

Meta-Analysis, Ideals of Objectivity, and the Reliability of Medical Knowledge

Saana Jukola

This paper focuses on the ideals of scientific objectivity as they emerge in discussions concerning meta-analyses and medical research. Stegenga (2011) has argued that meta-analyses fail to be objective because conducting them involves making judgments. I show that his reasoning is based on the so-called procedural ideal of objectivity, which can be questioned: this ideal is unattainable and does not capture some of the problematic issues of medical research. By introducing a case in research on selective serotonin reuptake inhibitors, I demonstrate why the so-called social view on objectivity succeeds better in accommodating 1) the way in which scientific research necessarily involves judgments, 2) the possible risks involved in research, and 3) the influence that the institutional context has on research activities. Adopting this ideal of objectivity helps us better appreciate the virtues of meta-analyses and pinpoint which practices threaten the reliability of meta-analyses' results.

Keywords: objectivity, meta-analysis, medical research

Introduction

In this paper, I shall provide a critical study of the ideas concerning objectivity, bias and reliability¹ that emerge in the discussions concerning meta-analyses and medical research in general. Meta-analysis is a method of synthesizing information from two or more studies by using statistical techniques. The synthesis of evidence is done by collecting a number of primary studies that satisfy the pre-specified inclusion criteria, measuring the effect of interest of each study, and then combining them to produce an overall measure for the studies. (Moore & McCabe, 2006: 598.) In evidence-based medicine (EBM) and policy, systematic reviews - often including

meta-analyses² - are placed on the top of the evidence hierarchies, which represent the assumed strength of different types of evidence. Meta-analyses are thought to provide more precise information on the effects of treatments than individual studies (Cochrane Collaboration, 2015: 1.2.2.³). They are meant to amalgamate evidence in a less biased way than other means of synthesizing studies: while a researcher conducting a qualitative review has to make judgments on the relevance of individual studies and formulate the summary based on her account on the chosen material, the formal rules of meta-analysis are supposed to ensure the objectivity of the process. Thus, results of meta-analyses are considered to be a reliable source of information

for forming policy recommendations and treatment guidelines. Here, the inherent assumption is that knowledge is best produced by following rules that enable “excluding individual [...] judgments from the process” (Douglas, 2004: 461).

This article has two aims. First, I argue that the discussions on the strengths of meta-analyses are guided by the so-called procedural ideal of objectivity. I approach this issue by addressing an article by Jacob Stegenga (2011) who argues that meta-analysis’s status as “the platinum standard of evidence” is not warranted since even in meta-analysis, the analyst must make diverse judgments, which might hamper the objectivity of the process. Second, I shall show how the procedural ideal may not capture some of the central issues that the ideal of objectivity should take into consideration in order to successfully guide scientific practices. The argument being that even if meta-analyses did satisfy the conditions of being objective in the sense of excluding the need for judgments, some biases detrimental to medical research could not be removed. In this way, I demonstrate why we need a contextual ideal of objectivity for evaluating the production of medical knowledge.

The use of the concept “objective” is eminently complicated, as recent philosophical (e.g., Douglas, 2004) and historical (e.g., Daston & Galison, 2010) analyses demonstrate. Individuals, modes of inquiry, as well as the outcomes of processes can be called objective. In general, the objectivity of results is thought to be a consequence of the method being objective (cf. Longino, 1990: 62–63). In this paper, unless otherwise noted, I discuss the objectivity of methods: I posit that we call a method objective if it produces results that do not unduly reflect the preferences of actors involved in the inquiry (Wilholt, 2009; Jukola, 2014). Which methods

best ensure this, is what debates on the objectivity of science tend to centre on. It is important to note that objectivity is an ideal for research. What this means is that even though objectivity could not be fully achieved, it is still worth striving for (e.g., Resnik, 2007: 52). The fact that a discussion about possible biases influencing the results and the limitations of the used methods is a central part of many scientific articles illustrates that the difficulty of achieving objectivity is widely acknowledged in scientific practice.

Objectivity is prized because the objectivity of methods is seen to be necessary for achieving the goals of science, both practical (e.g., predicting and controlling natural phenomena) and epistemological (e.g., explanation and emancipation from flawed beliefs). In this paper, the focus is on the practical consequences of different understandings of what kinds of practices are objective. When research is conducted in a way that we denounce as not being objective, we assume that it is not rational to base practical decisions on its outcomes. Ziman’s (2000: 157–161) discussion of the reliability of research captures this understanding: science is not only expected to produce true results but results that can be utilized successfully. In medical research this is particularly apparent as the results of studies are used for informing practical treatment guidelines.

In this way, discussions concerning objectivity are significant also from a non-academic perspective. Research results can have significant implications for the lives of the public. Because of this, maintaining the public’s trust in research is an important reason for striving for objectivity (e.g., Resnik, 2007: 57). Trust in science is partly based on the assumption that research results are not biased. In the case of medicine, eroding trust may have serious consequences if it results in behaviour

such as not vaccinating children against diseases such as measles (cf. Poland & Jacobson, 2011). Consequently, searching for ways and means of weeding out biases and cultivating practices that contribute to objectivity is a central task for both scientists and philosophers of science.

Historical studies of objectivity have highlighted the complexity of the concept. In this article, I do not aim at contributing to historical examinations of objectivity. However, these discussions help one to perceive how different ways of using the concept have justified the preferences for different methods, which in turn possess different epistemic virtues. For example, Theodore Porter (1992) has investigated the history of the ideal according to which following rigid rules helps to secure trust⁴. According to Porter, following guidelines and aiming at numerical description of reality were seen as ways of building trust by giving the impression of the absence of judgments. This was the case especially in those fields that were under suspicion: when members of the group committed to following common rules, they appeared more impartial than what had been the case if the actions had been based on their own judgments (Porter, 1992: 639). Following this ideal that secured apparent impartiality came with a price: as Porter (1992: 645) remarks,

quantification is a powerful agency of standardization because it imposes some order on hazy thinking, but this depends on the licence it provides to leave out much of what is difficult of obscure.

In the same way, in their study of scientific image making, Daston and Galison (2010: 179) describe how some diagnostic utility was lost when photographs replaced drawings made by expert artists. Even though

photographs could depict their object apparently without the involvement of human judgment, they could not illustrate colour or spatial depth. Aiming at objectivity - understood as independence from human judgment - did not secure the best possible employability.

In this article, I shall demonstrate that in a similar way aiming at procedural objectivity in the production of medical knowledge may omit some salient features of the evaluated practice. When we look for the conditions that a research process should fulfill in order to be objective, we should not focus on the researchers only. We also need to consider the institutional context of research. This is something that the concept we adopt should be capable of recognizing. Adopting this kind of a contextual view on objectivity as opposed to a merely procedural one can help us see under which conditions meta-analyses serve as a means of producing unbiased knowledge, and, correspondingly, which practices may undermine objectivity.

