



MAKING MEDICAL KNOWLEDGE

MIRIAM SOLOMON

Reviewed by Michael Wilde

Making Medical Knowledge

Miriam Solomon

Oxford: Oxford University Press, 2015, £35

ISBN 9780198732617

In *Making Medical Knowledge*, Miriam Solomon describes a variety of epistemological approaches, or 'methods', employed by medical researchers and practitioners. In particular, there is a detailed discussion of consensus conferences, evidence-based medicine, translational medicine, and narrative medicine. The book ends with a case study of the recent controversy over whether screening mammography leads to a reduction in breast cancer mortality. An important thesis of the book is that such a controversy is best explained by acknowledging that medical researchers rely upon a plurality of methods, and that sometimes these methods give conflicting results. In other words, in medicine there is 'a developing, untidy, methodological pluralism' (p. 208).

In cases where the methods conflict, Solomon argues against implementing a hierarchy of methods. Instead, she thinks that 'the [epistemological] strengths and weaknesses of each of the methods should be considered when exploring results further' (p. 229). Solomon contributes significantly to this task by considering in great detail the strengths and weaknesses of the various methods. She acknowledges that

lot of attention has been paid to evidence-based medicine. However, Solomon also points out that discussion has tended to consider each method in isolation from the others: 'There is no substantial study of the ways in which the different methodologies fit together, react to one another, sometimes disagree with one another, and are negotiated in the context of specific research and clinical questions' (p. 10). She aims to give a more complete account of the strengths and weaknesses of the various methods by paying closer attention to their histories.

I think it is right that this historical approach can be helpful in giving a more complete account. In a number of cases, a novel method has developed in response to the apparent weaknesses of an existing method. As a result, tracking this history is helpful in pointing out the apparent weaknesses and putative strengths of particular methods. Here are some examples: A key idea in translational medicine is to translate basic science research into effective health interventions by relying heavily upon mechanistic reasoning. In contrast, evidence-based medicine is often presented as downplaying the role of mechanistic reasoning in determining the effectiveness of interventions, and instead as playing up the results of comparative clinical studies, such as randomized controlled trials (pp. 105–32). Solomon thinks that translational medicine is a response to one of the weaknesses of evidence-based medicine (pp. 155–77). In particular, she thinks that downplaying mechanistic reasoning is problematic, because such reasoning is important in proposing hypotheses about effective interventions (pp. 124–6). If this account is on the right lines, it is plausible that translational medicine has putative strengths in proposing hypotheses about health interventions, at least relative to evidence-based medicine.

In turn, Solomon presents narrative medicine as a response to a weakness of both translational and evidence-based medicine, namely, that they do not pay enough attention to the therapeutic role of the relationship between the physician and the individual patient (pp. 192–5). Narrative medicine claims that narrative competences such as empathy are essential in effectively treating a patient as an individual. Again, if this account is on the right lines, narrative medicine promises to have strengths as a result of more appropriately attending to the relationship between the physician and the individual patient, at least relative to translational and evidence-based medicine. (Solomon also offers criticisms of narrative medicine, for example, that narrative medicine may focus on the individual at the expense of the social determinants of health (pp. 195–204).)

A significant portion of the book focuses on medical consensus conferences. The received view of such conferences is that they involve a group of experts or semi-experts engaging in rational group deliberation in order to develop a consensus and thereby helping to resolve a medical controversy. These controversies typically concern whether some intervention is effective at bringing about a particular patient-relevant health outcome, for example, whether a particular anti-hypertensive drug is most effective for reducing the risk of cardiovascular problems. Against this received view, Solomon argues that consensus conferences do not play a role in helping to resolve a given medical controversy. In support, she gives a detailed history of the medical consensus conference movement to argue that consensus conferences have tended to take place after the relevant medical controversy had been resolved. She argues that this is most explicit in the age of evidence-based medicine, where conferences are provided with a report of evidence-based results that seems to predetermine the conclusions of the conference (pp. 48–54). Her alternative explanation of the role of consensus conferences is that they are social-epistemic rituals designed only to communicate evidence-based results in an authoritative manner (p. 83). She does not think that this is a disappointment

controversy. To settle a controversy, they instead attempt to gather further evidence (pp. 86–90).

