Studies in History and Philosophy of Biological and Biomedical Sciences xxx (2015) 1–4

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

journal homepage: www.elsevier.com/locate/shpsc

Introduction

Prediction in epidemiology and medicine

Jonathan Fuller [a], Alex Broadbent [b], Luis J. Flores [c]

[a] Faculty of Medicine, University of Toronto, Canada
[b] Department of Philosophy, University of Johannesburg, South Africa
[c] Department of Philosophy, King's College London, United Kingdom

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Prediction
Epidemiology
Medicine
Philosophy of epidemiology
Philosophy of medicine

Here's a prediction: if Sam doesn't smoke, Sam probably won't have a stroke. Here's a similar sounding claim: if Sam *does* smoke, Sam probably still won't have a stroke. Finally: smoking will cause strokes in a large population of individuals similar to Sam (and including Sam). Each of these scientific predictions might be inferred from different components of the same body of epidemiological and experimental evidence. The first prediction might be inferred from the low rate of strokes in a non-smoking population; the second prediction from the low (but higher) rate of strokes in a smoking population; the last from the ensemble of research evidence supporting the association between smoking and strokes. All three inferences will likely consider other information about the study populations and the Sam-like population, including similarities and dissimilarities.

In order to understand the last prediction, the prediction that smoking will cause strokes in a Sam-like population, we cannot rely only on the vast causal inference literature in the philosophy of science. Studying causal inference might help us understand why the epidemiologists concluded from certain research data (including an association between smoking and strokes) that smoking *causes* strokes, and also whether they were justified in their conclusion. But it cannot help us assess the *predictive* inference(s) that led to the conclusion that smoking *will cause* strokes in a Sam-like population. The predictive inference could not have involved studying an association between smoking and strokes in the Sam-like population because the cases of stroke we predict in the Sam-like population have not yet occurred.

Nor can we understand any of the three prediction claims simply by labelling them as explananda, considering the epidemiologic evidence as an explanans, and straightforwardly applying our favourite account of explanation from the philosophy of science. The fact that most smokers in the study did not have a stroke does not explain why Sam will not have a stroke, even though it might help us to predict that Sam will not have one. There is much more to understanding prediction than that.

Compared to topics like causal inference and explanation, prediction has received scarce attention in the philosophy of science (Douglas, 2009).[1] Yet it is a central activity in many scientific

[1] To be more precise, the attention that has been given to prediction has focused primarily on its relation to other concepts of interest. For example, there was a debate in the mid-Twentieth century about the relation of prediction to explanation, and whether it was symmetrical, as the deductive-nomological model of explanation implies (Canfield & Lehrer, 1961; Hanna, 1969; Hempel & Oppenheim, 1948; Kim, 1964; Rescher, 1958; Rescher, 1963; Scheffler, 1957; Suchting, 1967). There is a live contemporary debate on the significance of predictive success for scientific realism (Barnes, 2008; Barnes, 2014; Carrier, 2014; Menke, 2014; Laudan, 1981; Lawson, 1985; Lipton, 2004, 2005; Mayo, 2014; Schindler, 2014; Schurz, 2014; Van Fraassen, 1980; Votsis, 2014; Worrall, 2014). Nicholas Rescher has offered a book-length treatment of prediction, and one of the present authors has published discussions and a theory of prediction (Broadbent, 2011, 2013, pp. 81–114). These exceptions notwithstanding, it is fair to say that the topic has not attracted sustained attention in its own right in the philosophical literature.

http://dx.doi.org/10.1016/j.shpsc.2015.07.001
1369-8486/

domains, especially those in which interventions are studied or implemented. In public policy, we wish to know which policies will work and how effective they will be, and the social sciences often provide the research needed to answer these important questions. Likewise, in medicine and in public health we must decide which dangerous exposures to avoid and which beneficial treatments to use; in this case the health sciences, including epidemiology, supply the data along with some of the inferential machinery.

