

14 New Directions in Philosophy of Medicine

*Jacob Stegenga, Ashley Graham
Kennedy, Şerife Tekin, Saana
Jukola, and Robyn Bluhm*

Introduction

As we have seen from the previous chapters, philosophy of medicine is a dynamic area of research, raising, and seeking to answer, a plethora of metaphysical, practical, and moral questions in medicine. Such questions are of importance not just for their intrinsic philosophical interest, but also because they have implications for medical research, practice, and policy. Recent work in philosophy of medicine has addressed such questions as the appropriate evidence base for medicine, the nature and definition of “health” and of “disease,” and the relative contributions of scientific research and patients’ experiences to an understanding of medical questions.

The purpose of this chapter is to describe what we see as several important new directions for philosophy of medicine. This recent work (i) takes existing discussions in important and promising new directions, (ii) identifies areas that have not received sufficient and deserved attention to date, and/or (iii) brings together philosophy of medicine with other areas of philosophy (including bioethics, philosophy of psychiatry, and social epistemology). To this end, the next part focuses on what we call the “epistemological turn” in recent work in the philosophy of medicine; the third part addresses new developments in medical research that raise interesting questions for philosophy of medicine; the fourth part is a discussion of philosophical issues within the practice of diagnosis; the fifth part focuses on the recent developments in psychiatric classification and scientific and ethical issues therein, and the final part focuses on the objectivity of medical research.

The Epistemological Turn

Some of the best scholarship in philosophy of medicine in the past two decades has been about the epistemology of medical research. This is a welcome

development after a period in which much good work in philosophy of medicine was focused on conceptual topics, most notably analyses of health and disease, or normative topics, especially ethical issues that arise in medical practice. Philosophers such as John Worrall, Nancy Cartwright, Alex Broadbent, Kirstin Borgerson, Phyllis Illari, and Jeremy Howick have been at the forefront of this turn to epistemology in the philosophy of medicine. Two of the most prominent topics in the epistemological turn were the epistemic merits of randomized controlled trials (RCTs) and the role of mechanisms in grounding various sorts of causal claims. Here we describe four ways in which present and future work in the philosophy of medicine has been and will continue to develop this turn toward epistemology.

Some of the central epistemological debates in philosophy of medicine have been about the epistemic status of RCTs and systematic reviews. RCTs are considered to be one of the pillars of the evidence-based medicine (EBM) movement, and systematic reviews are also placed at the top of EBM evidence hierarchies. However, both have recently been the targets of philosophical criticism (Worrall, 2002, 2007; Borgerson, 2009; Stegenga, 2011). Other philosophers suggest that knowledge of mechanisms can aid in making causal inferences, while theoreticians in the EBM movement tend to hold that reasoning from mechanisms is too often unreliable (Howick, 2011a; Russo and Williamson, 2007). These EBM methodological principles and their associated philosophical critiques have tended to articulate epistemological merits and vices at a rather coarse grain: is randomization necessary to support reliably causal hypotheses in medicine? Is meta-analysis the platinum standard of evidence? Is knowledge of mechanisms necessary to infer causation or to warrant extrapolation? Such questions could be fruitfully addressed at a finer grain.

An example of a fine-grained approach to methodology is the assessment of particular details of RCTs rather than arguing about the merits of randomization generally. Likewise, McClimans (2013) investigates the various measurement instruments employed in RCTs, thereby unpacking details about RCTs left unanalyzed when the merits of RCTs simpliciter are debated. Similarly, rather than argue about the merits of meta-analysis tout court, as Stegenga (2011) does, one could articulate the ways in which meta-analyses can be better or worse. Or, to take another example, rather than arguing whether knowledge of mechanisms is necessary for causal inferences in medicine, one could attempt to formulate precisely how mechanisms can aid in causal inferences and extrapolation from experimental populations to target populations. Steel (2007), for example, argues that “comparative process tracing” of mechanisms can aid in extrapolating causal knowledge.

Another example of fine-grained approaches in the epistemological turn is evidence hierarchies. Arguments for and against the standard evidence hierarchies have recently been articulated (Howick, 2011b; Stegenga, 2014). But

Osimani (forthcoming) argues that different kinds of hypotheses, such as hypotheses about harms of medical interventions, require different kinds of evidence compared to hypotheses about benefits of medical interventions; thus, general arguments about the justification of EBM evidence hierarchies might be too coarse-grained. Indeed, changing the metaphor, Bluhm (2005) argues that evidence hierarchies should be replaced by evidence networks, to take into account the rich information that various types of studies in clinical research can provide. One final example is the ways in which recent trial designs attempt to take into account some of the complexities of clinical practice. In short, a fine-grained analysis of the methodological details of medical research is on the forefront of the epistemological turn in philosophy of medicine.

The second aspect of the epistemological turn in philosophy of medicine involves articulating the intersection between social, ethical, and methodological aspects of medical research. An example of such a concern of present (and future) work in philosophy of medicine is to articulate the methodological and social conditions under which many of the problems of medical research are possible. For example (Jukola, 2015; see also the end of this chapter), argues that a compelling way to understand the shortcomings of meta-analysis requires not just an examination of the methodological details of meta-analysis, but also an examination of the social context in which this technique is employed. As another example, De Melo-Martin and Intemann (2011) argue for a feminist approach to understanding problems with contemporary biomedical research. A different sort of interest regarding the relationship between social, ethical, and methodological aspects of medical research investigates the influence that social or ethical values can have on the production and interpretation of evidence. In a widely discussed paper, Douglas (2000) argues for the central role of values in scientific reasoning, and this thesis is especially prominent in medical research. For example, Kennedy (2013) argues that ethics and evidence are “intertwined” in the practice of differential diagnosis. As another example, Tekin (2014) argues that psychiatric nosology is “at the crossroads of science and ethics.”

In the past decade, there have been many proclamations of a thesis that one could call “medical nihilism.” Various versions of medical nihilism emphasize the lack of reproducibility of many high-profile research findings in medicine, the nefarious activities of medical scientists associated with pharmaceutical companies, and perhaps most troubling, the low effectiveness of the vast majority of recent pharmaceuticals. Contributors to this literature include prominent physicians, epidemiologists, and journalists. Recent examples include books by Marcia Angell (2004), Moynihan and Cassels (2005), Carl Elliott (2010), Ben Goldacre (2012), and Peter Gøtzsche (2013), and articles by epidemiologists such as John Ioannidis, Lisa Bero, Peter Jüni, and Jan Vandenbroucke. A third task for present philosophy of medicine in the epistemological turn is to assess

just how deep and troubling the thesis of medical nihilism is. Stegenga (forthcoming), for example, argues that medical nihilism is a compelling thesis for much of recent medicine.

Medical nihilism is motivated, in part, by noting problems with the socio-political nexus in which medical research takes place. A fourth set of concerns for present and future philosophy of medicine in the epistemological turn is to address questions such as: How should medical research be modulated given the recent work on epistemology of such research? Who should pay for medical research? What kinds of projects should be prioritized by funders of medical research? Should the results and products of medical research be protected by intellectual property laws? As an example of recent work in this domain, Brown (2008) argues that medical research should be socialized and the results of medical research should not be protected by patent laws. A broadly similar proposal is suggested by Reiss (2010). The epistemological turn has uncovered numerous epistemological problems with contemporary medical research, and such problems call for normative guidance.

In what follows, we discuss some of these elements of the epistemological turn in the philosophy of medicine in more detail, including issues related to evidence hierarchies, extrapolation, diagnosis, the construction of psychiatric categories, and the pursuit of objectivity in medical research.

