Studies in History and Philosophy of Biological and Biomedical Sciences 54 (2015) 49–61

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

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

The Risk GP Model: The standard model of prediction in medicine

Jonathan Fuller [a,*], Luis J. Flores [b]

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

ARTICLE INFO

Article history:
Available online 26 July 2015

Keywords:
Prediction
Epidemiology
Medicine
Risk
Extrapolation
Probability

ABSTRACT

With the ascent of modern epidemiology in the Twentieth Century came a new standard model of prediction in public health and clinical medicine. In this article, we describe the structure of the model. The standard model uses epidemiological measures—most commonly, risk measures—to predict outcomes (prognosis) and effect sizes (treatment) in a patient population that can then be transformed into probabilities for individual patients. In the first step, a risk measure in a study population is *generalized* or extrapolated to a target population. In the second step, the risk measure is *particularized* or transformed to yield probabilistic information relevant to a patient from the target population. Hence, we call the approach the Risk Generalization–Particularization (Risk GP) Model. There are serious problems at both stages, especially with the extent to which the required assumptions will hold and the extent to which we have evidence for the assumptions. Given that there are other models of prediction that use different assumptions, we should not inflexibly commit ourselves to one standard model. Instead, model pluralism should be standard in medical prediction.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Biological and Biomedical Sciences*

1. Introduction

Predictions are central to medical practice. Doctors want to know what will happen to the patient in the future given their present condition (prognosis), and how treatment or prevention might alter the natural course of events (intervention). But is there a standard model of prediction in medicine, a dominant approach in which trainees are schooled and according to which doctors practice? What we are after is a prediction scheme similar to other models of prediction in the philosophy of science, the most classic and well-known of which is the Deductive-Nomological Model (DN Model) of Carl Hempel and Paul Oppenheim (Hempel & Oppenheim, 1948).[1]

Such an idealization is not to be found in medical textbooks. In fact, textbooks tend not to use the term 'prediction' to label a major category of clinical inference, but instead divide inferential activities into the traditional medical categories of diagnosis, prognosis, therapy and harm (Guyatt, Rennie, Meade, & Cook, 2008). Yet prognostic, therapeutic and harm-related inferences typically involve predictions, hypotheses about future outcomes. Even diagnosis can be conceptualized as a predictive activity; clinical textbooks speak of "clinical prediction rules" for diagnosis and the "positive predictive value" of a diagnostic test (Guyatt et al., 2008, pp. 491–505; Fuller, Sankar, & Upshur, 2013, pp. 580). Diagnosis is predictive in the wider sense of inferring an outcome that is not definitively known (i.e. the presence of a particular disease). It will be profitable to examine the shared structure of these distinct types of clinical inference.

An important clue to the existence of a standard model is that there seems to be a common target of several critiques of medical prediction, some of which will be explored in Sections 4 through 6. However, the received model lacks an explicit philosophical reconstruction—or a reconstruction of any sort, for that matter. Without a clear representation, it remains a nebulous target.

Abbreviations: GP, Generalization–Particularization; EBM, evidence-based medicine; AR, absolute risk; CVD, cardiovascular disease; ES, effect size; RR, relative risk; RD, risk difference; NHS, National Health Service.

* Corresponding author.
E-mail address: jonathan.fuller@mail.utoronto.ca (J. Fuller).

[1] Another notable model of prediction is found more recently in the work of Spirtes, Glymour, and Scheines (2000). Their approach uses directed graphical modelling to predict the probability distribution resulting from a targeted intervention on one or more variables.

http://dx.doi.org/10.1016/j.shpsc.2015.06.006
1369-8486/

Here, we reconstruct and examine the Risk Generalization–Particularization (Risk GP) Model, the standard model of prediction in medicine. Risk GP is standard in that it represents the dominant prescriptive model in contemporary practice (the gold standard), as well as the model that many practitioners implicitly rely upon when making evidence-based decisions. Risk GP is an epidemiological model, relying centrally on aggregate outcomes in populations. Like the science of epidemiology, the model is relatively new when framed against the long history of medicine, although rational approaches to prediction have been around since at least the time of Hippocrates (460-370 BCE). The Risk GP Model actually consists of two inferences in series: a *generalization* of a risk measure from a study population to a target patient population of interest; and a *particularization*, a transformation of this measure to yield probabilistic information about a patient within the target population.

There are well-known problems at both stages. Most worryingly, the necessary assumptions for generalization and particularization may not hold widely, and even when they do hold, we might not have evidence to warrant them. These problems are not an inevitable challenge for clinical practice, or even for epidemiological predictions, but are peculiar to the Risk GP Model. Of course, most models are imperfect, and their ideal assumptions will sometimes fail to represent reality. Those circumstances demand flexibility; we should not commit ourselves to a one-model-fits-all approach, but should be model pluralists instead.

