Causality at the Dawn of the ‘Omics’ Era in Medicine and in Nephrology

Carmine Zoccali, Diego Brancaccio, & Marco J. Nathan

Nephrology, Dialysis, Transplantation (Advance Access)

Abstract

Causality is a core concept in medicine. The quantitative determinacy characterizing today’s biomedical science is unprecedented. The assessment of causal relations in human diseases is evolving, and it is therefore fundamental to keep up with the steady pace of theoretical and technological advancements. The exact specification of all causes of pathologies at the individual level, precision medicine, is expected to allow the complete eradication of disease. In this article, we discuss the various conceptualizations of causation that are at play in the context of randomized clinical trials and observational studies. Genomics, proteomics, metabolomics and epigenetics can now produce the precise knowledge we need for 21st century medicine. New conceptions of causality are needed to form the basis of the new precision medicine.

Keywords: causality, dispositionalism, empiricism, Hume, manipulability theory, randomized trial

1 Introduction

Causality, a core concept in the philosophy of science, has been central throughout the history of medicine, from its Epicurean origins to the modern age. Yet, the unprecedented degree of quantitative determinacy characterizing contemporary biomedical science confronts us with fascinating new challenges. The epistemology of causality—the assessment of causal relations, especially in relation to human diseases—is constantly evolving, striving to keep up with the steady pace of theoretical and technological advancements. The holy grail of health research is arguably the development of a true ‘precision medicine,’ i.e. an exact specification of all causes of pathologies at the individual level, triggering, in the
long run, the complete eradication of disease (Collins and Varmus, 2015). The
goal of this essay is to present and discuss various competing conceptualizations
of causation that are at play, more or less explicitly, in the context of random-
ized clinical trials (RCTs)—the core instrument of evidence-based medicine.
Unravelling the ontological assumptions underlying the nature of causation is
important, we maintain, because methodological disputes in medicine are often
a consequence of different conceptions of causality.

2 Randomized Clinical Trials

Randomized experimental interventions, both in vitro and in vivo, and RCTs are
widely considered to be the gold standard for assessing the causes of diseases and
for establishing the efficacy of treatments. Accordingly, contemporary medical re-
search relegates ‘observational’—i.e. non-interventional—studies to a lower
rank on the evidence scale (Guyatt et al., 1995).

So-called evidence-based medical practice strives towards the full integration
of knowledge derived from experimental research—such as RCTs, observational
and mechanistic studies—with parameters such as doctor experience, patient
expectations and economic values set by public and private health government
bodies (Sackett et al., 1996). However, findings in large-scale RCTs and exten-
sive epidemiological surveys often fail to conform to the kind of knowledge re-
quired for the treatment of individual patients (DeMaria, 2008). Indeed, rarely,
if ever, are clinical trials designed to test the full range of the multiple effects
of an intervention, including both beneficial and potentially noxious ones, in-
cited by treatments. The heuristic imperfection of RCTs becomes evident in
notorious examples of drugs such as flecainide, which, despite being highly ef-
fective in controlling some of the main manifestations of a disease (flecainide
controls ventricular ectopic beats in patients with cardiac diseases), turned out
to increase mortality. To make things worse, the apparent benefit of a drug
in the context of clinical trials is often time-limited, i.e. it may not last for
periods longer than the actual duration of the reference trial. The case of drug-
eluting stents for the late risk of thrombosis is a famous example. Furthermore,
trials focus on restricted populations—individuals in a given age range, with
specific disease phenotypes, patients with proteinuria, patients without severe
hypertension, etc.—whereas, in real practice, doctors deal with patients from
a broad range of ethnic and socio-economic backgrounds, with morbidities of
various sorts, and with diseases of varying duration and severity. In short, trials
constitute a precious and irreplaceable background knowledge; however, their
limited and narrow results are often insufficient to tailor treatment at the level
of patients. Developing a more personalized approach to medicine presupposes,
among other things, an account of causality that is geared to- wards the individ-
al, not merely the population. This article considers some aspects of causality
that are important to emphasize in developing the precision medicine of the
future.
3 The Influence of Classical Empiricism