Beyond RCTs and Meta-analyses

In addition to these recent developments in philosophy of medicine, there are a number of developments in medical research itself that should be of interest to philosophers. As noted earlier, much of the work being done on philosophical questions raised by clinical research has examined RCTs. This is in large part due to the influence of EBM and, in particular, to its “hierarchy of evidence,” which stipulates that the best quality evidence comes from RCTs and systematic reviews of RCTs. The rationale behind the hierarchy of evidence is that designs higher on the hierarchy have greater internal validity and more precision in their estimates of outcomes than those lower on the hierarchy. Yet, as critics of EBM have pointed out, this precision may come at the cost of generalizability. First, because many RCTs have strict inclusion and exclusion criteria, many of the patients who might be treated with a drug in clinical practice would not have qualified for participation in the trial that tested that drug. Second, because RCTs compare average outcomes in the treatment and the control groups for a study, they give little information about differences in outcomes among the patients in the study. Meta-analyses exacerbate this problem, since they usually average the (average) results across studies to get a more precise estimate of outcomes.

While EBM itself has not tended to acknowledge the gravity of these problems, there have been a number of recent trends in medical research that have attempted to address the issue of variable treatment outcomes, either by conducting research in a setting that more closely resembles clinical practice or by studying outcomes specifically in groups of patients who have a particular demographic or physiological characteristic. This section surveys several of these trends and outlines their potential interest for philosophers of medicine.

Research Generalizability and Clinical Care

RCTs with strict inclusion and exclusion criteria tend to be so-called explanatory trials, which aim to provide evidence that an intervention causes an outcome of interest by showing that, in carefully controlled, “ideal” conditions, a therapeutic intervention provides better outcomes than a placebo control (Thorpe et al., 2009). The strength of these trials is that they are considered to provide evidence that the experimental intervention causes the outcomes(s) being measured; this is because the trial is designed to ensure that, as far as possible, the only difference between the treatment and the control groups is that the treatment group receives the intervention being tested. An efficient way to show differences between study groups is to minimize variability *within* groups. To achieve this, explanatory studies have strict inclusion and exclusion criteria; for example, limiting eligibility to a narrow age range and enrolling only patients with no comorbid conditions, who are not taking concomitant medications. As a result of these decisions about study eligibility, subjects who qualify for participation in explanatory trials are different in important respects from typical patients who are treated in clinical practice. It is therefore not clear how well causal relationships between treatment and outcome observed in subjects enrolled in explanatory studies might hold up in other kinds of patients.

Because of this, greater attention is now being paid to research that is more closely integrated with clinical practice. In a 2007 workshop report, the US Institute of Medicine called for the development of a “learning healthcare system,” defined as a system of healthcare “in which knowledge generation is so embedded into the core of the practice of medicine that it is a natural outgrowth and product of the healthcare delivery process and leads to continual improvement in care.” Although the workshop addressed a broad range of issues, one key point made was that RCTs are not sufficient to inform clinical practice, in part because of the concerns about generalizability noted earlier.

Learning healthcare systems would conduct a variety of kinds of research, including long-term observational studies and studies using administrative databases and patient records. They would still conduct RCTs, but these studies would tend to be pragmatic rather than explanatory; that is, they would be

designed to reflect the conditions in which an intervention is used in practice, rather than an idealized test environment. Pragmatic trials, for example, tend to enroll a broader range of patients than explanatory trials, as well as to involve physicians working in a wider variety of care settings. They may also, unlike explanatory trials, allow variability in treatment protocol.

While some studies that take a more pragmatic approach to research design are RCTs, others may examine outcomes in the context of actual clinical practice, for example, examining patient records or using nonrandomized designs to follow patients who receive an intervention of interest. Thorpe et al. (2009) categorize these as pragmatic features; they have developed a tool that characterizes study designs along a number of dimensions to determine whether they are more explanatory or more pragmatic in design, or even whether those studies are RCTs.

This spectrum of research designs raises interesting epistemological and ethical questions for philosophers of medicine. For example, work in research ethics has traditionally taken it to be the case that a sharp distinction can be made between research and clinical practice, but this can no longer be assumed. For example, according to Thorpe et al. (2009), one pragmatic feature a study could have is tracking patient outcomes over the long term using healthcare records (which is something generally done in the context of clinical practice) rather than using the kind of formal follow-up that is the norm in clinical research. Philosophers of medicine may also be interested in elucidating the strengths and weaknesses of different kinds of studies, building on the body of work that examines the epistemology of EBM. Work in this area will also have implications for research ethics, as it challenges the traditional research/practice distinction that is the foundation of the Belmont Report (United States, 1978; see also Largent et al., 2011; Kass et al., 2012).

Tracking Outcomes within Specific Patient Groups

A number of other developments in clinical research are related to the second problem identified with respect to RCTs, that is, the fact that RCTs tend to look only at average outcomes in the treatment and in the control groups. We briefly survey three of these developments here: gender medicine, “basket trials,” and the US National Institutes of Mental Health’s Research Domain Criteria (RDoC). These examples come from different areas of medicine and also differ in the kind of characteristics they use to “sort” patients into relevantly similar groups. While these examples are not concerned directly with treatment outcomes, they do aim ultimately to improve health interventions by ensuring that they target only groups of patients in which there is a higher-than-average chance of benefit. One reason for this broad development in clinical research is

that most new drugs being developed have very low treatment effects (a problem that has contributed to the medical nihilism mentioned earlier); the hope is that particular groups of patients might experience better results.

Gender medicine (also known as gender-specific medicine) has developed in order to understand better the effects of sex and gender differences on health.¹ This includes “prevention, clinical signs, therapeutic approach, prognosis, psychological and social impact” (Baggio et al., 2013). As with anything that focuses on differences between women and men, gender medicine includes not only biological dimensions but also social and political dimensions. Traditionally, clinical research has been conducted mainly on men, on the grounds that women’s hormone cycles present a serious confounding factor that complicates the interpretation of data. In 1982, Rebecca Dresser pointed out that medical research was conducted almost entirely on white men, and she argued that the near exclusion of other demographic groups was both ethically and epistemologically problematic. With regard to women in particular, she noted that in addition to the resulting lack of knowledge of diseases that affect only women (e.g., uterine cancer), women and men can have physiological differences that influence both the manifestations of disease and response to treatment. Currently, in the United States, the NIH requires studies to enroll participants from different demographic groups, so that study samples are more representative of the population. There is disagreement, however, about the epistemological value of analyzing treatment outcomes in different subgroups of participants in an RCT; gender medicine takes a clear stance on this issue by recommending that potential outcome differences between women and men be examined.

Another attempt to develop more clinically useful groupings of patients comes from oncology research. “Basket trials” are designed to provide quick information about the efficacy of drugs in groups of patients who have a specific, usually rare, mutation—regardless of the histology of the tumors. These studies build on evidence that, even in trials of therapies that do not appear to be effective in treating cancer, a small number of patients may respond to the therapy, sometimes quite dramatically. These patients were found to have similar mutations, suggesting that new treatments may target specific genetic markers (Lynch et al., 2004). Trials that utilize these findings enroll patients who have a particular mutation, and then group patients with the same kind of cancer into smaller “baskets” within the bigger trial. In some cases treatments are effective across different cancer types, although in others the picture is more complicated. For example, vemurfenib has been approved for treatment of melanoma, in which a mutation called *BRAF* V600E is fairly common. The drug is not effective in treating colorectal cancer associated with this mutation, but may be effective for metastatic papillary thyroid cancer (Willyard, 2013). More recently, it has been suggested that, based on the molecular mechanisms involved, vemurfenib may be effective for treatment of colorectal cancer when

used as part of a combination therapy (Prahallad et al., 2012). This work suggests that understanding the relationship between tumor type and mutation type may be important. Although advocates believe that basket trials have the promise to revolutionize cancer treatment, it should be noted that the available evidence is still very scant.