2. Models of prediction in historical perspective

A few distinctions will be useful upfront. Alex Broadbent identifies a "process/product" ambiguity in the concepts of prediction. He distinguishes two senses of the term: prediction as a *claim*, and prediction as an *activity* (2013, pp. 86, 89–93). The first sense of 'prediction' is a claim or hypothesis, such as: 'it will rain tomorrow'. The second sense of 'prediction' is an activity or argument, such as: 'following many previous weeks like this one it rained the next day; therefore, it will rain tomorrow'. Prediction activities are inferences involving prediction claims.

In cases like the meteorological prediction just mentioned, prediction activities are inferences with a prediction claim, a definite forecast, as their conclusion. We can call these prediction activities *predictive inferences* to distinguish them from prediction activities that do *not* have a definite prediction claim as their conclusion. For instance, take the inference: 'on 60% of previous weeks like this one it rained the next day; therefore, the probability that it will rain tomorrow is 60%'. The conclusion is not a prediction claim; in asserting it, we are not placing a bet or committing ourselves to the occurrence of some future event. If instead it was clear skies without any approaching storm fronts, the meteorologist would conclude that the probability of rain is low, which is obviously not a prediction that it will rain tomorrow. Yet we might still want to call this statistical inference a 'prediction activity' because the conclusion tells us the probability *of* a prediction claim.[2]

As previously alluded, there are at least two, non-exclusive types of prediction claims in natural language and medical discourse. The more inclusive type encompasses all hypotheses about unknown (unobserved) events or outcomes. It includes diagnostic hypotheses like: 'the patient has heart disease'. Meanwhile, the less inclusive type of prediction claim is a subtype of the former, and includes only hypotheses about the future (e.g. 'the patient will experience a cardiovascular disease event over the next ten years'). In these cases, the outcomes are unknown specifically because they have not yet occurred (Broadbent calls this less inclusive type "narrow prediction" (2013, pp. 93)). As we will see, the standard model can account for predictions in the broader sense. But since prognostic and therapeutic predictions are usually predictions in the narrow sense (they are hypotheses about what will happen in the future to the patient), narrow predictions will be our main focus.[3]

An informative prediction scheme would model both the prediction claims and the associated prediction activities in a given field. The DN Model (Hempel & Oppenheim, 1948) provides a good illustration. Given a physical phenomenon to be explained (the explanandum), we supply the laws of nature and particular facts that jointly entail it (the explanans). To explain why an object is accelerating at a particular rate of 1/2 m/s^2, we can deduce the rate from Newton's Second Law and some initial conditions:

$$\begin{array}{c}\text{Acceleration} = \text{Force/Mass} \\ \underline{\text{Force} = 1\ \text{N, Mass} = 2\ \text{kg}} \\ \text{Acceleration} = 1/2\ \text{m/s}^2\end{array}$$

In the DN scheme, explanation and prediction are symmetrical activities; a prediction is an explanation in which the explanans (above the line) is known but the explanandum (below the line) is not. So the DN Model is also a model of prediction in the wide sense. The entire model represents a prediction activity, while the conclusion represents a prediction claim about an unknown variable.

Unfortunately, the DN Model is of limited use in characterizing modern medical prediction.[4] Few universal laws are used in clinical practice, and there is no unifying theory akin to Newton's Laws. Yet the absence of any grand theory in contemporary medicine is peculiar from a historical perspective. The miasma and contagion theories of disease persisted well into the Nineteenth Century (Gillies, 2005), and from Ancient Greek medicine until the Renaissance, the Hippocratic Theory of the Four Humours, a paradigmatic example of a unifying medical theory, provided a theoretical basis for medicine (Duffin, 2010, pp. 42–45).

In the canonical interpretation of humoural theory, the balance of four bodily fluids—blood, phlegm, black bile and yellow bile—determines a person's state of health or disease. When each of the four humours is in equilibrium, the person is healthy; when any are in disequilibrium, the person is diseased. It follows that reversing disequilibrium in disease restores health. Thus, for bilious patients (with excess bile) and phlegmatic patients (with excess phlegm) the Hippocratic *Affectations* makes the following prescription: "In cleaning, employ medications according to the following principle: when patients are bilious, give medications that clean out bile; when they are phlegmatic, give medications that clean out phlegm" (Potter, 1988, pp. 43). Bloodletting, a therapy commonly used for thousands of years and for a wide range of

[2] Predictive inferences in medicine are typically also 'probabilistic' in that they warrant the definite prediction claim inductively. The essential difference is that a predictive inference concludes that the outcome or event will occur, while this statistical inference merely derives the probability of its occurrence.

[3] Narrow prediction claims include subjunctive conditionals, or 'counterfactuals' ('if T, then O'); specifically, counterfactuals in which the consequent refers to some future outcome or event. In order to decide on a course of action, especially when multiple alternative courses are open, physicians must often predict what will happen before the antecedents of the outcome are established. For instance, what will happen in the future to the patient *if* they are treated?

[4] Hempel (1962) also proposed a statistical model analogous to the DN Model that at first glance might seem more relevant, but because the model is intertwined with Hempel's interpretation of probability we will not discuss the details of the statistical model here.