

PERSPECTIVE

Etiology in a taxonomy of illnesses

Johann Steurer¹, Lucas M. Bachmann^{1,2} & Olli S. Miettinen^{3,4}

¹Horten Center for Patient-oriented Research and Knowledge Transfer, Department of Internal Medicine, University of Zurich, Zurich, Switzerland; ²Division of Epidemiology, Biostatistics, Institute for Social and Preventive Medicine, University of Bern, Bern, Switzerland; ³Department of Epidemiology, Biostatistics and Occupational Health, and Department of Medicine, Faculty of Medicine, McGill University, Montreal, Canada; ⁴Department of Medicine, Weill Medical College, Cornell University, New York, USA

Accepted in revised form 9 December 2005

Abstract. According to what Robert Koch termed the etiological standpoint, illnesses are best understood and controlled by focusing on their causes, including in their definitions and, thus, in the construction of their taxonomies. In some ways flawed, this standpoint has been misunderstood and misapplied. A taxonomy based solely on etiology was an unrealistic dream in the context of ‘the bacteriological revolution’, and it also is unrealistic in the present context of ‘the genetic revolution.’ We argue that the illnesses in a taxonomy of them are in some cases best defined directly in terms of their respective somatic anomalies,

in some others indirectly by the unique and universal etiology of that anomaly (left unspecified) in a ‘deeper’ somatic anomaly, and in yet others as a combination of these; and when the somatic anomaly for direct definition remains unknown, it is to be defined indirectly by the clinical syndrome that is its patient-relevant manifestation, possibly in conjunction with a somatic cause. We note, also, that these taxonomic issues have no material bearing on epidemiologists’ etiologic research for the knowledge base of community-level preventive medicine.

Key words: Definition, Etiology, Genetic illness, Infectious illness, Taxonomy

Introduction

K.C. Carter [1], a philosopher who ‘wanted to find out what doctors were trained to believe,’ opened his recent treatise on ‘causal concepts of disease’ thus:

Of the numerous changes that have occurred in medical thinking over the last two centuries, none have been more consequential than the adoption of what Robert Koch called the etiological standpoint [ref.].

Carter explained this standpoint to be ‘the belief that diseases are best controlled and understood by means of causes’ and, in particular, ones that are natural, necessary, and universal (same in each instance of the disease). ‘This way of conceiving disease has dominated medical thought for the last century,’ he noted.

H.R. Wulff [2], a physician famous for his ‘Rational Diagnosis and Treatment’ book, for example, characterized the 19th century redefinition of diseases on the basis of ‘the identity of the microorganism’ as ‘an important phase in the disease classification as it strengthened the false idea of monocausal determination of diseases’. He characterized the microorganism – specific, unique – as representing only one of the causes of an infectious disease, even though a necessary one.

A.S. Evans [3], a physician eminent in microbiology and epidemiology, took issue with the idea of ‘the’ microorganism in what already was understood to be the multifactorial causation of infectious diseases: ‘the idea that an infectious malady can be caused only by the action of a single agent is incorrect’; ‘it became apparent about 1960 that a number of common illnesses or syndromes existed in which several agents could produce the same clinical picture’. Highly illustrative of this is the recent history of the virology of hepatitis [4].

While the microbiological thinking about the etiology of infectious diseases has been evolving through revisions, two ideas in this have not: microbial agents continue to be viewed as causes of infectious diseases and such causation as definitional to infectious diseases. Koch, having ‘proven’ the tubercle bacillus (*mycobacterium tuberculosis*, var. *hominis*) to be ‘the’ cause of tuberculosis (of the pathology of it), ‘defined tuberculosis as infestation by the tubercle bacillus’ [1], and that definition scarcely would be challenged at present even if it be understood that ‘the tubercle bacillus is universally necessary for tuberculosis only because “tuberculosis” is defined as infection by this bacillus’ [1]. This dual role for the microorganism(s) in the concept of any infectious disease persists,

arguably at least, because of continual failure to appreciate that ‘an effect – the disease – should not be confused with its own cause’ [5].

The developments in the definitions of infectious diseases in the latter half of the 19th century soon became paradigmatic in respect to other illnesses also. Thus, the point was made in 1884 that ‘symptomatology and pathological anatomy could not significantly advance the comprehension of any disease’; that the comprehension ‘required following the model of bacterial theory’ [1]. Sigmund Freud, among others, paid heed; and so it was, for example, that before the century had drawn to its close, ‘the sexual etiology of hysteria had become definitional’ [1].