The most radical attempt to find new ways of grouping patients is the RDoC project, which aims to replace current, symptom-based diagnostic categories—and, by extension, prognostic and treatment categories—with new diagnostic groups that are rooted in genetics, behavioral sciences, and, especially, neuroscience.² The RDoC framework consists of a matrix, in which the rows represent constructs that are “the fundamental unit of analysis” in the National Institutes of Mental Health (NIMH) framework and which describe specific dimensions of psychological functioning, such as reward learning, cognitive performance monitoring, and attachment formation and maintenance. The columns of the matrix reflect units or levels of analysis (e.g., genes, cells, neural circuits, self-reports) and the cells of the matrix contain information about which specific genes (etc.) have been shown to be relevant to the construct. The NIMH notes explicitly that this may entail that patients with different diagnoses (according to the current system), but similar functional impairments related to a construct, may end up qualifying for the same research study. Similarly, using constructs as the basic categorization may entail that among patients with the same diagnosis, only some will be eligible for a study that examines a particular construct.

The NIMH notes that the RDoC project is still in its preliminary stages, and very much open to revision (including the possible addition of new constructs). It also notes that the true test of the framework as a whole and of particular constructs is their clinical usefulness: “the critical test is how well the new molecular and neurobiological parameters predict prognosis and treatment” (Insel et al., 2010). At the same time, it is clear that the NIMH is betting that the biological approach is the best way to achieve progress in psychiatry and to overcome the limitations of the current symptom-based approach. “If we assume that the clinical syndromes based on subjective symptoms are unique and unitary disorders, we undercut the power of biology to identify illnesses linked to pathophysiology and we limit the development of more specific treatments” (National Institutes of Mental Health, RDoC website).

Initiatives like RDoC, basket trials, and gender medicine raise philosophically interesting questions about the establishment of prognostic (and, particularly in the case of RDoC, of diagnostic) groups. For example, they may inform the debates in philosophy of medicine about the role of knowledge of mechanisms in clinical research and in patient care, which may ideally

take the “fine-grained” approach to medical epistemology we have described earlier.

Diagnosis

Compared with discussions of clinical research on treatments, which has also been the primary focus of the previous sections of this chapter, there is less philosophical work on diagnosis. This represents a significant research gap, as diagnosis is of pivotal importance in medical practice, and as such is the starting point of the clinical encounter. Before treatment or prognostic evaluation of a patient can begin, there must be at least a working diagnosis. If a clinician does not begin the clinical encounter by working to obtain an accurate diagnosis, then subsequent treatments prescribed for the patient are likely to fail, and prognoses to be inaccurate. In light of the important role of diagnosis in medical practice, it might seem somewhat surprising that the philosophy of medicine literature on diagnosis is sparse. A recent survey, for example, found that of the 627 articles published over a 10-year span in the two main philosophy of medicine journals, *Journal of Medicine and Philosophy* and *Theoretical Medicine and Bioethics*, only 4 included a discussion of diagnosis (Stempsey, 2008). Addressing this research gap is another new area in which philosophers of medicine are beginning to work. There are many issues in diagnosis and diagnostics that have as yet to benefit from philosophical attention. In the next section, we note a few of them and then examine one in detail: the question of how to evaluate the medical worth of a diagnostic test.

Philosophical Questions in Diagnostic Practice

Philosophical questions in diagnostic practice can be roughly divided between those that concern diagnostic *reasoning* and those that concern *diagnostics* (tests and procedures that are used in the process of medical diagnosis). In the first category, there are questions of whether there is a logic of diagnosis, and if so, whether this logic is computable, and whether the diagnostic reasoning process is generalizable.

In the second category, which we examine more closely here, the questions concern the diagnostic tests and procedures themselves, and how they should be evaluated. In the first instance, diagnostic tests need to be evaluated for accuracy. However, even once a diagnostic test is determined to be relatively accurate, the question of whether it is valuable remains. In order to answer this question, one needs at least a working theory of what counts as valuable in the

medical context. Thus the process of determining the medical worth of a diagnostic test has both epistemic and ethical components.

Diagnostic Accuracy and Patient Outcomes

Before we can determine whether a diagnostic test is valuable, we must first determine whether the test is accurate. Currently, there is discussion in the medical literature³ about the best way to determine diagnostic accuracy. While, as we have seen, RCTs are considered by the EBM movement to be the gold standard for determining both *treatment* and *prevention* efficacy, until recently they have not been used to determine test accuracy. However, RCTs are now being used, for example, to determine whether one diagnostic test is more accurate (i.e., more sensitive and specific) than another. This can be done, for instance, to test the comparative accuracy of a new, or lesser used, diagnostic test against the currently accepted clinical reference standard (the test that is deemed to be the most reliable available test for diagnosing a given condition).⁴ For example, one might be interested in determining the comparative accuracy of duplex ultrasonography versus angiography (the clinical reference standard) for diagnosing arterial stenosis in patients presenting with cervical bruit. As the sensitivity and specificity of diagnostic tests such as these can vary across population subgroups, the two tests must be evaluated in comparable groups, one of which is randomized to receive the older test, and the other of which is randomized to receive the newer test. Or, the trial could be designed so that the same group receives both tests (Bossuyt et al., 2006).

Once a diagnostic test's degree of accuracy is determined, that is, once it is found to provide reliable information about the condition in question, it must still be determined whether the test is worth using. Even when a diagnostic is accurate, this does not guarantee that it might improve patient outcomes. In fact, the information gained from diagnostic testing alone never has a direct impact on patient outcomes (although it can have positive or negative indirect effects, see, for instance, Cournoyea and Kennedy (2014))—only diagnostic, treatment, and preventative *decisions* made subsequent to obtaining test results have this kind of impact. Thus while the result (accurate or inaccurate) of a diagnostic test might in turn lead to a decision that has an impact on the patient, the information generated by the test alone does not have this power. Some have argued that, because of this, we ought not only to be concerned with determining test accuracy, but also with finding out whether performing a test ultimately improves the lives of patients. In order to determine this, we need to know whether a given diagnostic test is a good predictor, not only of the condition it is intended to diagnose, but also of treatment outcomes for this condition. This information can, at least in theory, be determined via clinical trial

evaluation. However, just how to design a trial that reliably provides this information is not immediately straightforward. The reason for this is that, unlike in the case of treatment trials,⁵ in the case of trials of the impact of a diagnostic on patient outcomes, the evaluation is not of the test alone, but rather of the way in which the test guides treatment decisions that will in turn affect the patient. Thus, a trial of the impact of a diagnostic test on patient outcomes is really a trial of the *test plus treatment* (and perhaps also of cost, impact of the information gained on the patient, etc.). In other words, while we can test the accuracy of one diagnostic against the accuracy of another, this, in itself, does not constitute a measure of the effect of the test on a patient.

Various ways have been suggested for designing trials that might yield reliable information about the way that diagnostic tests affect patients. One central concern in designing these trials is the decision of when or where to randomize. Different quantities might be measured depending upon whether one decides to randomize before the decision is made to perform the test being evaluated (in which case it would not be possible to distinguish between treatment versus prognostic value of the test), versus after deciding to perform the test but prior to test result (in which case the test results would not need to be revealed to investigators or participants), versus after receiving the test result, and then randomizing to treatment (for instance, only those who test positive for the condition in question, in which case investigators and trial participants would know the test results). In designing a test plus treatment trial in order to determine the effect of a diagnostic on patient outcomes, one must first decide exactly what is to be measured before deciding upon when to randomize the trial.

