothesis H_{eir} (bisphenol A causes adverse health effects in this strand of estrogen-insensitive rat). Obviously, there is inductive risk associated with the question of how much evidence is sufficient to accept the hypothesis. But for the purpose of determining whether to regulate bisphenol A, H_{eir} is not particularly salient. Ultimately, what we care about is whether bisphenol A causes adverse health effects in human beings, and human beings happen to be sensitive to estrogen and other related chemicals. Thus, the more important hypothesis is H_{esr} (bisphenol A causes some particular adverse health effect in strands of rats that are sensitive to estrogen). Unfortunately, E_{eir} is simply not relevant to H_{esr} . Thus, the main issue here is not how much evidence (E_{eir}) is sufficient to accept or reject a hypothesis (H_{eir}), but rather whether this evidence and hypothesis are relevant to another hypothesis (H_{esr}) that is particularly important to us.

I have argued thus far that some of the extensions of the classical argument from inductive risk do not fit well within the original model. I have not argued that these extensions *definitively* do not fit within the original

model, only that they do not fit cleanly within it; they can perhaps be made to fit, but this would require twisting, contorting, or just plain hammering. Each of these extensions highlights genuine *risks* and each identifies points in the research process where values (often ethical values) play an ineliminable role; as such, they provide important insights into the role of values in science. But they are not best seen as highlighting *inductive risks*. In the next section, I will provide an example that demonstrates even more clearly that there are important risks within the research process that fall outside of the domain of inductive risk, and I will use this example to subsequently introduce the notion of epistemic risk.

4. Overdiagnosis of Disease and Expansion of Disease Definitions

Overdiagnosis of disease is becoming a serious problem in health care. Overdiagnosis occurs when people without symptoms are diagnosed with a disease that will never lead them to experience symptoms or early death. Problems resulting from overdiagnosis include increased risk to patients who undergo procedures that they do not need, wasted time and energy that could have been spent elsewhere, and money. It is estimated that in 2011 in the United States alone, between \$158–226 billion were wasted on unnecessary treatments (Berwick and Hackbarth 2012). I will not take a stand on just how serious the problem of overdiagnosis is or whether it is a more serious problem than underdiagnosis; my aim in discussing it is simply to highlight an example of risk that is important and that very clearly falls outside of the domain of inductive risk.

There are a number of different causes of overdiagnosis (Moynihan et al. 2012). The first is technological; as the sensitivity of diagnostic tests has increased, so has the detection of “diseases” that will never progress to the point of causing symptoms or early death. For example, the United States National Cancer Institute has estimated that “a ‘perfect’ breast cancer screening test would identify approximately 10% of ‘normal’ women as having breast cancer, even though most of those cancers would probably not result in illness or death” (National Cancer Institute 2015). A second cause of overdiagnosis, which works in conjunction with the first, is the adoption of clinical practice guidelines and recommendations that encourage early testing. In order to catch disease at earlier stages, many have recommended that patients undergo testing for disease, whether or not they experience symptoms. These recommendations have undoubtedly had beneficial effects, as some patients have received treatment for diseases that they otherwise would not have received (or not received as early). But at the same time, they can lead to overdiagnosis and overtreatment. For example, a systematic review estimates that up to one third of all screening-detected cancers are overdiagnosed (Moynihan et al. 2012), and many

patients would not have undergone these tests in the absence of policies encouraging them to do so. A third cause of overdiagnosis comes from within the research establishment itself, in the form of expansions of disease definitions and lowering of thresholds for treatment. For example, in 2010 the definition of gestational diabetes was expanded, which had the effect of more than doubling the number of woman classified to almost 18% (Moynihan et al. 2012). Overdiagnosis due to expanded disease definitions is particularly relevant to the discussion of inductive risk and epistemic risk; as a result, the remainder of this paper will focus on it.