Epidemiology is the study of the distribution and determinants of disease and other health states in human populations by means of group comparisons for the purpose of improving population health (Broadbent, 2013, p. 1). This definition does not make obvious the distinctive features of epidemiology, which also make it philosophically interesting. Epidemiology is a detective-like science, approaching its topic piecemeal and in a pragmatic fashion—quite unlike economics, which one might imagine to be similar in character. Unlike economics, epidemiology does not formulate laws; it does not even formulate theories in the sense in which philosophers use the word. Rather than seeking a comprehensive and unified description of its domain, it is focused on particular problems (e.g. does living next to electricity lines increase the incidence of some kinds of cancer?). This makes it hard to apply the terms of many traditional philosophical debates to epidemiology.

The most pressing concerns in the philosophy of epidemiology concern causation, explanation, and—importantly—prediction. Epidemiologists often rely on 'observational' studies, meaning non-experimental studies (the investigator does not conduct an experiment, but rather 'observes' a large number of people, and makes inferences on the basis of these passive observations). The inability to intervene in many contexts involving human subjects raises conceptual questions about the meaningfulness and predictive usefulness of estimates that are arrived at on the basis of observational studies. Because the population in an observational study is a group of diverse people in diverse circumstances, questions can arise as to whether—for example—it is meaningful to estimate the excess mortality attributable to obesity. In particular, it is not clear that such an estimate allows us to predict the effect of reducing obesity in a population, since the various ways in which we might reduce obesity (exercise, diet, smoking, amphetamines) may produce different outcomes (Hernán & Taubman, 2008). Furthermore, epidemiology runs into philosophical difficulties concerning the role of causal knowledge in the classification of diseases. Many of the diseases of contemporary interest appear to be 'multifactorial', meaning that we cannot define them in relation to a specific cause or set of causes (Broadbent, 2009, 2013, p. 1453). This raises a question as to whether too much is claimed or hoped for sound causal inference. Identifying a risk factor—part of the multifactorial constellation of causes—and establishing that it is causal turns out not to be as useful for public health or clinical intervention as one might imagine. This calls for conceptual and methodological work on topics *downstream* of causal inference, such as how to use causal knowledge to predict.

We should pause for a moment to distinguish epidemiology from 'clinical epidemiology' and evidence-based medicine (EBM). While epidemiology is mainly oriented towards population and public health, clinical epidemiology is mainly interested in the health of individual patients as a science for clinical medicine (Bluhm & Borgerson, 2011). Meanwhile, EBM is the application of clinical research evidence to the care of patients. Although all three disciplines are reliant on a common set of research methods (population studies), clinical epidemiology and evidence-based medicine represent a different intellectual tradition compared to epidemiology. For instance, while epidemiology is especially reliant on observational studies, clinical epidemiology and EBM often emphasize experimental human studies, particularly randomized clinical trials.

EBM and clinical medicine raise their own interesting philosophical problems, which are explored in the philosophy of medicine. By developing explicit standards to guide diagnostic, prognostic and therapeutic inferences, EBM has prompted an increasing number of philosophical incursions into the largely unexplored terrain of clinical prediction. EBM's hierarchies of evidence—essentially hierarchies of evidence for clinical predictions—have fuelled philosophical discussion about the relative soundness of prediction activities based on qualitatively different kinds of medical evidence (Clarke, Gillies, Illari, Russo, & Williamson, 2014; Howick, 2011; Worrall, 2002). Philosophers have drawn attention to the need to study medical prediction activities in more detail, revealing that they often involve unexamined steps (Cartwright, 2012; Fuller & Flores, 2015). One of these activities is the extrapolation of results from medical research studies to a population or patient of interest. The problem of extrapolation or external validity is a general problem, but an especially venerable one in medicine. The benefits of philosophical attention to these problems are bidirectional. Philosophers can bring clarity and insight to genuine problems in medical research and practice, while examining philosophical problems that are of interest to philosophy of science.