Diagnostics and Medical Value

Related to the question of the relationship between an accurate diagnostic and patient impact is the question of whether a diagnostic test or procedure can be considered valuable when it has no direct effect on patient outcome. For example, one might ask whether an accurate diagnostic test for an untreatable disease has any medical worth.⁶ This question is, at least in part, an ethical one, and is currently debated within the medical community. While it has been established that many patients do want to know what is wrong with them, even when a treatment for their illness or condition is not available, clinicians are divided over whether an accurate test for an untreatable disease should ever be performed (Lijmer and Bossuyt, 2009). To resolve this issue depends upon how we understand value in the medical context—for example, whether we believe that value is tied inextricably to patient outcomes or whether we believe that a test that provides knowledge is valuable even when it does not lead to improved patient health. Some clinicians have explicitly argued that what

makes a test valuable *for the patient* is not its degree of correspondence with the truth—that is, they argue that diagnostic value is not in accurate diagnosis alone, but in whether that diagnosis, and subsequent treatment, can prevent the patient from suffering (Lijmer and Bossuyt, 2009). The argument is that patients undergoing a diagnostic test for arterial stenosis, for example, are not interested simply in knowing whether they have the condition, but in whether treating the condition might prevent a cerebrovascular episode.

The above issue cannot be addressed solely by an analysis of diagnostic accuracy. Instead, both the test's predictive and prognostic value must be evaluated as well. But this is an example of a condition that *is* treatable, and one might argue that while it is clear that in such a case the patient would be interested not only in a diagnosis but also in a treatment for the condition diagnosed, in other cases the patient's interests might not be as clear. Consider, for instance, the question of whether a predictive genetic test for Huntington's disease has medical worth. The test can be given to a healthy person with a family history of the disease and then allows the person to know whether he or she will develop the disease in the future. The test is very accurate; however, since there is no currently available treatment for Huntington's disease, it is arguable whether it has any medical value. On the one hand, one might argue that there is value in simply knowing one's future health fate. On the other hand, one could argue that since medicine is an applied practice that aims at improving patient health, and that, further, operates under limited resources, it would be a waste of both time and monetary resources to perform a test that will not in any way improve the health outcome of the patient. Thus, a determination of the worth of a diagnostic test is complex and depends not only on a determination of accuracy, but also on an estimation of the resulting value of performing the test, and this estimation may differ depending on how one analyzes value.

In summary, diagnostic reasoning and testing is a new area of research within the field of philosophy of medicine, and the questions to be analyzed in this area are both varied and complex, in many cases containing both epistemic and ethical, as well as theoretical and applied, components.

New Directions in Philosophy of Psychiatry

In contrast to other areas of philosophy of medicine, questions about diagnosis have been of central concern in the philosophy of psychiatry. More generally, the relationship between philosophy of psychiatry and philosophy of medicine is not as close as it perhaps should be. This state of affairs may reflect the general relationship between psychiatry and the rest of medicine. There is an ongoing debate about whether, why, and to what extent psychiatry is distinct

from medicine. Some people argue that psychiatry is conceptually no different than other branches of medicine (Guze, 1992), though they may acknowledge that our relatively poor understanding of the pathophysiology of mental illness entails that, in practice, psychiatry faces distinct challenges. Others argue that psychiatry is, and will always be, importantly different than other areas of medicine (Laing, 1985). We believe that there could be a productive exchange of ideas between philosophy of psychiatry and philosophy of medicine; however, this discussion is beyond the scope of the chapter. Here, we will focus on questions relevant to psychiatric diagnosis.

Mental illness is an urgent and growing public health problem, contributing to the global burden of disease throughout the world. Vast deficiencies in mental healthcare across the globe are matched by ongoing controversy over the nature, causes, and best treatments for disorders such as schizophrenia and depression. Mental disorder, its nature, its research, and its care have been of interest not only to philosophers who are concerned about the nature of the mind, and in the scientific explanation of complex human phenomena, such as mental disorders, but also to philosophers of medicine and ethicists who focus on issues relevant to the development of effective and ethical treatment of mental disorders (Gupta, 2014).

The goal of this section is to outline the issues concerning the scientific and ethical issues surrounding the psychiatric classification systems, as they have become the focus of increased controversy leading to and following the publication of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013). The DSM-5 offers the standard criteria for classification of mental disorders. It is designed for pragmatic use across a variety of settings to accomplish a plethora of tasks: to facilitate clinical treatment; to develop educational programs about mental illness; to provide clear criteria of eligibility for various administrative and policy related purposes, including the determination of insurance coverage and disability aid; and to further scientific research into mental disorder etiology, psychopharmacology, and forensics. One core concern that stems from the multiplicity of the purposes assigned to the DSM-5 is its questionable capacity in fully fulfilling these roles. This concern, which is getting stronger following the publication of the DSM-5, focuses on its usefulness in the clinical context during the diagnosis and treatment of mental disorders, and utility for advancing research on the etiology of mental disorders. With respect to clinical use, we highlight the problems associated with its symptom-based approach to classifying mental disorders. For instance, this feature has led to the removal of the bereavement exclusion criterion from the major depression category, allowing complicated grief to be diagnosed as depression on the grounds that grief related distress and depression are manifest through the same symptoms. We then turn to research related limitations of the DSM-5. Here we also review

RDoC, the alternative schema for psychiatric research created by the NIMH, as an alternative to the DSM-5 for research purposes.

A core concern about the fitness of the DSM-5 for clinical purposes is the descriptive approach to mental disorder classifications, in which mental disorders are individuated through symptoms and signs, as opposed to focusing on the individual's experience as a complex and multidimensional person (Sadler, 2005; Radden, 2009; Tekin, 2015; Tekin and Mosko, 2015). The symptom-based classification of mental disorders was adopted in the DSM-III (1980) and guided both the DSM-IV (1994) and the recently published DSM-5 (2013). The development of this approach was an expression of psychiatry's move toward an evidence-based scientific framework, away from the etiological approaches of the DSM-I (1952) and the DSM-II (1968). These earlier approaches relied on empirically undefended theoretical assumptions about the workings of the mind, rather than outwardly observable disease correlates. Mental disorders, also called "reactions" in these manuals, were represented in relation to the causal factors thought to underlie them (American Psychiatric Association, 1952). These causal factors were described in the framework of psychoanalysis and taken either to be a dysfunction in the brain or a general difficulty in adaptation to environmental stressors due to unresolved sexual conflicts of childhood.

In the DSM-III, a descriptive approach replaced this framework because clusters of symptoms and signs, by virtue of their observability and measurability, were thought to facilitate objective scientific research and reliable clinical diagnosis. A scientifically valid category of mental disorder requires external validators, such as symptoms and signs, not simply theories (Robins and Guze, 1970). Thus, symptom and sign clusters were resourceful constructs for scientists whose goal was to investigate better the neurological and genetic underpinnings of mental illness. The proponents of the descriptive approach have first come up with a broad list of signs and symptoms individuating mental disorders, knowing that these are only abstractions and that they do not capture the full complexity of mental disorders. The hope was that as psychiatry progressed symptoms and signs would be better delineated and more refined categories might be developed. However, this did not happen; rather the characterization of mental disorders as symptom clusters remained. The categories have departed further from the complex and real experiences of individuals with mental disorders.