I shall begin by briefly introducing the basic idea of meta-analysis and the motive for using the method. In section 3, I present the argument that Stegenga (2011) gives against the high evidential status of the method. In these two sections, I show how the ideal of procedural objectivity seems to underlie both the arguments for using meta-analyses and Stegenga's criticism of the method. Section 4 focuses on evaluating the success of the procedural ideal of objectivity in the context of medical research. By introducing a case in selective serotonin reuptake inhibitors (SSRIs) in subsections 4.1. and 4.2., I show that the ideal does not capture some of the factors that may cause detrimental biases in research. This case serves as a test for appraising how different ideals of objectivity succeed in evaluating current research practices. In this way, the conceptual analysis typical of philosophical

investigations is reinforced with an analysis of actual practices of science. In section 5, I first argue that the conception of objectivity that is adopted should be capable of accommodating a wider spectrum of practices than the procedural view does. I present the so-called social view on objectivity and then argue that the pertinent features of medical research are better captured by this view on objectivity, which 1) offers tools for discerning between judgments that can be taken as acceptable and unacceptable, 2) takes notice of the way in which risks, both epistemic and non-epistemic, are involved in the process of producing knowledge, and 3) pays attention to the way in which the institutional context of research can either improve or hinder objectivity.

Meta-Analysis — What Is It and Why Is It Used?

Meta-analysis was introduced by Gene Glass in 1976 and it is used for bringing some order to acquired results in those fields of research where there are many studies looking into the same topics. The numerous primary studies may result in diverse and even contradictory outcomes while producing huge amounts of data. In social and medical sciences, especially, the information resulting from research is in demand, for policy makers and practitioners need it in order to support decision making and practice. However, because of the volume of research conducted, decision makers cannot explore all available evidence on a given topic. Thus, summaries of research are needed. (Shercliffe et al., 2009: 413–414; Stegenga, 2011: 498.)

The basic principle of meta-analysis is simple: calculating a weighted average of a measured effect of interest. For instance, researchers might be interested

in knowing if taking certain medication improves subjects' Hamilton Depression Rating Scale scores. In this case, the meta-analysis would involve measuring the mean difference in the pre and post intervention scores. Stegenga outlines the four steps of conducting meta-analysis as follows:

Meta-analysis is performed by (i) selecting which primary studies are to be included in the meta-analysis, (ii) calculating the magnitude of the effect due to a purported cause for each study, (iii) assigning a weight to each study, which is often determined by the size and the quality of the study, and then (iv) calculating a weighted average of the effect magnitudes. (Stegenga, 2011: 498.)

As a result of pooling evidence from multiple sources, coherent patterns of interest can be established.

Systematic means of synthesizing evidence are thought to minimize the possibility of subjective biases entering the process, and thus provide more reliable knowledge that could be used to support decision making (Cochrane collaboration, 2015: 1.2.2.). Meta-analysis is hailed as superior since it is considered to make subjectivity redundant in amalgamating evidence by supplying a systematic and explicit method. This conception of securing the reliability of research is related to an ideal which Heather Douglas (2004) calls *procedural objectivity*⁵. The essence of this ideal is that there is a process that “allow[s] for individual interchangeability and exclude[s] individual idiosyncrasies or judgments from processes” (Douglas, 2004: 461). If the outcome of a process is the same regardless of the preferences of the person conducting the analysis, the method is objective and the ensuing results are suitable for guiding our actions.

The procedural ideal can be seen in action also in evidence-based medicine (EBM)⁶, in which meta-analyses play an important role. According to advocates of EBM, clinical decisions should be grounded on strong scientific evidence. The idea is that instead of personal expert judgments, decisions could be based on guidelines that are composed by collecting evidence from soundly performed scientific studies, preferably systematic reviews of randomized controlled trials (RCTs) (Timmermans & Mauck, 2005: 19). By offering guidelines for clinicians, EBM aims at “limiting idiosyncrasies in [...] clinical procedures” (Timmermans & Mauck, 2005: 20). Moreover, EBM is thought to provide a scientific basis for public policy (Timmermans & Mauck, 2005: 20–21).

The status of randomized controlled trials as the “gold standard” of evidence in medical research has generated plenty of discussion: for instance, the external validity of the results of these trials is argued to be limited (e.g., Cartwright, 2007; Cartwright & Hardie, 2012⁷). I shall be engaging in this debate only in passing. However, my aim is somewhat similar to the one that the critics of RCTs have, i.e., examining whether an ideal that guides activities is fruitful in practice. As mentioned in the introduction, one reason for striving for objectivity is that objective methods are thought to produce knowledge that can be utilized successfully. Advocates of meta-analyses and EBM seem to consider aiming at the procedural ideal to be the best way of producing unbiased, applicable knowledge. This makes it particularly interesting to study whether meta-analysis satisfies this ideal, and whether the ideal itself is successful in producing knowledge for pragmatic purposes.

Stegenga’s Argument Against the Objectivity of Meta-Analysis

Next I proceed to discussing Stegenga’s (2011: 498) criticism of meta-analysis as “the platinum standard of evidence”. The core of Stegenga’s argument is that it is unwarranted to praise meta-analysis for being objective since analysts need to make multiple decisions involving judgments at different stages of the process. For instance, when choosing what primary evidence to analyse, a researcher needs to consider at least the following questions: What methodological quality criteria should the included studies meet? Can study parameters diverge? How can the problems caused by publication bias be solved? The rules for conducting analyses leave space for deciding how to proceed with respect to these questions. According to Stegenga, the subjectivity included in the process is indicated by the meta-analyses that have reached contradictory conclusions on the same hypotheses. In other words, contrary to what the promoters of meta-analysis say, using this method does not free scientists from personal judgments, which means that the procedure is not objective, according to Stegenga. It needs to be stressed that Stegenga (2011: 505) is not against using the method as such, but states that “the epistemic prominence given to meta-analysis is unjustified”. (Stegenga, 2011: 497–505.)

As mentioned, judgments are involved at different stages of conducting meta-analyses. Two additional issues discussed by Stegenga (2011: 502) need to be highlighted as relevant to the argument of this paper. The first is publication bias: studies that show positive and statistically significant results are published more often than studies with negative or inconclusive results. This bias affects the results of meta-analyses by limiting the pool of primary evidence. Duplicate publication, namely

that the same trials are published multiple times in different journals, can further skew the picture that the published studies paint of the efficacy of treatments. There are methods⁸ for trying to correct the effect of publication bias (cf. Shercliffe et al., 2009: 420; Torgerson, 2006: 95–96), but because analysts differ in how much effort they put in trying to fix the bias the results of their analyses vary, which in turn denotes that subjectivity has entered the process (Stegenga, 2011: 502). Thus, according to the ideal of procedural objectivity, publication bias threatens the objectivity of meta-analyses by introducing a further need for judgments.

The second issue relevant to the focus of this paper is the homogeneity of evidence that is included in meta-analyses. The dominant view is that only evidence from RCTs should be included in meta-analyses. This means that other types of statistical (e.g., cohort studies) or non-statistical (e.g., results from pathophysiological studies) evidence are excluded. In addition, the included studies should be sufficiently similar with respect to subjects, results and interventions (Cochrane Collaboration, 2015: 9.5.1.). As Stegenga (2011: 501) acknowledges, it is justifiable to demand a certain amount of homogeneity from the included studies as the purpose of meta-analysis is to measure a causal relation between the studied treatment and the effect of interest. If original studies are designed to detect completely different effects, or study populations differ significantly, conducting a meta-analysis does not make sense. According to Stegenga (2011: 500–502), however, the objectivity of the process is threatened because of the judgments researchers need to make while deciding which data are suitable to be included in analyses. In addition to the threat that judgments pose to objectivity, the external validity of meta-analyses may

be limited by the lack of evidential diversity. Relying on evidence from RCTs while ignoring other types of evidence, such as the outcomes of case-control studies, “risks making uninformed judgment [...] on a hypothesis” (Stegenga, 2011: 501).