The account of causality typically presupposed in experimental medicine derives, more or less directly, from the framework developed by empiricists such as Hume (1748), which has been—and still is—extremely influential across the natural sciences. To wit, the influential set of criteria for determining causal patterns suggested by Bradford-Hill (1965) is clearly inspired by Hume’s so-called ‘regularity approach,’ which, roughly speaking, purports to reduce attributions of causality to constant conjunctions between events. At the same time, Bradford-Hill notes, the strongest evidence for inferring a true causal connection comes from experimental intervention in conjunction with ‘negative’ feedback, i.e. no change must be observed when the intervention is missing. As we shall see, this constitutes a substantial departure from the classic empiricist framework.

Let us begin by addressing Bradford-Hill’s methodology for inferring causal relations and drawing some connections to contemporary medical practice. In experimental settings, causes are typically determined on the basis of effects produced by intentional manipulation—the administration of a drug or an environmental modification, such as a dietary change—in a sample population of reference. Generally speaking, such experimental interventions can be conceived as the superimposition of a manipulation on a specific natural regularity, which is expected to produce a change in that same regularity, e.g. a modification in the health status of a population of patients or in a relevant pathophysiological process. RCTs are a particularly effective way to assess causal claims in this fashion since, in principle, randomization may eliminate any possible bias and, when properly conducted, it generates matched groups that differ only in the application of the intervention under scrutiny (active group versus control group). In these experimentally controlled settings, any recorded intra-group difference can be safely attributed to the intervention. However, as noted above, an observed positive change in a regularity subsequent to the intervention is not yet sufficient to determine the causal nature of the modification. To strengthen his criterion for causality, Bradford-Hill adds as a further necessary condition a form of negative feedback, namely, that no change must occur when the intervention is not applied. However, as we discuss below, there are various alternative ways to spell out the relevant notion of causality.

Significant as they have become, RCTs are not the magic bullet that can be applied to test all alleged causal relations. In various cases, causality must be assessed through other means, e.g. via purely observational, non-interventionist approaches that register data in large groups of patients (Jager et al., 2007). As an illustration, RCTs cannot be employed, for obvious reasons, to test alleged toxic exposures, which, however, can be validly assessed through observational studies, supplemented by the relevant in vitro and in vivo knowledge. This approach has been used to determine that smoking is a cause of both cardiovascular disease and neoplasia and that dioxin causes skin cancer and various other adverse health effects. Note how the conceptualization of causality adopted, more or less explicitly, in these observational studies accords perfectly to the Humean regularity framework, according to which a causal claim is justified if
the purported cause invariably occurs before the effect, and the two events are reasonably close in time. This conclusion is further strengthened if the effect in question (e.g., a disease) is less frequently observed when the cause under scrutiny (smoking, toxins, etc.) is rare or absent, in a comparable setting. However, while observed correlations and associations may suggest causal relations, they seldom provide decisive evidence in favour of or against them—a point starkly in contrast with Hume’s dictum that there is nothing more to attributing causation than observing constant conjunction of cause and effect.

As a general model of causation, Hume’s radically empiricist approach is overly restrictive. To be sure, it is sometimes possible to infer a bona fide causal relation when two events are strongly correlated, especially when mechanistic studies (in vitro or in vivo) support this causal link, as in the case of smoking and cancer. Accordingly, clinical epidemiologists tend to regard weak correlations as non-causal (Bradford-Hill, 1965). Yet, this purely quantitative (statistical) criterion, which rules out causality when the association between exposure and adverse effects is weak—i.e., when only a small fraction of exposed individuals contracts the disease—may hinder the discovery of genuine causal links. For example, population-based studies suggest that the use of oral contraceptives is associated with incident thrombosis only in a relatively small number of cases: 115 cases per 1000 women per year in contraceptive takers and 53 cases per 1000 women per year in non-takers (Various Authors, 1978). A strict application of the principle of regularity would likely treat the connection as non-causal. However, such a conclusion contrasts with findings in robust genetic and mechanistic studies supporting a causal role of contraceptives in thrombosis at the individual level (Vandenbroucke et al., 1997). All of this shows that, whereas it might be possible to effectively apply Hume’s principle of regularity at the population level, the classic empiricist approach becomes problematic in the case of individuals.