One significant disadvantage of operationalizing a symptom-based approach is that the symptom clusters fail to represent certain complexities involved in mental disorder, which are neither immediately observable nor readily measurable. In a mental disorder experience, the individual's relationship with herself, her physical environment, and her social environment is strained or severed, adding many layers of complexities to mental disorders. These include the

developmental trajectory of mental disorder in the individual from childhood to adulthood; the individual's particular life history; interpersonal relationships; biological and environmental risk factors; gender, race, and socioeconomic status; the first-person-specific dimension of the symptoms, such as what the individual hears when she hears voices; and the meaning the individual ascribes to these elements of life in her sociocultural context. DSM-5 categories, by virtue of highlighting symptoms, abstract (or bracket) the self-related and context-specific aspects with mental disorders. By saying little about how the disorder experience is integrated into the patient's life, the categories are simply a "repertoire of behavior" (Radden, 2009). Such neglect of the complexity of the experiences of those with mental disorders has jeopardized the DSM-5's project as an effective tool for clinical diagnosis and treatment.

These worries are escalated with the DSM-5's removal of the bereavement exclusion criterion from the depression category, which allows an individual experiencing complicated grief to be diagnosed as having depression. The argument for this change is as follows: since there are significant overlaps between symptoms and signs of depression and the experiences of those experiencing complicated grief, and science, there is no scientific evidence for characterizing bereavement related distress and depression as distinct conditions; hence, whatever treatment helps the latter might also help the former (Zisook and Kendler, 2007; Zisook et al., 2001). Those arguing against the change insist that the cited evidence base is slim and that there are significant differences between complicated grief and depression (Horwitz and Wakefield, 2007; Wakefield and First, 2012; Kleinman, 2012; Frances, 2013; Wakefield, 2015). Concerns have also been raised about the clinical efficacy of this change, with the argument that folding complicated grief into depression does not facilitate the development of psychotherapeutic approaches to address complicated grief clinically (Tekin, 2015; Tekin and Mosko, 2015).

The DSM-5 has also not satisfied those interested in developing a psychiatric taxonomy system that advances research in psychiatry. Just before the publication of the DSM-5 in 2013, the NIMH abandoned the DSM-5 for research purposes (Insel, 2013). The argument put forward was that the DSM-5 categories are not sufficient for research purposes because they lack validity, and that a diagnostic system aiming to scrutinize mental illness should more directly reflect modern neuroscience, as "mental illness will be best understood as disorders of brain structure and function that implicate specific domains of cognition, emotion, and behavior" (Insel, 2013).

As an alternative to the DSM-5, the NIMH announced the RDoC project, which attempts to create a new conceptual framework to describe psychiatric research. RDoC brings together the resources provided by various basic sciences, including genetics and neuroscience. It lays out a model of basic psychological capacities that are believed to uncover the biological mechanisms

underlying psychopathology. Instead of organizing psychopathology into DSM-5 categories like schizophrenia and major depressive disorder, the RDoC explicates psychopathology in terms of basic psychological processes (e.g., declarative memory, perception) and their underlying biological mechanisms.

Insel (2013) provides three fundamental tenets of the RDoC. First, RDoC is a diagnostic approach based on biology, not on observable signs and symptoms. Second, it takes mental disorders to be biological disorders involving brain circuitry that implicate the specific domains of cognition, emotion, or behavior. It is expected that scientists can better identify and investigate the circuits implicated in mental illness as neuroscience and genetics advance. Third, the mapping of the cognitive, circuit, and genetic aspects of mental disorders can yield new and better targets for treatment.

The success of RDoC as a useful guide for research remains to be determined, as the NIMH task force works to complete the project. Skeptics question the ability of RDoC to rescue psychiatry from its crisis, suggesting that it would fail to increase validity as it is grounded on assumptions about how the brain works rather than on actual scientific facts about the complex mechanisms by which the brain operates (Hoffman and Zachar, in press). Some critics have suggested that the primacy of neuroscientific and genetic research in psychopathology continues an unfortunate trend that ignores the crucial role of the phenomenology of mental illness, and this may have negative implications for treatment (Graham and Flanagan, 2013). Finally, some critics worry that the developers of the RDoC are making the same mistake as those who were instrumental in developing the symptom-based criteria for mental disorders. As discussed earlier, the proponents of the descriptive approach characterized mental disorders through a list of signs and symptoms, knowing that these are only abstractions and that they do not capture the full complexity of mental disorders. However, as the DSM project evolved, they dropped the recognition that such characterizations are abstractions, leaving the categories of mental disorders further away from the true complexities of mental disorders. Similarly, the worry is that the RDoC's proposed molecular and neurobiological mental disorder parameters, which are expected to be better identified and investigated as neuroscience and genetics develop, may be uncritically accepted as the true targets of research (Bluhm, in press), leading to disorientation in the field of psychiatric research.

Objectivity and Medical Research

A final area that we want to highlight as an important new direction in philosophy of medicine is that of the objectivity of medical research. We noted earlier that philosophy of medicine is beginning to take account of the relationship

between medical research and the broader social context within which it occurs. The social context of medical research, in part, helps to constitute the objectivity of medical research and, in part, serves to threaten that very objectivity.

As a discipline whose results often bear direct social relevance, medical research attracts considerable interest among the general public. For instance, 58 percent of Americans say they are interested in new medical discoveries (National Science Foundation, 2014) and, according to the Wellcome trust monitor report (2013), 75 percent of adults in the United Kingdom reported being curious about medical research. Media reports on new cures, clinical guidelines, and possible health threats may influence the behavior of the members of the public. Scandals such as the ones related to the link between the painkiller Vioxx and cardiovascular events (Biddle, 2007) and selective serotonin reuptake inhibitors and suicidal ideation (Healy, 2012) have raised doubts about the trustworthiness of medical research and the integrity of scientists working in the field. Similarly, the recent outbreaks of vaccine-preventable diseases—such as measles, whooping cough, and polio—have been associated with the rise of the antivaccine movement and the distrust of established medical expertise (Poland and Jacobson, 2011). The public's reactions to these events show that one of the most important current challenges for the community of medical scientists is securing the public's trust in its research. One of the conditions for maintaining trust in science is that research is conducted in a way that is thought to be objective. Thus, searching for the means that best support objectivity is a central assignment for medical scientists and philosophers of medicine.

Commercial interests have been associated with many of the scandals that have threatened the apparent trustworthiness of medical science, since industry-funded drug trials report favorable results for company products when compared to trials funded by independent agencies (Lundh et al., 2012). Furthermore, reports of secrecy and dubious practices in medical research, particularly in the development of new drugs, have started to emerge (Sismondo, 2008). Biomedical research is a highly commercialized field. In 2012, 58 percent of research in the United States was funded by private sources (Moses et al., 2015). Consequently, critics have argued that the objectivity of the discipline is being compromised. Yet, it is not self-evident what actually is called for when objectivity of research is demanded.

Traditionally the objectivity of research has been associated with the integrity of individual scientists: objectivity is a trait that researchers need to cultivate in themselves.⁷ According to this understanding, researchers should be on guard against their own biases and withdraw from assignments that might involve conflicts of interests that could undermine their impartiality (Shamoo and Resnik, 2009). However, studies on implicit biases have shown that individuals are not very good at recognizing their own biases (Uhlman and Cohen, 2007). Furthermore, even when individuals do their best not to let

their preferences influence their reasoning, human actions tend to be affected by extraneous factors. For example, even small gifts can have an impact on the prescribing practices of physicians (Katz et al., 2003). Thus, it seems that relying on individuals' integrity is not enough to ensure that research results are unbiased. Because of this, when the grounds for conducting objective medical research are investigated, the perspective has to be shifted from evaluating the attitudes of individuals to examining practices and methods that would best ensure unbiased outcomes.