To repeat, Stegenga’s claim is that despite the guidelines for conducting meta-analyses, diverse judgments are involved in the process and thus subjectivity cannot be removed. This makes it unlikely to achieve the objectivity of the process in the sense that the promoters of meta-analysis strive for. In the next section I proceed to discussing the shortcomings of this procedural view. Later, in section 5, I shall argue that instead of condemning meta-analysis for not fulfilling the procedural ideal, the concerns arising in the discussions concerning the method, its weaknesses and strengths can be taken as demonstrations of the inability of the procedural ideal to capture some of the problematic issues involved in medical research. These problems are better accommodated by the so-called social view on objectivity. It has to be emphasized that the view I am presenting is not to be understood as opposing the concerns that Stegenga raises. Instead, I argue that it would be conducive to approach the issues he discusses with a different ideal of objectivity.

Weaknesses of the Procedural Ideal

The implicit assumptions underlying the ideal of procedural objectivity, i.e., the account that judgments threaten objectivity and thus diminish our chances of achieving results that can be utilized without complications, are 1) that it is possible to conduct scientific inquiry by following a formally specified rule, and 2) that striving for performing a process that does not involve judgments is the best way for making sure that the preferences of involved

parties are not reflected in the outcomes. However, as Douglas (2004) has argued, this procedural objectivity is only one of the senses in which the term objective is used, and objectivity in one sense does not necessarily guarantee objectivity in other senses. Associating the objectivity of science with this ideal is debatable, and several authors (e.g., Longino, 1990; Carrier, 2010; Hammersley, 2013) in social epistemology have argued that the objectivity of scientific inquiry may not be best understood as mechanical application of rules.

Why, then, does the ideal of procedural objectivity fall short as a model of objectivity for meta-analysis and medical research in general? First of all, as is well known, there cannot be formal rules of proceeding for every step of the research process (e.g., Longino, 1990; Wilholt, 2013). Second, even if there were such rules and following them without judgments was possible, some of the problematic issues related to medical research could still not be removed. This will be demonstrated shortly: it seems that controlling biases in producing knowledge in medical science requires something more than striving for amalgamating evidence without making judgments. Next I will present a brief case on research on selective serotonin reuptake inhibitors. By introducing this example, I demonstrate why the procedural view on objectivity does not capture some of the essential issues related to what objectivity should comprise in the context under discussion. The problematic practices that are introduced by the SSRIs example are not specific to research on this type of drug, but to pharmaceutical research in general. Because of this, discussing this case is of relevance to the philosophy of medicine and commercialized research more broadly.

Research on selective serotonin reuptake inhibitors

According to the statistics of WHO (2012), more than 350 million people worldwide suffer from depression. Nowadays the illness is often treated with selective serotonin reuptake inhibitors (SSRIs) that were accepted for use in the 1980s (Lawlor, 2012: 176). SSRIs have become one of the most commonly used drugs in the world (Fergusson et al. 2005). Lately, dissenting voices have begun to question the extensive use of the drugs. Some critics (e.g., Horowitz & Wakefield, 2007) worry that people facing hardships that are an inseparable part of human life are excessively diagnosed as suffering from pathological depression and prescribed medication, others (e.g., Kirsch, 2010) have even claimed that the effectiveness of SSRIs is based on the placebo effect. Here, however, I discuss the debate on the possible side-effects of the drugs.

The discussion on the adverse effects of SSRIs began in the 1990s when stories about suicidal behaviour that was linked to the use of the drug started to appear (Healy & Whitaker, 2003: 332). However, it was not before October in 2004 that America's Food and Drug Administration (FDA) decided that a black box warning should be added on the packages of antidepressants to inform patients about the increased risk of suicidal tendencies in children and adolescents taking these drugs (FDA, 2004)⁹. Therefore, questions that need to be asked are: Why did it take almost 15 years before the risks were officially recognized? Which factors have had an impact on the discussions on SSRIs?

In a recent article Pigott et al. (2010) examined four meta-analyses of the efficacy of antidepressants. All reviewed studies analysed trials registered by FDA before they were started. Thereby, researchers were able to compare published results with the outcomes that could be drawn from

the data sent to FDA. Analyses “document a profound publication bias that inflates [SSRIs’] apparent efficacy” (Pigott et al., 2010: 267). In addition, published studies accentuated positive results for outcome measures that were not pre-specified as primary - or pre-specified at all - while leaving the negative results for pre-specified outcome measures unpublished. (Pigott et al., 2010: 267.) For instance, Turner et al. (2008), who compared data on 12 antidepressants (including SSRIs such as Citalopram, Escitalopram, Fluoxetine and Paroxetine), concluded that negative or inconclusive studies were often published as reporting positive results.

Insufficient reporting of data was established also when Whittington et al. (2004) analysed published and unpublished data on treating depressed children with SSRIs: even though published data gave support for treating children with SSRIs, unpublished evidence suggested not only that citalopram, venlafaxine, paroxetine and sertraline may not be efficacious but also that using the products may involve the risk of suicidal behaviour. Similarly, Fergusson et al. (2005) conducted a systematic review of RCTs to examine the possible connection between SSRIs and suicide attempts. Their result was that there is a “more than a twofold increase in the rate of suicide attempts in patients receiving SSRIs compared with placebo or therapeutic interventions other than tricyclic antidepressants” (Fergusson et al., 2005: 398). According to the authors, the increased risk may not have been noticed in individual studies because of the small trial sizes and the rarity of suicides and attempted suicides. Likewise, the studies analysed by Whittington et al. were not designed to detect suicides. Even though the increased risk of suicidal behaviour would be low, it is a cause for concern at the population level due to the commonness

of SSRI-treatment. Fergusson et al. (2005: 399) remark that the duration of clinical trials tends to be short, and thus the long term benefits and risks of treatment may go unnoticed.

David Healy (2002: 259) - one of the most vocal critics of the nonchalant use of SSRIs - has claimed that the financial interests of the pharmaceutical industry played a role in the problems of the drugs being unnoticed. His claim is that since the companies sponsoring trials had similar interests, i.e., reaching outcomes according to which the drugs are safe and effective, the studies were designed in ways that made the detection of adverse events less likely¹⁰ (Healy, 2002: 259). Healy is not alone in claiming that pharmaceutical research might be skewed by commercial interests. For example, Als-Nielsen et al. (2003) examined 370 RCTs to determine if an association existed between the source of funding and conclusions. The study concluded that in for-profit-trials the quantitative results tended to be interpreted more positively, i.e., the tested drug was recommended for use more often, than in other trials. A recent review by Lundh et al. (2012) states that methods of industry sponsored trials may be chosen in ways that lead to results favoured by the sponsors. According to Sergio Sismondo (2008a), the sponsors may influence the outcomes of studies both indirectly and directly. First of all, the prevalence of pro-industry findings could be partly explained by unrecognized obligations that the researchers paid by the industry feel towards their sponsors: their judgments can be influenced by the feeling of needing to compensate for the benefits they receive from their employer. Second, industry employs more direct ways of influencing results. Publication planning, or “ghost management of medical research includes practices such as choosing test subjects, endpoints, comparators, the doses of tested products and comparators, and trial

duration in a way that makes adverse events less likely to appear” (Sismondo & Doucet, 2010: 275)¹¹. Also, withholding negative data or interpreting it in a questionable way can be used to conceal side effects. In addition, trials may be prematurely aborted and study protocols altered when the study is already in the making. (Sismondo, 2008a: 1910–1912.) One approximation is that 40% of publications on new drugs have been ghost-managed in this way by the industry (Sismondo, 2009: 172).