In particular, the topic of scientific prediction throws new light on well-studied philosophical problems like induction and explanation, but raises novel questions as well. What is a prediction? What makes a good prediction activity; should we evaluate prediction activities according to how reliable they are, or how well justified they are? Can we identify a set of exclusive and exhaustive types of scientific prediction, such as 'statistical prediction' and 'theoretical prediction'? Are the types of prediction used in scientific practice discipline-specific or discipline-nonspecific?

In epidemiology, how can we derive predictions about future distributions of disease or the results of an intervention in a population? Should we be guided by standard methods and inference schemes, or by more general epistemic principles (and what might these methods, schemes or principles be)? In clinical medicine, how can we make predictions about individual patients from population data? What sort of evidence might various sources of knowledge, including knowledge of biological mechanisms and knowledge from population research, provide for medical prediction? Are some kinds of evidence stronger than others? And how might we evaluate the strength of particular kinds of evidence?

The papers in this special section illustrate the philosophical novelty and practical importance of examining prediction in epidemiology and medicine, and of increased philosophical attention to epidemiology and medicine in general. Jonathan Fuller and Luis Flores (2015) represent the Risk Generalization-Particularization (Risk GP) Model, which they argue is the standard model of prediction in contemporary medicine. The model involves generalizing or extrapolating the risk of the outcome or the effect size from a population study, and subsequently transforming (a) the risk or (b) the effect size into (a) the probability or (b) the change in probability of an outcome for a particular patient. Fuller and Flores discuss several well-known problems in the context of their model, including the problem of extrapolation and the reference class problem. One might wonder whether the authors have oversimplified medical prediction by describing a univocal standard. The authors themselves discuss many other models of prediction that are sometimes used in practice, including mechanistic reasoning and induction from personal clinical experience. Though the authors present these other models as alternatives, it may be worthwhile to ask how various methods of medical prediction could be complementary or might provide mutual confirmation for a single prediction claim.

Please cite this article in press as: Fuller, J., et al., Prediction in epidemiology and medicine, Studies in History and Philosophy of Biological and Biomedical Sciences (2015), http://dx.doi.org/10.1016/j.shpsc.2015.07.001

Jacob Stegenga (2015a) discusses problems in measuring the effectiveness of medical interventions, building on another paper in this issue of the journal (2015b) in which he is concerned with conceptual issues related to medical effectiveness. In his conceptual paper (2015b), Stegenga argues that an effective medical intervention must treat disease (versus, for instance, a non-disease bodily state or a social ill) or else it is not a medical intervention. He argues for a hybridist account of disease, which includes both a constitutive causal (natural) component and a harmful (normative) component. He further argues that an effective medical intervention is an intervention that (successfully) targets either the causal basis of disease or the normative basis of disease. Stegenga also places a little-discussed necessary condition on effectiveness: an effective medical condition must improve health. As the notion of health is just as controversial as the notion of disease, his account leaves an important question unanswered: what does it mean for an effective medical intervention to improve health?

In his measurement paper (2015a), Stegenga discusses three epistemic challenges in measuring effectiveness: the choice of measuring instrument (e.g. a subjective symptom rating scale), the choice of outcome measure (measure of association), and the problem of extrapolation from the study to a target population. His paper illustrates that predicting effectiveness in a target population requires addressing prior problems that are not so straightforward. Stegenga argues that contemporary research fails to adequately address these challenges, which contributes to an overestimation of the effectiveness of our medical interventions. He makes a strong case for the existence of a sort of generalised bias towards overestimating predicted effectiveness. Perhaps the main question the paper leaves unanswered concerns the origin of this generalised bias. Is it driven solely by extrinsic factors—such as financial interest? Or is there a problem more intrinsic to scientific methods and inferences in this domain that tends to make predictions overestimate effectiveness no matter what the social and political context? In other words, do medical research methods themselves tend to inflate effectiveness predictions?

