Studies in History and Philosophy of Biol & Biomed Sci xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biol & Biomed Sci

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

Universal etiology, multifactorial diseases and the constitutive model of disease classification

Jonathan Fuller[a,b,*]

[a] Faculty of Medicine, University of Toronto, Canada
[b] Research Associate, African Centre for Epistemology and Philosophy of Science, University of Johannesburg, South Africa

ARTICLE INFO

Keywords:
Disease classification
Multifactorial
Etiology
Monocausal model
Philosophy of medicine

ABSTRACT

Infectious diseases are often said to have a universal etiology, while chronic and noncommunicable diseases are said to be multifactorial in their etiology. It has been argued that the universal etiology of an infectious disease results from its classification using a monocausal disease model. In this article, I will reconstruct the monocausal model and argue that modern 'multifactorial diseases' are not monocausal by definition. 'Multifactorial diseases' are instead defined according to a constitutive disease model. On closer analysis, infectious diseases are also defined using the constitutive model rather than the monocausal model. As a result, our classification models alone cannot explain why infectious diseases have a universal etiology while chronic and noncommunicable diseases lack one. The explanation is instead provided by the Nineteenth Century germ theorists.

1. The causes of disease

Multifactorial thinking pervades modern epidemiology and medicine, from the way we describe modern diseases as having multiple and variable etiology (Krieger, 1994; McMahon, Pugh, & Ipsen, 1960; Susser, 1985) to the way that we measure causal risk factors for diseases and customize medical classification, prognosis and prevention based on those risk factors (WHO, 2005; 2014). Nancy Krieger argues that "notions of multiple causation and multivariate analysis are so commonplace and so embedded in modern epidemiologic reasoning that they hardly merit discussion as a model or as an approach to understanding disease" (1994, pp. 891). As an example of multifactorial thinking, the major modifiable risk factors for cardiovascular disease, including stroke, are: smoking, obesity, physical inactivity, dyslipidemia, hypertension, diet and diabetes mellitus (Hennekens, 2015). Individually causal risk factors are not sufficient for disease (not everyone who smokes has a stroke); nor are they necessary (not everyone who has a stroke smokes).[1]

It is not only cardiovascular diseases like stroke and heart attack that are multifactorial, but also chronic diseases like diabetes and dementia, injuries like bone fracture, and even symptoms like back pain. The rise in prominence of multifactorial diseases is partly explained by medicine's own success in controlling infectious diseases and other acute health conditions (WHO, 2015). People are living longer and are increasingly afflicted with chronic diseases and noncommunicable diseases (NCDs) as they age (WHO, 2011; 2015). Chronic and noncommunicable diseases are now the leading killers worldwide, and are paradigmatically multifactorial in their causation.

The multifactorial etiology of modern ailments only seems noteworthy when set against a historical background. In the late Nineteenth and early Twentieth centuries, the paradigm medical maladies were infectious diseases, which are often described as having a single universal etiology. Particular infectious diseases are caused by a particular germ. The particular germ is even necessary for the particular disease; without variola virus, one cannot contract smallpox.

Alex Broadbent (2009; 2013; 2014) calls this turn-of-the-Twentieth Century understanding of diseases the "monocausal disease model" to emphasize the privileging of one particular cause. In contrast, the model of disease popular among epidemiologists and public health authorities beginning in the second half of the Twentieth Century is a "multifactorial model" that recognizes the contribution of multiple causal risk factors to the development of each type of disease. Broadbent argues that the monocausal model is as much a model of definition as it is a model of discovery. Not only do scientists discover a specific cause of a specific type of disease, they define that specific disease as the disease produced by that specific cause.

Abbreviations: NCD, noncommunicable disease; TB, tuberculosis; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; MTBC, *Mycobacterium tuberculosis* complex; HPV, human papillomavirus; RRP, recurrent respiratory papillomatosis

* University of Toronto, Toronto, Ontario, M5S 1A8, Canada.
E-mail address: jonathan.fuller@mail.utoronto.ca.

[1] The epidemiologist Kenneth Rothman (1976) illustrated this relationship between etiologic factors and disease by using 'causal pie' diagrams. In Rothman's diagrams, complete causal conditions are pies, individual etiologic factors are slices, and alternative pies can cause the same disease.

https://doi.org/10.1016/j.shpsc.2017.11.002
Received 6 March 2017; Received in revised form 1 November 2017; Accepted 3 November 2017

The model of nosology or disease taxonomy that the monocausal ideal is thought to have supplanted is one in which types of disease were defined constitutively, in terms of the components that comprised them.[2] In the early Nineteenth Century, these components were typically symptoms, but by the middle of the Nineteenth Century they often included pathological anatomical lesions (Porter, 2002; Carter, 2003). For the simple reason that different causes can give rise to the same symptoms and lesions, the causes of these diseases were not singular and universal but multiple and variant. From this perspective, it looks like modern epidemiology and medicine have been dragged back to an era when disease etiology was multifarious, complex and unwieldy. The long and eclectic lists of etiological factors in early Nineteenth Century medicine exasperated the famous physician Jacob Henle, who called for the discovery and reporting of causes of disease that were "universal, necessary and sufficient" (Carter, 2003, pp. 25).

According to Codell Carter (2003), the monocausal model, which he calls "the etiological standpoint", came to characterize modern Western medicine beginning in the late Nineteenth Century. The story often goes that a monocausal understanding of infectious diseases allowed for breakthroughs in their treatment and prevention (Carter, 2003; Evans, 1993). Jeremy Greene and colleagues write: "Motivated by breakthroughs in cellular pathology, pathophysiology, and especially bacteriology, doctors increasingly came to see diseases as specific entities, each with its own specific causes, manifested as characteristic syndromes. This new model prompted doctors to seek therapies tailored to the disease and not the patient", a "therapeutic revolution" (2012, pp. 1080).[3] Thus, it is perhaps a letdown that many modern diseases are multifactorial and not monocausal.