Objectivity of research on SSRIs

The published outcomes of the trials on SSRIs painted an overly positive picture on the efficacy and safety of these drugs, which was at least partly brought about by practices that were due to ghost-management. As the discussions on objectivity focus on what kinds of actions best ensure that the outcomes of inquiry are not unjustly influenced by the inclinations of involved parties, the case can be said to demonstrate a violation of objectivity. Next I shall itemize which factors delayed the recognition of SSRIs’ risks.

Firstly, as the studies by Fergusson et al. (2005) and Pigott et al. (2010) demonstrate, due to the publication bias, the published data indicated that the products were safer and more effective than later research has shown them to be. Since there is evidence of systematic disappearance of negative data, and since this seems to be connected to commercial interests, we should denounce this phenomenon as a violation of objectivity: the interests of involved parties have unduly guided research towards certain kinds of outcomes. As mentioned above, the existence of the difference between published and unpublished data has been demonstrated by meta-analyses, such as the one conducted by Whittington et al. (2004). However, these analyses can be carried out only if researchers get access

to unpublished material. This, in turn, requires institutional practices such as the preregistration of trials. What this means is that exposing the real scope of the bias is also dependent on processes that concern communal practices of science widely understood: structural factors, which are not traditionally perceived as relevant to the justification procedures (e.g., Brown, 2010), turn out to be relevant with respect to checking the objectivity of the process. The availability of produced data, for instance, can hinge on the institutional arrangements of the community.

Secondly, primary studies were too small and short in duration to detect rare adverse effects (Fergusson et al. 2005: 399; Vandembroucke & Psaty, 2006: 2417). As the promoters of meta-analysis and Stegenga (2011: 498–499) state, this is a problem that meta-analyses can help to correct.

Thirdly, empirical work on commercial research suggests that studies were designed in a way that made it more difficult to detect some of the effects of the drugs: for instance, due to the lack of suitable categories, some adverse events may have been misleadingly coded (Healy & Cattell, 2003; Healy, 2011: 151) and study protocols altered (Pigott et al., 2010). These are evident violations of good scientific practice, and conducting a meta-analysis on the data does not help to remove these biases. Moreover, methodological choices that raise the chances of tested products appearing more effective seem to be connected to the financial interests of the sponsors of studies (Lundh et al., 2012). It is also noteworthy that information on the dubious practices of industry has become available via legal actions involving pharmaceutical companies (e.g., Healy & Cattell, 2003), which highlights the importance that institutional extra-scientific factors have on the conditions for critical activities of science.

If the above approximations of prevalence of ghost-management are correct, a considerable part of published information on treatments is shaped by commercial interests. Even if they followed the ideal of procedural objectivity, meta-analyses alone could not solve the problem of biases as they emerge in this case: if the pool of primary evidence is already skewed, the outcome of the meta-analysis will not be unbiased either. However, they can help researchers to identify that problems exist, for instance by showing how published and unpublished data give rise to different conclusions. Procedural objectivity strives for removing personal biases from evidence synthesis by making judgments redundant. Even if meta-analysis succeeded in this - which it does not, as Stegenga's analysis demonstrates - it could not eliminate the more systematic biases of research. The case of SSRIs also exemplifies how institutional practices that are not part of justification procedures, such as preregistration of trials or legal action, may be needed for making biases or their sources visible. As mentioned, researchers had to have access to unpublished data before they could verify that some of the adverse events had not been reported. In the following section I shall argue that the procedural ideal of objectivity does not accommodate these points, and thus we should pursue a different ideal of objectivity for promoting reliable research.

The Social View on the Objectivity of Medical Research

The procedural ideal of objectivity seems to be both on the one hand practically unreachable, and on the other, insufficient in solving some of the problems of medical research. One of its weaknesses is that it cannot discern between acceptable and unacceptable judgments since judgments

per se are seen as destroying objectivity. However, since judgments are a necessary part of scientific activities, in order to be applicable to actual practices, the conception of objectivity should be capable of making these distinctions. For example, referring to the above discussed case, the adopted conception should clarify what it is that makes ghost-management problematic.

I have argued elsewhere (Jukola, 2014) that when discussing the objectivity of research in fields with applicable and socially relevant outcomes, we should adopt a view of this virtue that considers also community-level actions and structures: *the social view on objectivity*. This view accommodates the intuition that judgments do not have to be destructive to objectivity while, at the same time, making it possible to discern between acceptable and non-acceptable judgments. Further, this view helps one to understand how non-epistemic considerations can be incorporated into the research process without letting them steer research unduly. According to the social view, judgments as such do not destroy objectivity because, unlike those who praise meta-analysis for its alleged procedural objectivity seem to imply, they are not private and impossible to evaluate by nature (Hammersley, 2013: 63). This means that it is possible to examine whether the decisions made during the course of research have been warranted.

The social view and the need for making judgments

The social view holds that it is not enough that researchers strive for "thinking and drawing conclusions based on strict logical adherence to relevant facts" (Smith, 2004: 152) or following rules to the letter. The social view on objectivity takes as its starting point the realization that, in practice, conducting scientific inquiry always seems

to involve the need for making judgments and choosing between different ways of proceeding. Even if scientists are familiar with the guidelines they are to follow - i.e. they know the rules of the trade - and are willing to adhere to them, methodological conventions do not fully determine how research should be carried out. This is lucidly demonstrated by Stegenga's discussion on meta-analysis. However, unlike Stegenga claims, this does not mean that the objectivity of a process is necessarily under threat.

The question of how to incorporate the necessity of making judgments into our notion of objective inquiry has been comprehensively addressed in recent social epistemology. According to Helen Longino (2002: 184), different theoretical and methodological preferences may lead researchers to apply rules in disparate ways. Whether one finds a given application to be acceptable depends on the views one has concerning the theoretical and methodological assumptions that have given reasons for the application.

James Tabery's (2014: ch. 4 & 6) analysis of a dispute about whether depression is related to the gene-environment interaction between the serotonin transporter gene and stressful life events offers an example of this in the context of meta-analyses. In this case, different meta-analyses on an apparently same hypothesis reached contradictory outcomes due to different background assumptions. According to Tabery, the contradictory results emerged because researchers in different teams had diverging views on how to define the variables of the hypothesis and, thus, disagreed on which studies should be included in analyses. For example, as one team focused only on primary studies that had included cases where individuals had suffered many stressful life events, another team had included also those studies where

individuals had suffered one stressful event. (Tabery, 2014: 162-163.) Pointing out that judgments were made when these meta-analyses were conducted does not seem to be a fruitful way of criticizing either one of these studies. In this example, the differences in conducting the meta-analyses seem to be unrelated to non-epistemic interests of analysts, but were motivated by different theoretical assumptions. In order to evaluate the acceptability of the ways in which each analyst applied the rules of conducting meta-analyses, constructive criticism of research would have to heed these assumptions.