Lastly, Alex Broadbent (2015) reports on and critiques a methodological movement (he says "revolution") in contemporary epidemiology, which he calls the *Potential Outcomes Approach* (POA). The elements of this movement are not new, but the social dynamics around it are such that it challenges the established epidemiological paradigm of causal inference. That paradigm, exemplified by the history of smoking and lung cancer, emphasises the use of multiple evidentiary sources collated in a qualitative judgement. The POA instead focuses on a set of formal methods, thus switching the focus to single studies, and insists that causal questions be framed as explicitly as possible in reference to contrast classes. Broadbent identifies four conceptual commitments that he argues underpin the POA, and argues against them. One of these theses is the implicit commitment that causal knowledge entails (suffices for) predictive ability. Broadbent argues that prediction is typically a separate step from causal inference. The most obvious critique of Broadbent's paper is a charge of straw man, a charge he is at pains to rebut, but which may persist nonetheless. How much 'rational reconstruction' of the various writings and sayings of various epidemiologists is permissible, and how does one go about assessing whether these writings and sayings amount to a univocal movement? Does anyone adhere, whether knowingly or unwittingly, to the tenets of the POA as Broadbent identifies them? Since the chief proponents of the POA have not formulated it as an explicit position, these are difficult questions to answer fairly.

Together, these papers form an interesting picture of prediction in epidemiology and medicine. In physics, prediction is often portrayed as the tumbling out of a hypothesis from a theory. In comparison, prediction in epidemiology and medicine looks much more like a complicated assembly line that starts with crude data from some individuals and winds up with a refined claim about what will happen to other individuals (collectively or individually). The assembly line is perhaps longer than we appreciate; along the way choices are made about what to measure, how to conceptualize 'causation', 'effectiveness' and 'prediction', and how to infer causal claims, effectiveness claims and prediction claims. Quality checks at various points along the line reveal deep philosophical problems with the way we understand causation and intervention, with the way we measure effectiveness and with the way we extrapolate and probabilize study results.

Theory- or model-construction and explanation are the main goals in sciences like physics, biology and psychology. Theorizing and explaining are often *not* the primary objects in epidemiology or medicine. In order to understand epidemiologic and medical practice, philosophers must adjust their focus to new topics and new questions. 'Prediction' is a term broad enough to encompass many activities and problems in philosophy of epidemiology and philosophy of medicine, from extrapolation to effectiveness and mechanistic reasoning. These problems are each held together—and as a group hang together—in general ways, which makes them ideal candidates for the most general analysis; in other words, for philosophical study.

Acknowledgements

This special section arose out of a workshop titled Prediction in Epidemiology and Healthcare that took place at King's College London on 20 June 2014. Thanks for productive discussion and inspiration go to the workshop attendees, as well as the speakers: Maël Lemoine, Barbara Osimani, Federica Russo, Elselijn Kingma, Jacob Stegenga, Jeremy Howick, Jonathan Fuller, Luis Flores and Alex Broadbent. Further thanks are owed to the workshop's cosponsors: the British Society for the Philosophy of Science and the Centre for the Humanities and Health at King's College London.