Recently, there has been a consistent trend of researchers expanding disease definitions (Welch et al. 2011; Hoffman and Cooper 2012). For example, a recent systematic review analyzed publications between 2000 and 2013 from national and international panels that made decisions regarding disease definitions and diagnostic threshold criteria for common conditions in the United States (Moynihan et al. 2013). Upon examining sixteen publications on fourteen common conditions, it found that ten of these publications proposed changes widening definitions and/or diagnostic criteria; one proposed changes narrowing definitions, and five were unclear. There are a variety of different (and mutually consistent) explanations for this phenomenon. One is that researchers are expanding disease definitions out of an interest in benefiting patients—that is, in diagnosing and treating diseases before they cause serious harm. Many researchers are no doubt motivated by this goal. But there are also financial interests at play. Expanding the definition of a disease can be financially very lucrative for companies that sell treatments, as it can expand significantly the number of potential patients. The aforementioned systematic review included data that correlated expansion of disease definitions with financial conflicts of interest. Of the fourteen panels with conflict of interest disclosures, the average proportion of members with financial ties to industry was 75%. Most of the industries to which members had ties were active in the markets for the respective panels' conditions. Among members with ties, the median number of companies to which they reported ties was seven. Finally, almost all of the committees that disclosed conflicts of interest were chaired by people with ties to industry (twelve out of fourteen). Another study finds that 69% of the panel members of the *Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5)*—who set disease definitions and diagnostic criteria—reported financial ties to the pharmaceutical industry; this is up from 57% of *DSM-IV* panel members (Cosgrove and Krinsky 2012).

Moynihan et al. (2013, 9) are careful to assert that their study does not demonstrate a causal connection between financial ties and expanded disease definitions. Nonetheless, they emphasize that their findings should

be interpreted in light of the extensive literature on the prevalence and effects of financial conflicts of interest in biomedical research. Numerous studies, including two large systematic review articles, have confirmed the existence of a “funding effect” –i.e., an association between industry-sponsored research and pro-industry conclusions (e.g., Bekelman et al. 2003; Lexchin et al. 2003).³ Given this background, it is not unreasonable to believe that financial ties play a causal role in the expansion of disease definitions.

The phenomenon of expanding disease definitions serves as an illustration of an important risk that falls outside of the domain of inductive risk. There is clearly a risk involved in expanding the definition of a disease (just as there is a risk involved in narrowing the definition of a disease). Defining a disease relatively broadly will likely increase early diagnosis, which can lead the prevention of symptoms later, but it will also increase overdiagnosis. Defining a disease relatively narrowly, on the other hand, will tend to decrease both early diagnosis and overdiagnosis. There are different non-epistemic consequences that result from different disease definitions, and there are different values and interests that might motivate researchers to propose particular definitions. Because of this, identifying the risks of defining diseases in different ways helps to shed light on the role of values in biomedical research and hence is a worthwhile endeavor for philosophers of science. Yet the risk associated with defining a disease is very different from the question of how much evidence is sufficient to accept or reject a hypothesis; as a result, it is not a case of inductive risk.

This is not to say that defining a disease has nothing to do with evidential considerations; clearly, defining a particular disease requires that we take into account evidence regarding the effects of particular conditions. For example, in setting criteria for determining who has osteoporosis, it is important to consider evidence on the effects of that condition. But ultimately, the decision to define the disease more broadly or more narrowly will depend significantly on value judgments regarding the costs of living with the condition.⁴ More precisely, it will (or at least should) be based on judgments regarding the costs of living with symptoms from the condition versus the costs of overdiagnosis and treatment of a condition

3. For a more recent examination of the relationship between industry sponsorship and outcomes of cancer trials, see Djulbegovic et al. 2013. Thanks to an anonymous reviewer for bringing this study to my attention.

4. It should be emphasized that, in this paper, I am interested in the question of defining particular diseases, not the question of defining disease as such. There is a large literature on the question of whether or not the concept disease can be understood in purely objective terms (e.g., Boorse 1977). Whether or not this is the case, it should be clear that defining particular diseases requires making value judgments.

that will never cause symptoms. Defining a disease, thus, is not primarily an issue of accepting a hypothesis on the basis of evidence, and thus is not best thought of as a case of inductive risk. An example will help to illustrate this.