In addition to epistemic assumptions, different ways of applying a rule can be based on different understanding of the non-epistemic consequences of research. In her influential paper, Heather Douglas (2000) argued for a way of incorporating value judgments in research in a way that does not compromise the reliability of results. Her discussion is based on the idea of inductive risk: Data never provide conclusive certainty on the correctness of a hypothesis and the rules for conducting inquiry are not binding, and because of this uncertainty, researchers need to ponder what kind of consequences may follow from possible mistakes. Some research projects are expected to have non-epistemic consequences - for example, permissions to market new pharmaceutical products are granted on the grounds of research outcomes. In these cases, conducting inquiry should involve making value judgments concerning which risks are worth taking and which possible non-epistemic outcomes should be avoided. (Douglas, 2000.) Torsten Wilholt has argued for a similar view. According to him, our trust in scientific research is based on the assumption that when methodological rules leave space for judgments, the decisions that are made during research

are “based on the presumption of shared ideas about the values of true results and the dangers inherent in errors” (Wilholt, 2013: 248). In other words, the theoretical and methodological ideas and views concerning the goals of research help to bridge the gap between a methodological rule and its application in a given situation (see also Intemann, 2005: 1010–1011). For criticizing the judgments concerning how to proceed with research, it is necessary to scrutinize the epistemic and non-epistemic assumptions they are based on.

It should be noted that adopting the social view does not mean that one should abandon the idea of striving for following procedures altogether – the ideal of procedural objectivity would just lose its status as the ruling ideal. The social view can very well incorporate the goal of following a procedure as far as possible as a means of counteracting some problems of knowledge production. As individuals are prone to reasoning fallacies, guidelines for conducting experiments and analyses need to be formulated to guard against biases operating at the individual level (e.g., Ioannidis, 2005: 0698; Howick, 2011: 166). In this way, the social view invites the use of communally accepted guidelines, such as the Cochrane handbook, for regulating the activities of individuals. In other words, the social view can incorporate elements of the procedural ideal. However, unlike the procedural view, the social view does not require that a rule should be applied similarly in every situation.

In this case, objectivity does not mean that all researchers have to apply the rules in exactly the same ways or that research is value-free¹². Does this mean that anything goes? No. What objectivity requires, according to the view advocated here, is that when researchers make the judgments and decisions concerning how to conduct research, they need to consider which way

of proceeding would best support the goals of inquiry. Yet, as humans are fallible and prone to biases (e.g., Uhlmann & Cohen, 2007), this individual effort alone is not enough. As Longino (1990: ch. 4; 2002: ch. 6) has argued, a social backup mechanism is needed for auditing the practices of individual scientists, i.e., questioning the assumptions and goals their actions are based on. This means that the institutional practices in communities must enable critical evaluation of different stages of research. Practices that improve the chances of critical exchanges can be labelled as supportive with respect to objectivity and, obversely, factors that limit the possibility of criticism pose a threat to objectivity.

In the case of conducting meta-analyses, the social view holds that researchers need to be aware of the communally accepted guidelines and conventions for conducting analyses (for instance, *The Handbook of Cochrane Collaborations*, 2015) and do their best in obeying them. When the rules need to be applied, researchers should consider which ways of proceeding best contribute to the aims of inquiry. For example, as Stegenga (2011: 499) remarks, there is a link between the financial ties of researchers and the outcomes of the meta-analyses they have conducted. While both the social view and the procedural ideal of objectivity recognize these incidents as violations of objectivity, the social view offers more tools for analysing these cases. This is because while the procedural view finds the process to be non-objective once judgments are involved, the social view is capable of examining the situation further – and without judging all commercial research to be biased. According to the social view, inquiry is not objective if the decisions involved in the research process have been made for reasons that are not in line with the generally acknowledged goals of research¹³. Thus, if research is conducted in a way that

is motivated by interests inconsistent with pursuing these goals, it can be criticized for not being objective. For instance, ghost-management of medical studies becomes problematic when commercial interests override the concerns for finding the causes of and cures for illnesses. In these cases, it becomes less likely that the produced knowledge can be successfully applied in healthcare decisions. In this way, the social view shows us why some of the decisions made by actors involved in the research on SSRIs can be condemned.

The social view and publication bias

A major issue related to research on SSRIs was that a considerable part of studies was never published. Because it acknowledges the institutional context of research and evaluating practices at the community level, the social view is better at capturing the troublesome features of publication bias than the procedural view. The problematic nature of this bias becomes evident only when research on a given topic is regarded in its entirety - i.e., it cannot be recognized by evaluating individual studies only. In addition, this bias arises as a result of a prevalent communal practice. (Jukola, 2014.) Thus, its features are not best understood from a perspective that focuses on whether individuals follow rules to the letter¹⁴. The social view can help us identify how the current system of medical research encourages practices leading to publication bias and why this can be denounced as reprehensible: The imbalances between negative and positive outcomes that are supported by choices that have been motivated by interests not in line with the accepted goals of research are clearly reprehensible. One suggested reason for the prevalence of publication bias is that journal editors may be unwilling to publish negative or inconclusive results (e.g., Resnik, 2007). As researchers are work-

ing in an environment where the number of publications can make or break their careers, we cannot simply require them to start submitting their negative findings if those are likely to be rejected. However, a recent study found no decisive empirical evidence for the claim that editors were biased against negative studies (Chan et al., 2014). Instead, in medical research, the absence of studies with negative conclusions seems to be connected to financial interests, as was already mentioned. Consequently, new drugs are systematically favoured in industry sponsored trials, which affects the literature in general, and thus meta-analyses as well. (Bekelman, Li & Gross, 2003; Ioannidis, 2005; Sismondo, 2008a: 4; Sismondo, 2008b.)

From the perspective of the social view, publication bias is problematic as it disturbs critical interactions within and between communities. When some data are not published, other researchers cannot properly evaluate the soundness of research outcomes. In addition, replication becomes nearly impossible (Glasziou et al., 2014). Thus, publication bias not only leads to a distorted picture of the object under study, but it also further undermines the conditions for objectivity by violating the openness of research. Also, withholding evidence may increase the need for judgments involving non-epistemic elements: when researchers have to make their decisions on the grounds of less data than would otherwise be necessary, the uncertainties involved increase.