However, it is not clear that an evidence report alone is always enough to resolve the controversy in these medical consensus conferences. This is perhaps made most explicit by those cases in which consensus conferences on the same medical controversy give conflicting results (pp. 70–1). The problem is that the evidence report itself is often controversial, because there may be disagreement about the information that has been included in the report. Arguably, even the most apparently objective evidence reports—namely, those involving meta-analyses—may nonetheless be considered controversial due to disagreements about which studies are included and disagreements about how the included studies should be weighted (Stegenga [2011]). In turn, this may lead to disagreement about the conclusions licensed by the evidence report and, in particular, about external validity: whether the evidence-based results are transferable to the distinct population under consideration. An evidence report will typically consist only of the results of comparative clinical studies, studies which may have narrowly defined exclusion criteria (p. 112). It may be that these studies help to show that a particular anti-hypertensive drug is effective in certain populations; however, disagreement may persist since there is evidence that different pro-hypertensive mechanisms are at work in different ethnic groups (Clarke *et al.* [2014], p. 347). It may be that rational group deliberation is a good way of settling this sort of disagreement about external validity.

This suggests an alternative explanation of the fact that consensus conferences are still held in an age of evidence-based medicine: it may be that the members of the conference panel can help to resolve a controversy by coming to a consensus regarding the external validity of the reported evidence-based results. This proposal also seems to make sense of the fact that consensus conferences are often held on the same medical controversy in different locations (p. 74). Against this proposal, Solomon thinks that ‘the panel format [...] is not especially well suited to tasks of “extrapolation” or determination of external validity, which can be done more easily by the domain experts involved with producing the evidence synthesis’ (p. 71). However, it looks like the truth of this claim depends on a number of features particular to a given consensus conference. Elsewhere, Solomon argues that determining external validity requires a good amount of background knowledge (pp. 140–8), and she points out that consensus conferences may consist of panel members with exactly the relevant background knowledge (pp. 66–7). Thus it may be that a given conference consists of panel members with greater expertise in determining external validity. Solomon provides a detailed discussion of the differences between particular consensus conferences (pp. 63–83). However, more could have been said about whether these differences have a bearing on the inability of a consensus conference to better determine external validity. As it stands, it is not exactly clear why the determination of external validity is best performed by those putting together the evidence report, rather than a group of experts deliberating on the evidence in the report.

Solomon may think that group deliberation cannot reliably help determine external validity because it is susceptible to a variety of biases (pp. 96–100). However, even if this susceptibility to bias is problematic for group deliberation, it may be that group deliberation appropriately constrained by a report of evidence-based results is less problematic, because it is then less susceptible to the variety of biases. This is a place where more could have been said about the epistemological strengths and weaknesses of combinations of methods, in this case, the combination of consensus conferences and the methods of evidence-based medicine. This is important because the strengths of one method may help to address the weaknesses of another.

32). Solomon makes a useful distinction between mechanistic reasoning and evidence of mechanisms, where mechanistic reasoning is a process of proposing a hypothesis about the effectiveness of an intervention on the basis of evidence of mechanisms (pp. 121–4). She points out that such reasoning has a pretty bad track record, even in cases where the evidence establishes the existence of the relevant mechanisms. The problem is that ‘we could have strong evidence that the mechanisms operate, yet no evidence [...] that a particular proposed therapy will have the desired effect’ (p. 123). As a result, mechanistic reasoning does not provide evidence that might help to establish the effectiveness of an intervention for some health outcome. In other words, there is no evidential role for evidence of mechanisms (p. 124).