Which methods best promote the objectivity of research have been a debated issue in philosophy of science. The controversies are partly rooted in the intrinsic complexity of the concept of objectivity. As recent historical and philosophical analyses have suggested, objectivity is a multifaceted concept, and its different meanings can be used for promoting diverse practices. For example, Porter (1995) and Daston and Galison (2010) have traced the transformations of the concept and its different meanings through several centuries. According to these scholars, different virtues have been attached to objectivity. Daston and Galison (2010), for example, describe how the ideal of mechanical objectivity encouraged atlas makers in the late nineteenth and early twentieth centuries to find ways of depicting nature without the interference of human judgment and interpretation, even if it happened at the expense of describing the details of the object of interest. They detail how bacterial cultures and other objects could be illustrated using photography in a way that apparently had not been affected by subjective interpretations, but earlier methods of depicting nature, such as the drawings of expert artists, were better at portraying spatial depth and color. Thus, the photographs had less diagnostic utility than drawings, despite the fact that the latter were always influenced by the subjectivity of the artists. Similarly, Porter (1995) shows how in the discipline of accounting, portraying the object of interest in a quantified form has been used as a means of guarding the line of business from outsiders' accusations of corruption. What is common in these descriptions of the use of the term "objectivity" is that pursuing certain practices that are labeled objective or that ensure objectivity is a way of trying to build trust—both between the members of the communities following these practices and among outsiders.

Achieving objectivity is thought to require removing detrimental subjectivity, and the biases associated with it, from the process, and thereby improving the trustworthiness of results (cf. Daston and Galison, 2010, pp. 373–74). According to Douglas (2004, p. 454), the "implicit call to trust" is still common in the diverse ways in which the term "objective" is used today. She has specified eight different senses of objectivity that can be used to refer to outcomes of different kinds of processes. Due to the complexity of the term, it is possible to praise a method for producing objective results or denounce it as biased—depending on which of the senses of objectivity is chosen as the ideal to follow.

Because of the rhetorical force behind stating something to be objective, these choices can influence methodologies and have bearings on philosophical discussions concerning medical research.

Douglas's analysis can be used to evaluate the discussions concerning the objectivity of medical research as it elucidates the fact that not all senses of objectivity are applicable in every context. For example, one way of trying to constrain subjectivity and avoid biased results is to aim at making individual judgments redundant by establishing guidelines and procedures that give detailed instructions on how to carry out research. Douglas (2004, pp. 461–62) calls this ideal "procedural objectivity." This understanding of objectivity has led to discussions over the use of meta-analyses in amalgamating evidence. Stegenga (2011) has argued that meta-analysis falls short of being the platinum standard of evidence: despite the communally accepted guidelines, performing a meta-analysis necessarily involves numerous judgments on the analyst's part, which, in turn, opens the door for individual biases and subjective preferences and, thus, mitigates the method's objectivity. Stegenga's analysis demonstrates why following the ideal of disposing of the need for individual judgments fails in a context where the reliability of medical knowledge is sought: it is practically impossible.

In addition, focusing on the way in which studies and analyses are conducted, that is, on the "internal" processes of scientific inquiry, overlooks an important area that influences science, namely, the impact an institutional context of research has on knowledge production. For example, Young et al. (2008) worry that medical science may be biased by the current publication practices, which encourage researchers to pursue projects that are likely to be accepted by the most prestigious journals. Publication bias—studies showing positive results are published more often than studies with negative results—is perhaps the most significant bias affecting medical publishing, which systematically biases the literature (Godlee and Dickersin, 2003). Brown (2010), in turn, has argued that the funding structure in the field of biomedicine, that is, the prevalence of industry funding and the pressure to produce commercially applicable results, is creating lacunae in published literature. If the available funding guides research toward searching for only certain types of explanations for phenomena and the interests that motivate planning research in this way are not in line with generally accepted goals of research, then it might be the case that the outcomes of research regarded in their entirety have been biased. In other words, medical research is not objective. For instance, if funding is available only for those studies on mental illnesses that may produce patentable outcomes, important features of the phenomena may be left unstudied (Musschenga et al., 2010).

According to critics who urge taking notice of practices related to the allocation of funding and dissemination of results, focusing on individual studies and

assessing only their internal quality leaves invisible certain features of research activities that may systematically bias scientific knowledge. In the context of medical research, this is particularly worrisome. According to the principles of EBM, the results of meta-analyses and other systematic reviews trump individual studies, and treatment guidelines are usually based on amalgamated evidence (Howick, 2011b). Because of this, the objectivity of medical research should be evaluated by using a concept of objectivity that accommodates the possibility of systematic biases that are caused by factors not belonging to the internal stages of research.

According to a traditional understanding, practices that take place in the context of discovery, that is, when research questions and hypotheses are developed, do not have an influence on the objectivity of knowledge production because through rigorous testing any possible biases can be removed in the so-called context of justification. In the light of recent empirical studies on the pharmaceutical industry (Sismondo, 2008) and philosophical analyses (Brown, 2010), however, this assumption should be questioned. Contrary to the traditional understanding of the distinction between the contexts of discovery and justification, testing of claims may not weed out all biases. The objectivity and trustworthiness of medical research can be severely compromised because of publication bias or problematic practices that occur in the phases of the allocation of resources, which fall on the discovery side of the discovery-justification distinction. If the way of framing research questions and projects results in a skewed understanding of the object of interest, it is possible to talk about the violation of objectivity (Brown, 2010). Consequently, the objectivity of medical research has to be examined from a perspective that takes notice of a wider spectrum of practices and factors than is traditionally considered.

A viable candidate for a new perspective on the aforementioned issues comes from social epistemology. Social epistemological theories have highlighted the importance that the institutional context of inquiry has for the reliability of produced information. For example, Longino (1990, 2002) has argued that the objectivity of research is dependent on the institutional and social context in which inquiry takes place and because of this, instead of exclusively evaluating individual scientists or single studies, those who are interested in addressing biases should pay attention to the organizational structure of science. Likewise, Biddle (2007) has argued that epistemic problems of biomedical research should be addressed with institutional arrangements.

In sum, different understandings of objectivity direct our attention to assessing diverse features of scientific practice. In the light of recent analyses that have disclosed problematic practices related to current biomedical research, it seems that those ideals of objectivity that are applicable for assessing individual scientists or particular methods for testing hypotheses are insufficient

for capturing which factors may bias the eventual results. Because of this, it is essential to develop analyses of objectivity further that examine how conditions for this virtue are constituted at the level of research communities.

Conclusion

Philosophy of medicine has developed greatly over the past decade or so, with a new focus on epistemological questions (and their relationship with ethical questions). This chapter has identified several areas in which new and interesting philosophical questions are being addressed, and it has described how this research is beginning to build on and expand philosophical research on medicine, medical research, and the social context in which these activities are conducted.

Notes

1. See also chapter 10 of this volume for a discussion of gender medicine.
2. Here we discuss RDoC further, placing it in the larger context of discussions in philosophy of psychiatry. We focus only on RDoC as an attempt to develop ways of categorizing patients, which are intended to result in improved care.
3. See, for example, Bossuyt et al. (2000, 2006); and Ferrante di Rufano et al. (2012).
4. As with the debate regarding the epistemic merits of RCTs with respect to assessing the efficacy of therapeutic interventions, the requirement that comparative accuracy of diagnostic tests be assessed with RCTs is debatable—though we do not articulate the debate here.
5. RCTs of therapeutic treatments are subject to well-known epistemic and ethical problems of their own. However, as these issues have been elsewhere addressed in the philosophy of medicine literature, we will not discuss them here.
6. In such cases, a test might have a negative effect on patient outcomes, perhaps due to the intervention of the test, the cost, the anxiety from information gained, and so on. On the other hand, it might have a positive effect. Some studies (Lijmer and Bossuyt, 2009, for example) show that patients want to know what is wrong with them even when there isn't the possibility of doing anything about it.
7. See, for example, Smith (2004) for a defense of a view according to which the objectivity of science can be reduced to individuals.