Identifying publication bias as a community-level problem can also facilitate finding solutions to it, and many of the already suggested improvements reflect the spirit of the social view. Even if it is not usually explicitly stated, suggestions for improving the reliability of medical research often aim at improving the conditions of critical evaluation of studies. For example,

in 2004 the International Committee of Medical Journal Editors stated that trials on drugs cannot be published unless they have been registered before the beginning of the trial. (De Angelis et al., 2004). The motivation for this policy was to facilitate the detection of missing data and terminated trials, and thus govern publication bias. There are other suggestions for improving the objectivity of medical research that are perfectly in line with the ethos of the social view as well. As an example an influential paper by Ioannidis (2005) can be mentioned. Ioannidis argues that the body of published medical literature is seriously skewed as a result of the way individual studies are conducted, the prevalence of bias, and the way in which competition between research teams encourages practices that do not contribute to producing reliable knowledge. Ioannidis states that in addition to methodological improvements and adhering to common standards of conducting studies, for improving the situation it is important to evaluate the totality of evidence in a field of research, instead of focusing on evaluating individual studies. Preregistration of studies is one of the measures that can be taken to enable this. (Ioannidis, 2005: 0701.) Likewise, Young et al. (2008) state that the current publication practices distort science: Published results tend to give an exaggerated picture of the state of research because in biomedical research, the few prominent journals have high rejection rates. Because publishing in prestigious journals is a precondition for building a successful career, it is not rational to submit research that is likely to be rejected, e.g., replications of previous work (Young et al., 2008: 1419). According to Young et al., this situation is maintained somewhat artificially, as print page limits can be taken to be an excuse in the current age of online publishing. Digital publication

could facilitate the publication of a greater number of articles: those papers that pass peer review but are not considered to take priority could be published online. (Young et al., 2008: 120–121.) The social view explains why these policies can support the objectivity of research. By changing the institutional context of science, such as the publication system, it is possible to improve the conditions for critical interaction and the evaluation of research¹⁵.

The social view and the lack of evidential diversity

The social view not only helps us to identify factors that can bias research, it also helps to scrutinize why the lack of evidential diversity in meta-analyses can be seen as undermining the applicability of results. This is one of the points that Stegenga raises against meta-analysis's high status. According to him (2011: 500–502), the objectivity of process is threatened when researchers need to judge which data are suitable to be included in the analysis. However, as has been argued, we should adopt a conception of objectivity that recognizes how judgments *per se* do not impair objectivity, but that the problem lies elsewhere: the principle of total evidence may be violated when a decision on whether to reject a hypothesis or not is based only on RCTs. There are two diverse reasons for arguing that the homogeneity of included evidence may threaten the reliability of treatment guidelines: First, as was already mentioned in subsection 4.2., there is evidence on widespread bias in medical RCTs. If trials are systematically designed to reach positive results (e.g., Sismondo, 2008a; Sismondo & Doucet, 2010), and meta-analyses do not succeed in filtering out these biases, basing treatment guidelines on meta-analyses of RCTs is problematic. Second, similarly implied when the objectivity of research on SSRIs was

discussed in subsection 4.2, even properly conducted meta-analyses of ideal, unbiased RCTs produce evidence that might not alone suffice to answer some questions with considerable practical relevance.

A motivation for aiming at an objective method is the wish to gain reliable knowledge that does not lead us astray in our actions. In order to achieve reliable knowledge to back up treatment guidelines, evidence on the possible side effects of drugs is needed. According to critics, RCTs may not be able to deliver this information. For instance, as mentioned in subsection 4.2., trials are often too short in duration to detect the effects that appear later during the treatment, and the number of subjects may not be high enough to enable the appearance of rarer effects (Vandenbroucke & Psaty, 2006). As already stated, this is a problem meta-analyses can help to solve by pooling together more patient-level data. Another worry is the representativeness of trials. If the recruited subjects are young and in relatively good health, while the members of the eventual target group for the tested product tend to suffer from several illnesses, the outcomes of the RCT may not offer knowledge that could be applied without complications. (Vandenbroucke & Psaty, 2006: 2417.) As Cartwright and Hardie (2012: 122) state, properly designed and run RCTs are good for evaluating if a certain policy or a drug causes a certain effect of interest in a certain population but they are narrow in scope¹⁶. Yet another reason for questioning the authority of RCTs on providing evidence on adverse effects is that since side effects are usually unintended, they are difficult to record systematically (Vandenbroucke & Psaty, 2006: 2417).

According to the social view on objectivity, the epistemic and non-epistemic risks need to be considered when methodological choices are made, and the context and aims of research should

to be acknowledged when evidence is assessed. Along these lines, it has been suggested that avoiding errors with severe consequences may legitimize setting the criteria of sufficient evidence differently for establishing that a drug is efficacious and for establishing that it has side effects. Since the consequences of accepting a false hypothesis on the safety of a drug can be severe, the constraints on the acceptability of evidence should be “highly flexible” (Osimani, 2013: 457). For this reason, in the case of unintended side effects, collecting and analysing all available evidence, statistical and non-statistical, is a better approach to establishing if a drug has possible negative effects (Osimani, 2013: 459). In a similar vein, Stegenga (2011) suggests adopting the Hill strategy for evaluating medical evidence: Sir Bradford Hill argued that instead of accepting evidence only from RCTs, good arguments on a possible causal connection could be made if separate pieces of evidence support the claim, including, for instance, the plausibility of the suggested causal connection given the existing knowledge of possible biological mechanisms, the coherence between the causal interpretation and existing knowledge, and experimental evidence. The only desideratum necessary for establishing a causal relation is that the cause has to precede the effect. Otherwise, all of the points do not have to be met. (Stegenga, 2011: 504–505.) Particularly in the context where the commercial interests of industry have an influence on how RCTs are carried out, it is relevant to consider how these interests may impact the evidence that is produced, and then, modify methodological conventions of evaluating evidence accordingly (Osimani, 2013: 460). In this way, the risks that are related to research and its consequences can be acknowledged and minimized.

Because the social view allows the possible non-epistemic risks that are involved in research to be taken into consideration, it can help to solve some of the problems that arose in the case of SSRIs discussed in the section 4. According to this view, considering the results of RCTs only is not acceptable if there are reasons to assume that the possible existence of some serious side effects could not be established by them alone. In addition, the social view grants us the view to consider the possibility that decisions concerning how the RCTs were carried out were motivated by interests that were in conflict with generally accepted goals of biomedical research, e.g., finding a safe and effective treatment for an ailment. Moreover, this view acknowledges that conditions for producing reliable knowledge are partly dependent on contextual factors. In the SSRIs case, widely shared financial interest motivated actions that gave rise to serious publication bias, which could be avoided by previously mentioned actions such as preregistration of trials.

Conclusion

I have argued that the ideal of procedural objectivity as the guiding rule in medical research should be abandoned. This is because the ideal, on the one hand, is practically unattainable, and, on the other hand, does not help to evaluate all of the practices that are relevant in producing reliable medical knowledge. The issues emerging when knowledge is produced for solving medical problems are better dealt with by invoking the social view on objectivity. The social view offers us tools for separating legitimate judgments from illegitimate ones, and allows us to take notice of the risks involved in research. In addition, it takes into account the possibility of systematic biases, such as the publication bias.

As was mentioned in subsection 5.1., the social view is not fully incompatible with the procedural ideal. However, the procedural ideal is insufficient in science. Comprehending this and the way in which it may be useful to establish some codes of conduct according to the procedural ideal, helps one to better appreciate the role that meta-analyses have in producing reliable knowledge. Meta-analyses are no miracle tools that can be conducted without judgments. However, the rules for conducting them can be taken as an example of guidelines that curtail idiosyncratic preferences to a certain degree, i.e., methodological conventions that “facilitate epistemic reliance within science” (Wilholt, 2013: 244), but cannot fix all methodological steps. Adopting the social view helps us to see why the evidence produced by meta-analyses may be more reliable than the results of some other means of amalgamating evidence without having to adhere to the unattainable ideal of procedural objectivity.