In this article, I will reconstruct the monocausal disease model and ask why modern multifactorial diseases refuse to conform to the monocausal pattern. I will argue that multifactorial diseases are defined according to a constitutive model of classification, which explains why they are multifactorial. However, I will propose that – contrary to popular opinion – we can also understand the classification of infectious diseases according to the constitutive model. As an upshot, our classification models cannot fully explain the difference between those diseases with a specific universal environmental cause and those diseases without one; it is an old idea, the germ theory, that partly explains why infectious diseases have a universal etiology.

2. The monocausal ideal

According to Carter, "The etiological standpoint can be characterized by the belief that diseases are best controlled and understood by means of causes and, in particular, by causes that are *natural* (that is, they depend on forces of nature as opposed to the wilful transgression of moral or social norms), *universal* (that is, the same cause is common to every instance of a given disease), and *necessary* (that is, a disease does not occur in the absence of its cause)" (2003, pp. 1). The first criterion, the requirement that the cause of the disease is natural, immediately suggests a strategy for *discovering* a disease's etiology: empirical research, especially research in the natural sciences. Meanwhile, the third criterion, the criterion that the cause of the disease is necessary, suggests a principle for *defining* a disease category: one should define the disease according to the cause that was discovered. (The third criterion implies the second criterion of universality: if a certain cause is necessary for the disease, then that cause will always occur whenever the disease occurs.)

The etiologic standpoint is an ideal that Carter and many other authors believe guides etiologic research and faithfully describes our paradigmatic infectious diseases. Epidemiologist Mervyn Susser argues that Nineteenth Century discoveries by Pasteur and Koch "led to the redefinition and reclassification of many disease entities [disease types] by the criterion of cause … By current definition, tuberculosis is caused by the tubercle bacillus" (1973, pp. 23). Similarly, Rothman claims: "Necessary causes are often identifiable as part of the definition of the effect. For example, … infection with the tubercle bacillus is a necessary cause for tuberculosis" (1976, pp. 588). And philosopher Caroline Whitbeck notes that after the success of the germ theory in the Nineteenth Century "the name of the disease came to reflect the type of entity thought to cause it, the so-called etiologic agent, and etiology soon came to be definitive (i.e., to be regarded as essential) for those diseases for which it was known" (1977, pp. 622).

More recently, Alex Broadbent (2009; 2013; 2014) has referred to this principle of disease classification as the "monocausal model". He too emphasizes that "[t]he special status that the monocausal model offers to certain causes is not an empirical status, but a conceptual one. Certain causes *define* the disease in question" (2013, pp. 156). According to Broadbent's reconstruction, the monocausal model places a necessity requirement on the defining cause: "putative cause C is a cause of every case of disease D" (2003, pp. 150). Adopting this requirement, we can represent the monocausal model as follows:

a is case of disease *D only if* an *E* caused *a*.

In a case of infectious disease or poisoning, *E* refers to a specific etiologic agent (a specific germ or a specific toxin, respectively); in a disease of deficiency, it instead refers to the *absence* of a specific agent like a specific nutrient. The two key features of *E* – and thus of the monocausal model – are *causal specificity* and *causal necessity*. *E* is specific because it refers to one particular kind of causal agent; it cannot refer to a disjunction of several kinds of etiologic agents, or else the disease would not be *mono*-causal. *E* is necessary because *D* only occurs if *E* caused it.

Presenting the monocausal model in the above form draws attention to its role as a model for *defining* particular disease types/taxa such as anthrax or typhoid fever. Applied to the example of anthrax, an instance of infection (*a*) is a case of anthrax (*D*) *only if* the germ *B. anthracis* (*E*) caused the infection. As a necessary cause, *E* is a cause of every instance of *D*. Although this necessity arises because *D* is defined in terms of *E*, we cannot define *D* in terms of just any factor. The factor we choose must be a cause of *D*.[4] Whether or not a particular factor is a cause of *D* is an empirical matter, to be settled through empirical research rather than by stipulation.

Although the condition that *E* caused *a* is necessary for *a* to be a case of *D*, it is not sufficient. *B. anthracis* can cause many things – an immune response in those who have been vaccinated against the bacterium, the death of livestock, public hysteria. These occurrences are not thereby cases of anthrax. As a model for defining diseases, the monocausal model as I have presented it is incomplete, yet the constraint that it places on disease classification – the requirement of defining disease types according to a specific cause – is mighty nonetheless.

Broadbent's reconstruction of the monocausal model places a second requirement on the defining cause, a circumstantial sufficiency requirement: "given certain circumstances, which are not sufficient to cause D, every occurrence of C causes a case of D" (2013, pp. 150).[5,6] As

[2] Paul Thagard (1999) calls the change from one organizing taxonomic principle to another principle "tree switching".

[3] Greene et al. (2012) note that this "therapeutic revolution" was more complicated than this simple story might suggest, and that new therapies often took decades to arrive.

[4] One who holds that diseases form natural kinds (e.g. Lange, 2007) might want to add that for our disease classifications to be natural we must choose the *right* causes. Whether or not diseases form natural kinds – and if so, how we go about defining diseases accordingly – is a further issue for another paper.

[5] J.L. Mackie (1965) calls circumstantially sufficient causal conditions "*minimally* sufficient conditions" to emphasize that they contain no idle parts; were any cause missing, then the remaining causes would no longer be sufficient.

[6] Broadbent (2009) offers a different formulation of the second requirement: "given certain circumstances, a *C*-event is not a cause of any ¬*D* event (i.e. other diseases or good