References

- American Psychiatric Association. (1952) *Diagnostic and Statistical Manual: Mental Disorders*, 1st edition. Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (1968) *Diagnostic and Statistical Manual of Mental Disorders*, 2nd edition. Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (1980) *Diagnostic and Statistical Manual: Mental Disorders*, 3rd edition. Washington, DC: American Psychiatric Association.

- American Psychiatric Association. (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition. Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013), *Diagnostic and Statistical Manual: Mental Disorders*, 5th edition. Washington, DC: American Psychiatric Association.
- Angell, M. (2004), *The Truth about the Drug Companies: How They Deceive Us and What to Do about It*. New York: Random House.
- Baggio, G., Corsini, A., Floreani, A., Giannini, S., and Zagonel, V. (2013). Gender medicine: A task for the third millennium. *Clinical Chemistry and Laboratory Medicine* 51(4): 713–27.
- Biddle, J. (2007). Lessons from the Vioxx debacle: What the privatization of science can teach us about social epistemology. *Social Epistemology* 21(1): 21–39.
- Bluhm, R. (2005). From hierarchy to network: A richer view of evidence for evidence-based medicine. *Perspectives in Biology and Medicine* 48(4): 535–47. doi:10.1353/pbm.2005.0082.
- Bluhm, R. (in press). Evidence-based medicine, biological psychiatry, and the role of science in medicine. In *Extraordinary Science: Responding to the Current Crisis in Psychiatric Research*, eds. Jeffrey Poland and Şerife Tekin. Cambridge, MA: MIT University Press.
- Borgerson, K. (2009). Valuing evidence: Bias and the evidence hierarchy of evidence-based medicine. *Perspectives in Biology and Medicine* 52(2): 218–33. doi:10.1353/pbm.0.0086.
- Bossuyt, P. M., Irwig, L., Craig, J., and Glasziou, P. (2006). Comparative accuracy: Assessing new tests against existing diagnostic pathways. *British Medical Journal* 332: 389–92.
- Bossuyt, P., Lijmer, J., and Mol, B. (2000). Randomized comparisons of medical tests: Sometimes invalid, not always efficient. *Lancet* 356: 1844–47.
- Brown, J. (2008). The community of science®. In *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, eds. M. Carrier, D. Howard, and J. Kourany. Pittsburgh: University of Pittsburgh Press.
- Brown, J. (2010). One-shot science. In *The Commercialization of Academic Research*, ed. H. Radder. Pittsburgh: Pittsburgh University Press, pp. 90–109.
- Brozek, J. L. Akl. E. A., Jaeschke, R., Lang D. M., Bossuyt, P., Glasziou, P., Helfland, M., Ueffing, E., Alonso-Coello, P., Meerpohl, J., Phillips, B., Horvath, A. R., Bousquet, J., Guyatt, G. H., Schünemann, H. J., and GRADE Working Group. (2009). The GRADE approach to grading quality of evidence about diagnostic tests and strategies. *Allergy* 64: 1109–16.
- Cournoyea, M., and Kennedy, A. (2014). Causal explanatory pluralism and medically unexplained physical symptoms. *Journal of Evaluation in Clinical Practice* 20(6): 928–33.
- Daston, L., and Galison, P. (2010). *Objectivity*. New York: Zone Books.
- De Melo-Martin, I., and Intemann, K. (2011). Feminist resources for biomedical research: Lessons from the HPV vaccines. *Hypatia* 26(1): 79–101. doi:10.1111/j.1527-2001.2010.01144.x.
- Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science* 67(4): 559–79.
- Douglas, H. (2004). The irreducible complexity of objectivity. *Synthese* 138: 453–73.
- Dresser, R. (1982). Wanted: Single, white male for medical research. *Hastings Center Report*, 22(1): 24–29.
- Elliott, C. (2010). *White Coat, Black Hat: Adventures on the Dark Side of Medicine*. Boston: Beacon Press.
- Ferrante di Rufano L., Hyde, C. J., McCaffrey, K. J., Bossuyt, P. M. M., and Deeks, J. J. (2012). Assessing the value of diagnostic tests: A framework for designing and evaluating trials. *British Medical Journal* 344.
- Frances, A. (2013). Last plea to DSM-5: Save grief from the drug companies. *Huffington Post*, January 7.

- Godlee, F., and Dickersin, K. (2003). Bias, subjectivity, chance, and conflict of interest in editorial decisions. In *Peer review in Health Sciences*, 2nd edition, eds. F. Godlee and T. Jefferson. London: BMJ Publishing, pp. 91–117.
- Goldacre, B. (2012). *Bad Pharma: How Drug Companies Mislead Doctors and Harm Patients*. New York: HarperCollins.
- Gøtzsche, P. (2013). *Deadly Medicines and Organized Crime*. London: Radcliffe Publishing.
- Graham, G., and Flanagan, O. (2013). Psychiatry and the brain. Oxford University Press blog; August 1. <https://blog.oup.com/2013/08/psychiatry-brain-dsm-5-rdoc/>.
- Gupta, M. (2014). *Is Evidence-Based Psychiatry Ethical?* Oxford: Oxford University Press.
- Guze, S. B. (1992). *Why Psychiatry Is a Branch of Medicine*. New York: Oxford University Press.
- Healy, D. (2012). *Pharmageddon*. Berkeley: University of California Press.
- Hoffman, G., and Zachar, P. (in press). RDoC's metaphysical assumptions: Problems and promises. In *Extraordinary Science: Responding to the Current Crisis in Psychiatric Research*, eds. Jeffrey Poland and Şerife Tekin. Cambridge MA: MIT University Press.
- Horwitz, A. V., and Wakefield, J. C. (2007). *The Loss of Sadness: How Psychiatry Transformed Normal Sadness into Depressive Disorder*. New York: Oxford University Press.
- Howick, J. (2011a). Exposing the vanities—and a qualified defense—of mechanistic reasoning in health care decision making. *Philosophy of Science* 78(5): 926–40.
- Howick, J. (2011b). *The Philosophy of Evidence-Based Medicine*. Chichester UK: Wiley-Blackwell, BMJ Books.
- Insel, T. (2013). Director's Blog: Transforming Diagnosis. <http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml>.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., and Wang, P. (2010). Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry* 167(7): 748–51.
- Institute of Medicine. (2007). IOM roundtable on evidence-based medicine. In *The Learning Healthcare System: Workshop Summary*, eds. L. Olsen, D. Aisner, and J. M. McGinnis. Washington, DC: National Academies Press.
- Jukola, S. (2015). Meta-analysis, ideals of objectivity, and the reliability of medical knowledge. *Science and Technology Studies* 28(3): 101–120.
- Kass, N. E., Faden, R. R., Goodman, S. N., Provenost, P., Tunis, S., and Beauchamp, T. L. (2013). The research-treatment distinction: A problematic approach for determining which activities should have ethical oversight. *Hastings Center Report* 43(1): S4–S15.
- Katz, D., Caplan, A., and Mertz, J. (2003). All gifts large and small. Toward an understanding of the ethics of pharmaceutical industry gift-giving. *American Journal of Bioethics* 3(3): 39–46.
- Kennedy, A. G. (2013). Differential diagnosis and the suspension of judgment. *Journal of Medicine and Philosophy* 38(5): 487–500. doi: 10.1093/jmp/jht043.
- Kleinman, A. (2012). The art of medicine: Culture, bereavement, psychiatry. *Lancet* 379(9816): 608–609.
- Laing, R. D. (1985). *Wisdom, Madness and Folly: The Making of a Psychiatrist 1927–1957*. London: Macmillan.
- Largent, E. A., Joffe, S., and Miller, F. G. (2011). Can research and care be ethically integrated? *Hastings Center Report* 41(4): 37–46.
- Lijmer, J., and Bossuyt, P. (2009). Various randomized designs can be used to evaluate medical tests. *Journal of Clinical Epidemiology* 62: 364–73.
- Longino, H. (1990). *Science as Social Knowledge*. Princeton: Princeton University Press.
- Longino, H. (2002). *The Fate of Knowledge*. Princeton: Princeton University Press.
- Lundh, A., Sismondo, S., Lexchin, J., Busuioc, O. A., and Bero, L. (2012). Industry sponsorship and research outcome, Cochrane Database of Systematic Reviews 2012, Issue 12. Art. No: MR000033. doi: 10.1002/14651858.MR000033.pub2.

- Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J., and Haber, D. A. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *NEJM* 350: 2129–39.
- McClimans, L. (2013). The role of measurement in establishing evidence. *Journal of Medicine and Philosophy* 38(5): 520–38.
- Moses, H., III, Matheson, D. M. Cairns-Smith, S, George, B. P., Palisch, C., and Dorsey, E. (2015). The anatomy of medical research: US and international comparisons. *JAMA* 313(2): 174–89. doi:10.1001/jama.2014.15939.
- Moynihan, R., and Cassels, A. (2005). *Selling Sickness: How the World's Biggest Pharmaceutical Companies Are Turning Us All into Patients*. Vancouver: Greystone Books.
- Musschenga, A., van der Steen, W., and Ho, V. (2010). The business of drug research: A mixed blessing. In *The Commodification of Academic Science*, ed. H. Radder. Pittsburgh: Pittsburgh University Press, pp. 110–31.
- National Institutes of Mental Health. NIMH Research Domain Criteria (RDoC), available at <http://www.nimh.nih.gov/research-priorities/rdoc/nimh-research-domain-criteria-rdoc.shtml> (accessed March 12, 2015).
- National Science Foundation. (2014). *Science and Engineering Indicators 2014*. Arlington: VA (NSB 14-01).
- Osimani, B. (forthcoming). Hunting side effects and explaining them: Should we reverse evidence hierarchies upside down? *Topoi*.
- Poland, G., and Jacobson, R. (2011). The age-old struggle against antivaccinationists. *New England Journal of Medicine* 364(2): 97–99.
- Porter, T. (1995). *Trust in Numbers*. Princeton: Princeton University Press.
- Prahallad, A., Sun, C., Huang, S., Di Nicolantonio, F., Salazar, R., Zecchin, D., Beijersbergen, R. L., Bardelli, A., and Bernards, R. (2012). Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* 483(7387): 100–103.
- Radden, J. (2009). *Moody Minds Distempered: Essays on Melancholy and Depression*. New York: Oxford University Press.
- Reiss, J. (2010). In favour of a Millian proposal to reform biomedical research. *Synthese* 177(3): 427–47. doi: 10.1007/s11229-010-9790-7.
- Robins, E., and Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry* 126(7): 983–87.
- Russo, F., and Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science* 21: 157–70.
- Sadler, J. (2005). *Values and Psychiatric Diagnosis*. Oxford: Oxford University Press.
- Shamoo, A., and Resnik, D. (2009). *Responsible Conduct of Research*, 2nd edition. Oxford: Oxford University Press.
- Sismondo, S. (2008). How pharmaceutical industry funding affects trial outcomes: Causal structures and responses. *Social Science and Medicine* 66(9): 1909–14.
- Smith, T. (2004). “Social” objectivity and the objectivity of values. In *Science, Values, and Objectivity*, eds. P. Machamer, and G. Wolters. Pittsburgh: Pittsburgh University Press, pp. 143–71.
- Steel, D. (2007). *Across the Boundaries: Extrapolation in Biology and the Social Sciences*. New York: Oxford University Press.
- Stegenga, J. (2011). Is meta-analysis the platinum standard? *Studies in the History and Philosophy of Biological and Biomedical Sciences* 42: 497–507.
- Stegenga, J. (2014). Down with the hierarchies. *Topoi* 33(2): 313–22. doi:10.1007/s11245-013-9189-4.
- Stegenga, J. (forthcoming). *Medical Nihilism*. Oxford: Oxford University Press.

- Stempsey, W. (2008) Philosophy of medicine is what philosophers of medicine do. *Perspectives in Biology and Medicine* 51(3): 379–91.
- Tekin, Ş. (2014). Psychiatric taxonomy: At the crossroads of science and ethics. *Journal of Medical Ethics* 40(8): 513–14. doi:10.1136/medethics-2014-102339.
- Tekin, Ş. (2015). Against hyponarrating grief: Incompatible research and treatment interests in the DSM-5. In *The Psychiatric Babel: Assessing the DSM-5*, eds. P. Singy and S. Demazeux.
- Tekin, Ş., and Mosko, M. (2015). Hyponarrativity and context-specific limitations of the DSM-5. *Public Affairs Quarterly* 29(1): 109–45.
- Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G., Tunis, S., Bergel, E., Harvey, I. Magid, D. J., and Chalkidou, K. (2009). A pragmatic-explanatory continuum indicator summary (PRECIS): A tool to help trial designers. *Journal of Clinical Epidemiology* 62(5): 646–75.
- Uhlmann, E., and Cohen, G. (2007). “I think therefore it’s true”: Effects of self-perceived objectivity on hiring discrimination. *Organizational Behavior and Human Decision Processes* 104: 207–23.
- United States. (1978). *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Bethesda, MD: The Commission.
- Wakefield, J. C. (2015). The loss of grief: Science and pseudoscience in the debate over DSM-5’s elimination of the bereavement exclusion. In *The Psychiatric Babel: Assessing the DSM-5*, eds. P. Singy and S. Demazeux. *History, Philosophy and the Theory of the Life Sciences Series*. Springer Press.
- Wakefield, J. C., and First, M. B. (2012). Validity of the bereavement exclusion to major depression: Does the empirical evidence support the proposal to eliminate the exclusion in DSM-5?, *World Psychiatry* 11: 3–10.
- Wellcome trust. (2013). *Welcome trust monitor wave 2. Tracking public opinions science, biomedical research and science education*. Retrieved from http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_grants/documents/web_document/wtp053113.pdf.
- Willyard, C. (2013). “Basket studies” will hold intricate data for cancer drug approvals. *Nature Medicine* 19(6): 655.
- Worrall, J. (2002). What evidence in evidence-based medicine? *Philosophy of Science* 69: S316–S330.
- Worrall, J. (2007). Why there’s no cause to randomize. *British Journal for the Philosophy of Science* 58(3): 451–88.
- Young, N., Ioannidis, J., and Al-Ubaydli, O. (2008). Why current publication practices may distort science. *PloS Medicine* 5(10): e201. doi:10.1371/journal.pmed.0050201.
- Zisook, S., and Kendler, K. S. (2007). Is bereavement-related depression different than non- bereavement-related depression? *Psychological Medicine* 37: 779–94.
- Zisook, S., Shuchter, S. R., Pedrelli, P., Sable, J., and Deaciuc, S. C. (2001). Bupropion sustained release for bereavement: Results of an open trial. *Journal of Clinical Psychiatry* 62: 227–30.