The ‘negative clause,’ according to which in a truly causal relation the effect should not be observed in the absence of the cause, is insufficient, by itself, to save the day. To illustrate, consider a hypothetical study where chronic kidney disease (CKD) patients are randomized to two different anti-hypersensitive drug combinations. Patients treated with these combinations achieve normotension in 70% (first group) and 50% (second group) of cases, respectively. Given this statistically significant difference, the standard conclusion is that the first combination should be preferred to the second one as a treatment for CKD patients. Specifically, the fact that 50% of patients (a significant value!) in the control group achieve the blood pressure (BP) target is used to attribute a causal role to the BP decline in the active group. Notably, in the context of this trial, the apparent effect in the control (second) group remains without any causal underpinning. Furthermore, if we scotomize the control group and consider the first group only, we still have 70% of patients who achieve normotension. Still, causation cannot yet be established.

In sum, modern statistical techniques are a powerful means for controlling confounding factors and emulating RCTs in observational settings, in the real world of population studies and in medical practice (Stel et al., 2013). The
principle of regularity plays an important role in answering questions related to the superiority of a given treatment, compared with placebo or alternative treatments. However, this methodological principle is too restrictive to capture the nature of causation across the board—a well-known philosophical lesson that, we maintain, has significant normative implications for medical practice.

4 Beyond Strict Empiricist Notions of Causality

Philosophers have long been aware of the shortcomings of conceiving of causation in terms of pure regularities and have offered various alternative analyses. In general, the relevant literature distinguishes two families of concepts of causality. Reductive theories purport to describe causal relations in non-causal terms. For instance, probabilistic approaches develop the idea that causes need not be sufficient for their effects; they only need to raise the probabilities of their effects (Eells [1991]). Not all smokers develop lung cancer, and not all instances of lung cancer are caused by smoking; nevertheless, one can legitimately treat smoking as a cause of lung cancer because the former raises the probability of the latter. Another influential reductive account of causality is the so-called counterfactual approach developed, in its modern form, by Lewis [1973]. Counterfactual theories essentially stress the ‘negative clause’ presented above, according to which causes are conditions sine qua non for their effects: what underlies a causal attribution of the form ‘C causes E’ is the counterfactual statement ‘If C had not occurred, E would not have occurred.’ Other strategies do not attempt to eliminate causal jargon tout court. Non-reductive theories of causality presuppose some kind of inherently causal notions—such as interventions, manipulations, interferences, etc.—and employ these causal concepts to construct a general theory of causality. Influential non-reductive approaches to causation include dispositionalism—which essentially treats causes as dispositions, propensities to trigger their effects (Cartwright [1989], Mumford and Anjum [2011])—and the manipulability theory, according to which causes are means for producing their effects. According to the manipulability theory, agents can exploit the link between cause C and effect E as a ‘handle’ for bringing about E, and this can occur only if this relation remains stable under this sort of intervention (Woodward [2003]).

A detailed discussion of the main tenets (and significant shortcomings) of these theories of causation—and the many other variants and alternatives present in the growing relevant literature—lies beyond the scope of this work (for a succinct overview, see Hitchcock [2008]). The important point, for the present purposes, is that there is a whole host of contemporary accounts of causation, which, contrary to Hume’s overly restrictive empiricism, do not presuppose that causal attributions can be reduced to the constant conjunction of events and, therefore, that causes must invariably follow their effects. Despite all their problems and limitations, all the accounts presented above allow us to make better sense of the fact that hypercholesterolaemia disposes towards cardiovascular disease, but not all individuals with high cholesterol levels are affected by
myocardial infarction or by stroke. When the correlation between cause and effect is strong, the causal inference is straightforward, as in the inference from hypercholesterolaemia to myocardial infarction. But whether or not the effect will actually occur typically depends on which other causal factors are involved. Smoking may not cause cancer in a subject who lacks a predisposition to the disease, while it may cause it in another subject with a stronger genetic risk for neoplasia.