I think it is right that on this unqualified account of mechanistic reasoning, it is unlikely that such reasoning would provide much evidence for the effectiveness of a medical intervention. The problem here is that establishing that there is a mechanism by which the intervention makes a difference to the health outcome does little to raise the probability that the intervention will be effective. There may exist further unknown mechanisms by which the intervention cancels out any difference in the health outcome. Here is an example: Although there is a mechanism that links taking a particular contraceptive pill to developing thrombosis, it may be that taking the pill does not make an overall difference to thrombosis; there may exist a further mechanism by which the pill prevents thrombosis, namely, by preventing pregnancy (Hesslow [1976]). This is sometimes called the problem of masking (Illari [2011]). It has been argued that the problem of masking is not a problem for high-quality mechanistic reasoning, where an instance of mechanistic reasoning counts as high quality only if it also makes plausible that there do not exist further masking mechanisms, and it does this through the accumulation of more complete evidence of the relevant mechanisms (Howick [2011a]). Solomon thinks that this proposal is not particularly instructive because it is difficult to determine whether the evidence is sufficiently complete (pp. 122–3). However, there is another way to make it plausible that there are no masking mechanisms, a way that is not considered by Solomon, namely, securing evidence that there is a correlation between the intervention and the health outcome (Illari [2011], pp. 144–8).

According to this proposal, high-quality mechanistic reasoning is a process of coming to believe a hypothesis about the effectiveness of an intervention not only on the basis of evidence of mechanisms, but also some evidence that the intervention is appropriately correlated with the health outcome. It may be that this alternative is overlooked because it is natural to construe the dialectic in terms of a competition between reasoning on the basis of evidence of correlation and reasoning on the basis of evidence of mechanisms. However, proponents of an evidential role for evidence of mechanisms typically see evidence of mechanisms as complementing rather than competing with evidence of correlation (Clarke *et al.* [2014]). In this case, the evidence of correlation helps to address a characteristic weakness of mechanistic reasoning, namely, the problem of masking. As long as evidence of mechanisms is combined with evidence of correlation, it is plausible that a qualified form of mechanistic reasoning can provide evidence that may help to establish the effectiveness of a medical intervention. (Elsewhere, Solomon seems to agree: ‘the more we know about basic and other mechanisms, and the more we know about comparative physiology (of laboratory animals and humans) the more likely we are to make accurate predictions and avoid drug failure by focusing on those interventions with the greatest probability of success’ (p. 175).)

In response, it might be argued that evidence of mechanisms is unnecessary as soon as there is evidence of correlation, provided by comparative clinical studies. Indeed, Solomon maintains that ‘health care

outcomes' (p. 117). However, it does not seem quite right that the effectiveness of an intervention can be established simply by establishing an appropriate correlation between the intervention and the relevant health outcome. In particular, it is not enough to believe a hypothesis about the effectiveness of an intervention on the basis of having established only that the intervention is appropriately correlated with the health outcome. This is because there are alternative, non-causal explanations of this correlation. An observational study may establish a correlation between hormone replacement therapy and lower rates of coronary heart disease, but this is not sufficient to establish that hormone replacement therapy caused the lower rates of coronary heart disease; for example, the women receiving hormone replacement therapy may have been generally healthier than those who did not receive the therapy (Howick [2011b], pp. 40–2). Of course, a correlation established by a randomized trial may be a more appropriate basis for establishing the effectiveness of an intervention. But even a randomized trial may establish a non-causal correlation. For example, a correlation between retroactive, intercessory prayer and shorter duration in hospital recovering from bloodstream infection was established by randomized trial (Leibovici [2001]). This correlation may rightly be deemed as spurious because the evidence suggests that there is no plausible causal mechanism to explain the correlation. In this case, it looks like evidence of mechanisms is complementing evidence of correlation. Unfortunately, Solomon does not provide much discussion on this point because she thinks that 'the precise conditions under which causation can be inferred from correlation are not of importance to the present discussion' (p. 117). I think that this is a mistake since it has been argued that evidence of mechanisms is important precisely because it helps to distinguish causation from mere correlation, by helping to rule out non-causal explanations of a correlation, confounding, bias, and chance (Russo and Williamson [2007]). Again, a discussion of the strengths and weaknesses of combinations of methods may have helped here.