The reliability of meta-analyses’ results depends not only on the way the analysts obey the rules, but also on contextual issues. First of all, the primary studies included in the analyses have to be soundly performed. Second, as RCTs may not be the best possible means of acquiring evidence on possible side-effects of drugs, outcomes of meta-analyses drawing from RCTs may not suffice to offer reliable knowledge on whether a given product should be used to treat a given ailment. The example on research on SSRIs highlighted the twofold relation between meta-analysis and the biases of original studies. On the one hand, publication bias impairs the reliability of meta-analyses’ results by skewing the pool of available primary evidence. At the same time, however, meta-analyses and other systematic reviews (e.g., Bekelman et al., 2003) have demonstrated how published

and unpublished data differ. Together with qualitative reports (e.g., Sismondo, 2008a) these analyses help us identify problematic practices and find means of counteracting them. Since the social view recognizes that the level of evidence required for establishing different kinds of claims may be different, it allows us to hold on to the idea that meta-analyses produce valuable information on the effectiveness of treatments while, at the same time, acknowledging the limitations that the method has with respect to detecting possible side-effects. The social view directs our attention to evaluating research in its context, and thus adopting this view gives us tools for criticizing a system that disregards evidence produced by other means than RCTs and meta-analyses without abandoning the goal of objectivity. In this way, the social view “preserves” the objectivity of meta-analysis.

Acknowledgements

The research for this paper was funded by a grant from the Kone foundation. I thank Aki Lehtinen, Petteri Niemi, Jani Raerinne, Kristina Rolin, Jacob Stegenga, Miira Tuominen, Torsten Wilholt, and Petri Ylikoski for their feedback.

References

- Als-Nielsen B, Chen, W., Gluud, C. & Kjaergard, L (2003) Association of Funding and Conclusions in Randomized Drug Trials: A Reflection of Treatment Effect or Adverse Events? *The Journal of American Medical Association* 290(7): 921-928.
- Bekelman J, Li J & Gross C (2003) Scope and Impact of Financial Conflicts of Interest in Biomedical Research. *The Journal of American Medical Association* 289(4): 454-465.
- Brown JR (2010) One-Shot Science. In: Radder H (ed.) *The Commodification of Academic Research*. Pittsburgh, PA: Pittsburgh University Press, 90-109.
- Carrier M (2010) Research Under Pressure: Methodological Features of Commercialized Science. In: Radder H (ed.) *Commodification of Academic Research*. Pittsburgh, PA: Pittsburgh University Press, 158-186.
- Cartwright N (2007) Are RCTs the Gold Standard? *Biosocieties* 2(2): 11-20.
- Cartwright N & Hardie J (2012). *Evidence-based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press.
- Chan A, Song F, Vickers A, Jefferson T, Dickersin K, Goetzsche P, Krumholz H, Gherzi D & van der Worp HB (2014) Increasing Value and Reducing Waste: Addressing Inaccessible Research. *The Lancet* 383(9913): 257-266.
- Cochrane Collaboration (2015). *Cochrane Handbook*. Available at: <http://handbook.cochrane.org/> (accessed 14.8.2015).
- Daston L & Galison P (2010) *Objectivity*. New York: Zone Books.
- De Angelis C, Drazen J, Frizelle F, Haug C, Hoey J, Horton R, Kotzin S, Laine C, Marusic A, Overbeke A, Schroeder T, Sox H & van der Weyden M (2004) Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. *New England Journal of Medicine* 351(12): 1250-1251.
- Douglas H (2000) Inductive Risk and Values in Science. *Philosophy of Science* 67(4): 559-579.
- Douglas H (2004) The Irreducible Complexity of Objectivity. *Synthese* 138(3): 453-473.
- FDA (US Food and Drug Administration) (2004) News Release: FDA Launches a Multi-Pronged Strategy to Strengthen Safeguards for Children Treated

- With Antidepressant Medications. Available at: <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2004/ucm108363.htm#> (accessed 14.8.2015).
- FDA (US Food and Drug Administration) (2007) Antidepressant Use in Children, Adolescents, and Adults. Available at: <http://www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/UCM096273> (accessed 14.8.2015).
- Fergusson D, Doucette S, Glass KC, Shapiro S, Healy D, Hebert P & Hutton B (2005) Association Between Suicide Attempts and Selective Serotonin Reuptake Inhibitors: Systematic Review of Randomised Controlled Trials. *British Medical Journal* 330(7488): 396–402.
- Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Moher D & Wager D (2014) Reducing Waste from Incomplete or Unusable Reports of Biomedical Research. *The Lancet* 383(9913): 267–276.
- Goldenberg M (2009) Iconoclast or Creed? Objectivism, Pragmatism, and the Hierarchy of Evidence. *Perspectives in Biology and Medicine* 53(2): 168–187.
- Hammersley M (2013) *The Myth of Research-Based Policy & Practice*. London: SAGE.
- Healy D (2002) Conflicting Interests in Toronto. Anatomy of a controversy at the interface of academia and industry. *Perspectives in Biology and Medicine* 45(2): 250–263.
- Healy D (2011) Science, Rhetoric, and the Causality of Adverse Events. *The International Journal of Risk and Safety in Medicine* 23(3): 149–162.
- Healy D & Cattell D (2003) Interface between Authorship, Industry and Science in the Domain of Therapeutics. *British Journal of Psychiatry* 183(1): 22–27.
- Healy D & Whitaker C (2003) Antidepressants and Suicide: Risk-Benefit Conundrums. *Journal of Psychiatry & Neuroscience* 28(5): 331–337.
- Horowitz A & Wakefield J (2007) *The Loss of Sadness. How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford: Oxford University Press.
- Howick J (2011) *Philosophy of Evidence-Based Medicine*. Oxford: BMJ Books—Wiley-Blackwell.
- Intemann K (2005) Feminism, Underdetermination, and Values in Science. *Philosophy of Science* 72(5): 1001–1012.
- Ioannidis J (2005) Why Most Published Research Findings Are False. *PloS Medicine* 2(8):e124.
- Jukola S (2014) Commercialization of Research and the Quest for the Objectivity of Science. *Foundations of Science* 26(3): 1–15.
- Jukola S (2015) Longino's Theory of Objectivity and Commercialized Research. In: Wagenknecht S, Nersessian N & Andersen H (eds) *Empirical Philosophy of Science: Introducing Qualitative Methods into Philosophy of Science*. Cham: Springer International Publishing, 127–143.
- Kirsch I (2010) *Emperor's New Drugs*. London: The Bodley Head.
- Lawlor C (2012) *From Melancholia to Prozac: A History of Depression*. Oxford: Oxford University Press.
- Longino H (1990) *Science as Social Knowledge*. Princeton, NJ: Princeton University Press.
- Longino H (2002) *The Fate of Knowledge*. Princeton, NJ: Princeton University Press.
- Lundh A, Sismondo S, Lexchin J, Busuioic OA & Bero L (2012) Industry Sponsorship and Research Outcome. *The Cochrane Library Methodology Review*.