Contemporary medical practice is moving away from its overly empiricist origins. In contemporary medicine, it is the individual— as opposed to the population—that takes centre stage. Accordingly, we need accounts of causality that allow causal relations to be observed, tested and assessed in individual patients, not merely at the population level. What all the causal theories mentioned above have in common is a clear-cut departure from the Humean framework, where there is nothing more to causal relations than effects invariably following their causes. Statistically unlikely events caused by specific causal exposures are better captured by tendencies, dispositions, probabilistic and counterfactual inferences than ‘brute’ regularities. Unfortunately, philosophy does not offer a ready-made, one-size-fits-all theory of causation that can be straightforwardly imported into medical practice. The debate over how to best conceptualize causation is still much debated among philosophers. In addition, we should not be too hasty in assuming that there is a single theory that can be applied across the sciences. Perhaps an account of causation that works fine in economics will not shed light on physics, and some insights might shed light on internal medicine but not on neuroscience. However, there is a whole host of philosophical tools at our disposal, which can help tailor a concept of causation that is appropriate for the future of individualized medicine. The classic empiricist approach based on regularities provides an inadequate metaphysical foundation for modern medicine, which aims at profiling exact, sometimes unique, interventions at the level of individual patients. Furthermore, as discussed, there is a growing awareness that classic RCTs face significant limitations in causal epistemology and in producing the detailed knowledge required to address health problems at the individual level.

5 The Dawn of Precision Medicine: Concluding Remarks

Fifteen years after having completed the decoding of the human genome, we are now witnessing an unprecedented progress in both biology and medicine (Green and Guyer [2011]). Genomics, proteomics, metabolomics, epigenetics and various other ‘omics’ fields realize the vision and dream of previous generations of physicians. New discoveries generate unique opportunities for understanding health and disease and for enhancing the treatment of human illnesses. Over the past decade, there has been an exponential acceleration in the invention of new diagnostic and therapeutic options. These innovations are on the verge of
revolutionizing the way we approach the diagnosis and treatment of infectious diseases [Cooms, 2012]. Indeed, the technology for allowing rapid and precise identification of viruses and bacteria and their sensitivities to chemotherapeutic interventions at point-of-care level already exists. Treatments could be started early, preventing the exposure of patients to broad-spectrum antibiotics and eventually curbing today’s disconcerting rates of antibiotic resistance.

Precision medicine is already a reality in oncology. In lung neoplasia, there are novel classifications founded on molecular testing, anatomic criteria and histologic criteria. Genetic markers guide the prescription of highly effective treatments [Kwak and Camidge, 2010]. The integration of genetics with advanced computing and imaging techniques, and breakthrough approaches—such as epigenetics, proteomics and metabolomics—has already expanded the potential of precision medicine, generating new classifications of human disease with improved prognostic and therapeutic implications. To illustrate, the precise molecular diagnosis of multiple endocrine type-2 neoplasia may now be adopted as a decisional basis for prophylactic thyroidectomy and a screening tool for medullary thyroid cancer, hyperparathyroidism and pheochromocytoma in patients harbouring this disease [Moore and Dluhy, 2005]. To focus on nephrological examples, kidney biopsy is presently held as the state-of-the-art test for diagnosing acute rejection. It is likely that in the medium term, the biopsy will be replaced by a safer, non-invasive approach based on the study of urinary cells. Indeed, a signature of just three genes robustly discriminates biopsy specimens showing acute cellular rejection from rejection-free specimens [Suthanthiran et al., 2013].

We are at the dawn of a new era. Theoretical and technological advancements are producing the knowledge we need to found and develop the biomedical science of the 21st century. Precision medicine is now a foreseeable objective for medical practice, giving us the unprecedented possibility to assess causality directly in single cases and to design treatment at the individual level [Collins and Varmus, 2015; Jameson and Longo, 2015]. This precise knowledge will allow for the adoption of simple, immediate trials to test the efficacy of treatments at the individual level and may eventually replace proof of efficacy at the group level, the current pursuit of large-scale RCTs. In order for the great promise of the ‘omics’ sciences to become deployable and affordable in the world of clinical medicine, shifting from populations and probabilities to individuals and exactness, our current empiricist means of assessing causality in medicine will have to be revised to obtain a more solid philosophical foundation for evidence-based practice.
References