In spite of these reservations, this is an impressive and important book. It is full of useful examples and case studies from medicine that help to make original points. And many of these points seem to me to be correct: in medicine, there is a plurality of methods, and these methods should be judged on their merits rather than by appealing to a hierarchy of methods. Solomon's historical approach to these issues is a productive way to proceed. It furthers the debate about the strengths and weaknesses of the various methods used in medical research and practice by providing a more complete picture of the relationships between these methods. It is a must-read for anyone interested in the epistemology of medicine or philosophy of science more generally. It is also beautifully written, with clear and informative prose, and concise summaries of the conclusions at regular intervals. My criticism in this review has been that the approach advocated in the book might have been carried out more fully through the evaluation of the epistemological strengths and weaknesses of combinations of methods. It might be objected that this criticism is misplaced: as long as the methods are evaluated individually, these evaluations can be taken together to provide an overall assessment of a combination of the methods. The problem with this approach is that a combination of methods may be greater than the sum of its parts, because the weaknesses of one method may be addressed by the strengths of another. An approach that overlooks this point is more likely to give an inappropriately unfavourable evaluation of the individual methods, such as consensus conferences and mechanistic reasoning.

Michael Wilde
Department of Philosophy

Acknowledgements

Thanks to Brendan Clarke, Donald Gillies, Phyllis Illari, Mike Kelly, Christian Wallmann, and Jon Williamson.

References

Clarke, B., Gillies, D., Illari, P., Russo, F. and Williamson, J. [2014]: 'Mechanisms and the Evidence Hierarchy', *Topoi*, **33**, pp. 339–60.

Hesslow, G. [1976]: 'Two Notes on the Probabilistic Approach to Causality', *Philosophy of Science*, **43**, pp. 290–2.

Howick, J. [2011a]: 'Exposing the Vanities—and a Qualified Defense—of Mechanistic Reasoning in Health Care Decision Making', *Philosophy of Science*, **78**, pp. 926–40.

Howick, J. [2011b]: *The Philosophy of Evidence-Based Medicine*, Chichester: BMJ Books.

Illari, P. [2011]: 'Mechanistic Evidence: Disambiguating the Russo–Williamson thesis', *International Studies in the Philosophy of Science*, **25**, pp. 139–57.

Leibovici, L. [2001]: 'Effects of Remote, Retroactive Intercessory Prayer on Outcomes in Patients with Bloodstream Infection: Randomised Controlled Trial', *British Medical Journal*, **323**, p. 1450.

Russo, F. and Williamson, J. [2007]: 'Interpreting Causality in the Health Sciences', *International Studies in the Philosophy of Science*, **21**, pp. 157–70.

Stegenga, J. [2011]: 'Is Meta-analysis the Platinum Standard of Evidence?', *Studies in History and Philosophy of Biological and Biomedical Sciences*, **42**, pp. 497–507.

Acknowledgements

Thanks to Brendan Clarke, Donald Gillies, Phyllis Illari, Mike Kelly, Christian Wallmann, and Jon Williamson.

Michael Wilde
Department of Philosophy

References

Clarke, B., Gillies, D., Illari, P., Russo, F. and Williamson, J. [2014]: 'Mechanisms and the Evidence Hierarchy', *Topoi*, **33**, pp. 339–60.

Hesslow, G. [1976]: 'Two Notes on the Probabilistic Approach to Causality', *Philosophy of Science*, **43**, pp. 290–2.

Howick, J. [2011a]: 'Exposing the Vanities—and a Qualified Defense—of Mechanistic Reasoning in Health Care Decision Making', *Philosophy of Science*, **78**, pp. 926–40.

Howick, J. [2011b]: *The Philosophy of Evidence-Based Medicine*, Chichester: BMJ Books.

Illari, P. [2011]: 'Mechanistic Evidence: Disambiguating the Russo–Williamson thesis', *International Studies in the Philosophy of Science*, **25**, pp. 139–57.

Leibovici, L. [2001]: 'Effects of Remote, Retroactive Intercessory Prayer on Outcomes in Patients with Bloodstream Infection: Randomised Controlled Trial', *British Medical Journal*, **323**, p. 1450.

Russo, F. and Williamson, J. [2007]: 'Interpreting Causality in the Health Sciences', *International Studies in the Philosophy of Science*, **21**, pp. 157–70.

Stegenga, J. [2011]: 'Is Meta-analysis the Platinum Standard of Evidence?', *Studies in History and Philosophy of Biological and Biomedical Sciences*, **42**, pp. 497–507.

Share this review