- Moore D & McCabe G (2006) *Introduction to the Practice of Statistics*. New York: W.H. Freeman and Company. Fifth edition.
- Musschenga A, van der Steen W & Ho V (2010) The Business of Drug Research: A Mixed Blessing. In: Radder H (ed) *The Commodification of Academic Research*. Pittsburgh, PA: Pittsburgh University Press, 110–131.
- Osimani B (2013) Until RCT Proven? On the Asymmetry of Evidence Requirements for Risk Assessment. *Journal of Evaluation in Clinical Practice* 19(3): 44–462.
- Pigott HE, Leventhal A, Alter G & Boren J (2010) Efficacy and Effectiveness of Antidepressants: Current Status of Research. *Psychotherapy & Psychosomatics* 79(5): 267–279.
- Poland G & Jacobson R (2011) The Age-Old Struggle against the Antivaccinationists. *New England Journal of Medicine* 364(2): 97–99.
- Porter T (1992) Quantification and the Accounting Ideal in Science. *Social Studies of Science* 22(4): 633–652.
- Resnik D (2007) *Price of Truth. How Money Affects the Norms of Science*. Oxford: Oxford University Press.
- Ryś P, Władysiuk M, Skrzekowska-Baran I & Małecki M (2009) Review Articles, Systematic Reviews and Meta-Analyses: Which Can Be Trusted? *Polskie Archiwum Medycyny Wewnętrznej* 119(3): 148–156.
- Shercliffe RJ, Stahl W & Tuttle MP (2009) The Use of Meta-analysis in Psychology: A Superior Vintage or Casting of Old Wine in New Bottles? *Theory & Psychology* 19(3): 413–430.
- Sismondo S (2008a) How Pharmaceutical Industry Funding Affects Trial Outcomes: Causal Structures and Responses. *Social Science & Medicine* 66(9): 1909–1914.
- Sismondo S (2008b) Pharmaceutical Company Funding and its Consequences: A Qualitative Systematic Review. *Contemporary Clinical Trials* 29(2): 109–113.
- Sismondo S (2009) Ghosts in the Machine: Publication Planning in the Medical Sciences. *Social Studies of Science* 39(2): 171–198.
- Sismondo S & Doucet M (2010) Publication Ethics and the Ghost Management of Medical Publications. *Bioethics* 24(6): 273–283.
- Smith T (2004) "Social" Objectivity and the Objectivity of Values. In: Machamer P & Wolters G (eds) *Science, Values, and Objectivity*. Pittsburgh, PA: Pittsburgh University Press, 143–171.
- Stegenga J (2011) Is Meta-Analysis the Platinum Standard of Evidence? *Studies in History and Philosophy of Biological and Biomedical Sciences* 42(4): 497–507.
- Tabery J (2014) *Beyond Versus: The Struggle to Understand the Interaction of Nature and Nurture*. Cambridge, MA: MIT Press.
- Timmermans S & Mauck A (2005) The Promises and Pitfalls of Evidence-Based Medicine. *Health Affairs* 24(1): 18–28.
- Torgerson CJ (2006) Publication Bias: The Achilles' Heel of Systematic Reviews? *British Journal of Educational Studies* 54(1): 89–102.
- Turner E, Matthews A, Linardatos E, Tell R, Rosenthal R (2008) Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy. *The New England Journal of Medicine* 358(3): 252–260.
- Uhlmann E & Cohen G (2007) "I Think therefore It's True": Effects of Self-Perceived Objectivity on Hiring Discrimination. *Organizational Behavior and Human Decision Processes* 104(2): 207–223.

- Vandenbroucke J & Psaty B (2006) Benefits and Risks of Drug Treatments: How to Combine the Best Evidence on Benefits with the Best Data About Adverse Effects. *The Journal of American Medical Association* 300(20): 2417–2419.
- Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A & Boddington E (2004) Selective Serotonin Reuptake Inhibitors in Childhood Depression: Systematic Review on Published and Unpublished Data. *The Lancet* 363(9418): 1341–1345.
- WHO (World Health Organization) (2012) *Fact sheet no 369*. Available at: <http://www.who.int/mediacentre/factsheets/fs369/en/index.html> (accessed 14.8.2015).
- Wilholt T (2009) Bias and Values in Scientific Research. *Studies in History and Philosophy of Science* 40(1): 92–101.
- Wilholt T (2013) Epistemic Trust in Science. *The British Journal for the Philosophy of Science* 64(2): 233–253.
- Young N, Ioannidis J & Al-Ubaydli O (2008) Why Current Publication Practices May Distort Science. *PLoS Medicine* 5(10): e201.
- Ziman J (2000) *Real Science: What It Is, and What It Means*. Cambridge: Cambridge University Press.
- for synthesizing information from systematic reviews. Cf. Rys et al. (2009).
- 3 The Cochrane collaboration is an international network that aims to provide reliable medical information for the needs of policy makers and medical practitioners by conducting and publishing systematic reviews.
- 4 As Douglas notes (2004: 462, n 12), Porter calls this ideal mechanical objectivity.
- 5 The link between the ideal of procedural objectivity and meta-analysis has previously been noted by Hammersley (2013: 100).
- 6 See Howick (2011) for a philosophical account of EBM.
- 7 See Goldenberg (2009) for a review of criticism of RCTs high evidential status.
- 8 For example, funnel analysis can be used to detect publication bias. However, this method as such does not help to answer the central question, i.e. whether the tested treatment is effective or not.
- 9 Later, the warning was updated to include young adults between 18 and 24 years of age (FDA, 2007).
- 10 In another article (Jukola, 2015) I discuss in more detail the way in which the source of funding can steer research towards certain kinds of explanations and, thus, bias research. There, I argue that the conditions for objectivity are partly dependent on the funding structure of science. See also Resnik (2007), Carrier (2010), Musschenga, van der Steen & Ho (2010), Sismondo & Doucet (2010).
- 11 Sismondo (2009) examines the work of publication planners and their relationship with other actors in medical science.

Notes

- 1 Here the term reliability is not used in the sense of the discussions on the reliability of measuring instruments, i.e., the ability to produce consistent results.
- 2 Generally systematic reviews and meta-analyses are described in the following manner: A systematic review is a search for literature on a certain question by following explicit, predetermined criteria. A meta-analysis is a statistical technique

- 12 It should be noted that procedural objectivity does not secure value-freedom, either (see, e.g., Douglas, 2004: 462).
- 13 By 'generally acknowledged goals' I refer to goals that both researchers and non-scientists expect scientists to strive for. Some of these goals are shared by all research communities (e.g., empirical accuracy), some are community-specific (e.g., producing knowledge than can be used to develop practices for improving human health).
- 14 One could argue that following the procedural ideal does solve the problem of publication bias. After all, we could state that researchers should follow the rule of publishing all findings. However, it is likely that there will always be findings that, for a reason or another, are not worth publishing, and deciding which results should be made public requires judgment.
- 15 There have been more radical suggestions for tackling the roots of the bias in medical research via institutional changes: For instance, Brown (2010: 106); Carrier (2010: 164; 181) and Sismondo & Doucet (2010: 279) suggest changing the funding structure of the field.
- 16 The fact that both the studied effect and the context of application have to be clearly specified can doom generalizing the results of RCTs in social sciences (Cartwright & Hardie, 2012; Hammersley, 2013).

Saana Jukola
Department of Social Sciences
and Philosophy
University of Jyväskylä
Ylistönmäentie 33
40014 Jyväskylä, Finland
email: saanajukola@gmail.com