Since the 1990s, the definition of osteoporosis has expanded significantly.⁵ Prior to the 1990s, the definition was relatively narrow, such that it included primarily patients who suffered symptoms—in particular, painful fractures that were not associated with major trauma. In 1994, the World Health Organization (WHO) expanded the definition to include patients who suffer symptoms as well as patients who are asymptomatic but who have low bone mineral density. More recently, the National Osteoporosis Foundation (NOF) has expanded the definition further to include women with denser bones than were included under the WHO guidelines. The definition of osteoporosis is based on a “T-score” measurement of bone mineral density, which is a number that quantifies bone density relative to the average among 20–29 year old white women. According to the WHO guidelines, osteoporosis is defined as having a T-score of lower than -2.5 (the closer to zero the score, the better the bone density). According to the more recent NOF recommendations, treatment is recommended for all women who have a T-score of less than -2.0 and, for women with particular risk factors, less than -1.5 (Herndon et al. 2007, p. 1703).

It is estimated that these expanded definitions have increased the number of women treated over the age of 65 from 6.4 million (T-score less than -2.5) to 10.8 million (T-score less than -2.0) or 14.5 million (T-score less than -1.5). For women between 50–64 years old, the estimated increase is from 1.6 million to 4.0 million or 7.3 million. Under the WHO guidelines, 7 percent of women under the age of 65 are recommended for treatment; under the NOF guidelines, the percentages are approximately 20 (T-score less than -2.0) and 30 (T-score less than -1.5). At the same time, the risk of fractures to these additional women is quite low; for example, the ten-year risk of hip fractures (the most common type of fracture) for these women below the age of 65 is 0.5 percent to 2 percent (Herndon et al. 2007, p. 1707). The definition of osteoporosis has expanded, but not on the basis of new evidence regarding increased risk of fracture.

This expanded definition will likely result in some people benefiting from treatment that they otherwise would not have received. At the same time, there are significant costs associated with broadening criteria for treatment. One of these is financial; the net cost of treating the additional women with T-scores less than -2.0 is \$28 billion for women over the age of 65 and \$18 billion for women between the ages of 50–64. Another cost

5. My discussion of osteoporosis relies primarily upon Herndon et al. 2007.

is risk to patients from treatments. No treatment is risk free; in this case—treating low bone mineral density to prevent fractures—the treatment presents risks to patients, including esophagitis, thromboembolic events, or osteonecrosis (Herndon et al. 2007, pp. 1707–8). There is also relatively little data on the effectiveness of this treatment for many groups (Herndon et al. 2007, pp. 1707–8). Given the earlier discussion of conflicts of interest, it is probably not surprising that the committees that recommended the expanded definitions—both in the case of the WHO and the NOF—both received funding from the pharmaceutical industry (Herndon et al. 2007, pp. 1709–10).

The case of the expanding definitions of osteoporosis provides an illustration of how defining particular diseases presents risks that are not inductive risks. The definition of a particular disease is necessarily and heavily value laden; nature does not dictate how a disease should be defined. Rather, we determine criteria for diseases largely on the basis of value judgments concerning the costs and benefits of living with (or without) particular conditions. Given the discussion of overdiagnosis, it should be clear that there are risks to expanding (as well as to narrowing) disease definitions. But these are not best seen as inductive risks of accepting false hypotheses on the basis of evidence. A broader notion of risk, which I call *epistemic risk*, is required.

5. Epistemic Risk and Overdiagnosis of Disease

Epistemic risk is the risk of being wrong. One can wrongly accept or reject a hypothesis given evidence that is taken to support that hypothesis—thus the concept of epistemic risk includes the concept of inductive risk—but one can also wrongly accept or reject many other things, including a methodology, a background assumption, a set of test subjects, a policy and, as I will argue, a definition. None of these epistemic risks fall naturally within the concept of inductive risk, and yet they all have implications for the role of values in science and thus deserve to be taken seriously by philosophers of science. My primary aim in this section is to sketch a sufficient condition for wrongly accepting or rejecting a definition, but to do this I will first need to provide some sufficient conditions for wrongly accepting or rejecting a policy. Due to limitations of space, I do not attempt to provide a complete account of wrongly accepting or rejecting either a definition or a policy; rather, I aim to sketch some sufficient conditions that are particularly relevant to the case of overdiagnosis due to expanded disease definitions.⁶