References

Barnes, E. (2008). *The paradox of predictivism*. Cambridge: Cambridge University Press.
Barnes, E. (2014). The roots of predictivism. *Studies in History and Philosophy of Science, 45*(1), 46–53.
Bluhm, R., & Borgerson, K. (2011). Evidence-based medicine. In D. M. Gabbay, P. Thagard, & J. Woods (Eds.), *Handbook of the philosophy of science: Volume 16: Philosophy of medicine* (pp. 203–238). Amsterdam: Elsevier.
Broadbent, A. (2009). Causation and models of disease in epidemiology. *Studies in History and Philosophy of Biological and Biomedical Sciences, 40*, 302–311.
Broadbent, A. (2011). What could possibly go wrong? – A heuristic for predicting population health outcomes of interventions. *Preventive Medicine, 53*(4–5), 256–259.
Broadbent, A. (2013). *Philosophy of epidemiology. New directions in the philosophy of science*. London and New York: Palgrave Macmillan.
Broadbent, A. (2015). Causation and prediction in epidemiology: A guide to the "methodological revolution". *Studies in History and Philosophy of Biological and Biomedical Sciences* (in this issue)
Canfield, J., & Lehrer, K. (1961). A note on prediction and deduction. *Philosophy of Science, 28*(2), 204–208.
Carrier, M. (2014). Prediction in context: On the comparative epistemic merit of predictive success. *Studies in History and Philosophy of Science, 45*(1), 97–102.
Cartwright, N. (2012). Will this policy work for you? Predicting effectiveness better: How philosophy helps. *Philosophy of Science, 79*, 973–989.
Clarke, B., Gillies, D., Illari, P. M., Russo, F., & Williamson, J. (2014). Mechanisms and the evidence hierarchy. *Topoi, 33*, 339–360.
Douglas, H. (2009). Reintroducing prediction to explanation. *Philosophy of Science, 76*, 444–463.
Fuller, J., & Flores, L. J. (2015). The risk GP model: The standard model of prediction in medicine. *Studies in History and Philosophy of Biological and Biomedical Sciences* (in this issue)
Hanna, J. F. (1969). Explanation, prediction, description, and information theory. *Synthese, 20*(3), 308–334.
Hempel, C., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*(2), 135–175.

Hernán, M. A., & Taubman, S. L. (2008). Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *International Journal of Obesity, 32*, S8–S14.
Howick, J. (2011). *The philosophy of evidence-based medicine*. Chirchester: Wiley-Blackwell.
Kim, J. (1964). Inference, explanation, and prediction. *The Journal of Philosophy, 61*(12), 360–368.
Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science, 48*(1), 19–49.
Lawson, T. (1985). The context of prediction (and the paradox of confirmation). *The British Journal for the Philosophy of Science, 36*(4), 393–407.
Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London and New York: Routledge.
Lipton, P. (2005). Testing hypotheses: Prediction and prejudice. *Science, 307*(5707), 219–221.
Mayo, D. (2014). Some surprising facts about (the problem of) surprising facts. *Studies in History and Philosophy of Science, 45*(1), 79–86.
Menke, C. (2014). Does the miracle argument embody a base rate fallacy? *Studies in History and Philosophy of Science, 45*(1), 103–108.
Rescher, N. (1958). On prediction and explanation. *The British Journal for the Philosophy of Science, 8*(32), 281–290.
Rescher, N. (1963). Discrete state systems, Markov chains, and problems in the theory of scientific explanation and prediction. *Philosophy of Science, 30*(4), 325–345.
Scheffler, I. (1957). Explanation, prediction, and abstraction. *The British Journal for the Philosophy of Science, 7*(28), 293–309.
Schindler, S. (2014). Novelty, coherence, and Mendeleev's periodic table. *Studies in History and Philosophy of Science, 45*(1), 62–69.
Schurz, G. (2014). Bayesian pseudo-confirmation, use-novelty, and genuine confirmation. *Studies in History and Philosophy of Science, 45*(1), 87–96.
Stegenga, J. (2015a). Measuring effectiveness. *Studies in History and Philosophy of Biological and Biomedical Sciences* (in this issue)
Stegenga, J. (2015b). Effectiveness of medical interventions. *Studies in History and Philosophy of Biological and Biomedical Sciences* (in this issue)
Suchting, W. A. (1967). Deductive explanation and prediction revisited. *Philosophy of Science, 34*(1), 41–52.
Van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
Votsis, I. (2014). Objectivity in confirmation: Post hoc monsters and novel predictions. *Studies in History and Philosophy of Science, 45*(1), 70–78.
Worrall, J. (2002). What evidence in evidence-based medicine. *Philosophy of Science, 69*, S316–S330.
Worrall, J. (2014). Prediction and accommodation revisited. *Studies in History and Philosophy of Science, 45*(1), 54–61.

Please cite this article in press as: Fuller, J., et al., Prediction in epidemiology and medicine, Studies in History and Philosophy of Biological and Biomedical Sciences (2015), http://dx.doi.org/10.1016/j.shpsc.2015.07.001