6. One might wonder whether one can wrongly accept a methodology or a set of test subjects, as I stated earlier in this paragraph. While a full discussion of this issue is beyond the scope of this paper, I believe that one can wrongly accept a methodology or a set of test

There is at least one very clear sense in which one can wrongly accept a policy. If one affirms goal G , if one adopts policy P as a means of achieving G , and if pursuing P will not lead to the achievement of G , then one wrongly accepts P . Wrongly accepting a policy, in this sense, is to make its wrongness relative to a particular goal. There is also a stronger, non-relativistic sense of wrongly adopting a policy. In their well-known text on biomedical ethics, Beauchamp and Childress (2012) identify four mid-level ethical principles: non-maleficence, beneficence, respect for autonomy, and justice. They employ these principles primarily to evaluate individual actions in the biomedical arena, but they can also be used to evaluate policies more generally. Application of these principles does not always specify uniquely which action to pursue, or which policy to adopt, as these principles can be interpreted differently and can be weighted differently with respect to one another. For example, in determining whether to perform a particular intervention on a patient, a doctor will often have to weigh the principles of non-maleficence and beneficence and make a judgment as to which is most important in that particular situation. But there are some actions and policies that will not satisfy these principles, given any plausible interpretation of them and relative weighting. A policy of forced sterilization, for example, fails to satisfy any plausible interpretation and weighting of these principles.

Given this account of wrongly accepting a policy, we are in a position to provide a sufficient condition for wrongly accepting a definition of a disease: one wrongly accepts a definition of a disease if that definition contributes in a direct way to wrongly accepting or rejecting a policy. In the case of screening for a disease, a reasonable goal would be to prevent all and only those conditions that will cause symptoms or early death. If one defines a disease too narrowly, underdiagnosis will result; if one defines a disease too broadly, “diseases” will be diagnosed that never cause symptoms or early death. Either of these would count as wrongly accepting the definition of a disease. It can, of course, be difficult in some cases to ascertain whether a policy has been wrongly adopted, or whether a particular definition of a disease contributes in a direct way to wrongly accepting that

subjects, depending upon the goals of a particular research endeavor. For example, if one is conducting an experiment and one strongly wishes to avoid false positives, then choosing a p-value of 0.5 would be wrong. If one is overseeing a clinical trial in which one wishes to ascertain the frequency of a particular side effect in an elderly population, then it would be wrong to enroll only young people as test subjects. How one characterizes the risk of wrongly accepting a methodology, as opposed to wrongly accepting a definition or a policy or a hypothesis given a body of evidence, will vary depending on the case. But each of these risks can be thought of as epistemic risks, as they each involve the risk of being wrong. Thanks to an anonymous referee for encouraging further discussion on this point.

policy. Making either of these determinations will depend upon empirical evidence concerning the consequences of policies and the contributions of definitions toward those policies. But in many cases, such determinations can be made. Given the growing concerns about the problem of overdiagnosis, and given the increasing influence of financial interests in expanding disease definitions, one can make at least a *prima facie* case that researchers have, in some instances, wrongly expanded disease definitions.

6. Conclusion

I have argued that some of the purported extensions of the classical argument from inductive risk do not fit within the schema of the original argument, and I have discussed the example of overdiagnosis of diseases due to expanded disease definitions in order to provide a clear case of a risk that is both important and that falls outside of the domain of inductive risk. In order to characterize such risks, I have introduced the notion of epistemic risk, which allows for a more precise characterization of the decisions that are made in the research process, the variety of implications that these decisions can have on society more broadly, and on the role of values in science.

References

- Beauchamp, Tom and James Childress. 2012. *Principles of Biomedical Ethics*. New York: Oxford University Press.
- Bekelman, Justin E., Yan Li, Cary P. Gross. 2003. "Scope and Impact of Financial Conflicts of Interest in Biomedical Research: A Systematic Review." *Journal of the American Medical Association* 289: 454–465.
- Berwick, Donald and Andrew Hackbarth. 2012. "Eliminating Waste in US Health Care." *JAMA* 307: 1513–1516.
- Biddle, Justin. 2013. "State of the Field: Transient Underdetermination and Values in Science." *Studies in History and Philosophy of Science* 44: 124–133.
- Biddle, Justin and Eric Winsberg. 2010. "Value Judgements and the Estimation of Uncertainty in Climate Modeling." Pp. 172–197 in *New Waves in Philosophy of Science*. Edited by P. D. Magnus and J. Busch. Basingstoke, Hampshire: Palgrave MacMillan.
- Boorse, Christopher. 1977. "Health as a Theoretical Concept." *Philosophy of Science* 44: 542–573.
- Collins, Arthur W. 1996. "Moore's Paradox and Epistemic Risk." *The Philosophical Quarterly* 46: 308–319.
- Cosgrove, Lisa and Sheldon Krimsky. 2012. "A Comparison of DSM-IV and DSM-5 Panel Members' Financial Associations with Industry: A

- Pernicious Problem Persists." *PLoS Med* 9 (3): e1001190. doi:10.1371/journal.pmed.1001190
- Djulfbegovic, Benjamin, Ambuj Kumar, Branko Miladinovic, Tea Reljic, Sanja Galeb, Asmita Mhaskar, Rahul Mhaskar, Iztok Hozo, Dongsheng Tu, Heather A. Stanton, Christopher M. Booth, and Ralph M. Meyer. 2013. "Treatment Success in Cancer: Industry Compared to Publicly Sponsored Randomized Controlled Trials." *PLoS ONE* 8 (3): e58711. doi:10.1371/journal.pone.0058711
- Douglas, Heather. 2000. "Inductive Risk and Values in Science." *Philosophy of Science* 67: 559–579.
- Hempel, Carl. 1965. "Science and Human Values." In *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, 81–96. New York: The Free Press.
- Herdon, M. Brooke, Lisa M. Schwartz, Steven Woloshin, and H. Gilbert Welch. 2007. "Implications of Expanding Disease Definitions: The Case of Osteoporosis." *Health Affairs* 26: 1702–1711.
- Hoffman, Jerome and Richelle Cooper. 2012. "Overdiagnosis of Disease: A Modern Epidemic." *Archives of Internal Medicine* 172: 1123–1124.
- Jeffrey, Richard. 1956. "Valuation and Acceptance of Scientific Hypotheses." *Philosophy of Science* 23: 237–246.
- Levi, Isaac. 1960. "Must the Scientist Make Value Judgments?" *Journal of Philosophy* 57: 345–357.
- Lexchin, J., L. Bero, B. Djulfbegovic, and O. Clark. 2003. "Pharmaceutical Industry Sponsorship and Research Outcome and Quality." *BMJ* 326: 1167–1170.
- Moynihan, Raymond N., Georga P. E. Cooke, Jenny A. Doust, Lisa Bero, Suzanne Hill, and Paul P. Glasziou. 2013. "Expanding Disease Definitions in Guidelines and Expert Panel Ties to Industry: A Cross-Sectional Study of Common Conditions in the United States." *PLoS Med* 10 (8): e1001500.
- Moynihan, Ray, Jenny Doust, and David Henry. 2012. "Preventing Overdiagnosis: How to Stop Harming the Healthy." *BMJ* 344: e3502.
- National Cancer Institute. 2015. "Breast Cancer Screening (PDQ®) – Harms of Screening Mammography." <http://www.cancer.gov/cancertopics/pdq/screening/breast/healthprofessional/page8> (accessed 11 March, 2015).
- Rudner, Richard. 1953. "The Scientist *Qua* Scientist Makes Value Judgments." *Philosophy of Science* 20: 1–6.
- Welch, H. Gilbert, Lisa Schwartz, and Steven Woloshin. 2011. *Overdiagnosed: Making People Sick in the Pursuit of Health*. Boston: Beacon.
- Wilholt, Torsten. 2009. "Bias and Values in Scientific Research." *Studies in History and Philosophy of Science* 40: 92–101.