Through much of the 20th century ‘the model of the bacterial theory’ got to be followed quite literally by the many eminent ‘alienists’ – psychiatrists – who took various mental illnesses to be caused by ‘focal infection’ and on this basis committed, or condoned, unconscionable surgical enormities in mental hospitals [6].

Most recently this hubris, in terms of which ‘the practical payoffs of the bacteriological revolution seemed limitless’ [6], has been matched, if not superseded, by another:

The human genome sequence will dramatically alter how we define, prevent, and treat disease. We need to define the term disease so that it incorporates our expanding genetic knowledge, ... With recent dramatic advances in genetics and genomics, essentialists could argue that the essential lesion defining the disease state is a genetic abnormality [7].

Doctors in both clinical and community medicine, and clinical and epidemiological researchers also, indeed need to adopt a concept of disease (*morbus*) – and of defect (*vitium*) and injury (*trauma*) besides [8] – consistent with most recent understandings, including as for the proper role of etiology in this. But the larger need is to achieve a suitable taxonomy of the component entities of ill-health, again including in respect to the role of etiology in this – if any.

Given that epidemiological research typically is concerned with the etiology of entities of ill-health – of illnesses, that is [8] – it is particularly important for research epidemiologists to be clear on whether their etiologic research addresses something that already is, or in the future might become, definitional to the illness at issue in any given study; or whether, instead, at issue in each study is something dependent on a rational taxonomy of illnesses but having no bearing on such definitions. Toward properly answering these questions, the point of departure naturally is critical understanding of Koch’s ideas on etiologic definition of illnesses.

‘The etiological standpoint’

That ‘the etiological standpoint’ of Koch ‘has dominated medical thought for the last century’ is under-

standable, not only on the grounds of Koch’s great authority – very well deserved – but also because of its obvious correctness in the context in which he advocated this standpoint. For, no one can take issue with, for example, the idea that phthisis (consumption, wasting), and scrofula (brood sow, king’s evil) also, is ‘best controlled and understood by means of its cause’ in the proliferation of (a particular type of) mycobacteria in the lung parenchyma and lymph nodes, respectively.

To take one other, analogous and quite recent example, what was called hepatitis was a clinical concept of an aggregate of symptoms and signs attributed to the liver – though not to inflammation of it – in the Hippocratic Corpus already [4]. But the clinical syndrome obviously is best understood and controlled when appreciating its now-recognized causation in, specifically, inflammation of the liver, and the latter by means of its most recently understood causes in viral infections.

This standpoint has, however, been misunderstood and misused. Early on, Freud ‘repeatedly exploited the tactic of redefining diseases in terms of causes’ [1]. But the relation of clinical hysteria to sexual traumas in early childhood, to take this example, is not analogous with the relation of, say, phthisis to tuberculosis (tissue nodosity) nor with the relation of the latter to the mycobacterium’s proliferation in the affected tissue. Those purported causes of hysteria are matters of the past and thereby not, even in principle, subject to removal as a therapeutic intervention. By contrast, even though the notion of ‘focal infection’ as a ubiquitous cause of mental illness was definitely wrong and utterly tragic in its implications for the practice of ‘scientific’ medicine, it could in principle have been a salient example of the usefulness of the etiological standpoint.

Misunderstanding of the etiological standpoint has resulted, we suggest, from a common failure to appreciate an implicit but important subtlety in it: Etiology in Koch’s standpoint has two meanings. It refers, for one, to the causation of the patient’s manifest ill-health – sickness [9] – in its underlying somatic anomaly. In the main, however, his etiological standpoint concerns a sickness-producing *somatic anomaly’s etiology in its underlying somatic anomaly* and the latter in the particular meaning of the ‘deeper,’ primary anomaly being still present and continually *sustaining* the more ‘superficial,’ secondary one or its progression. A prime example is the etiology of tuberculosis in mycobacteriosis.

Failure to exemplify this special, limited meaning of etiology makes Freud’s etiologic definitions of neuroses and psychoses not to conform to Koch’s etiological standpoint and, thereby, not to possess the practical and conceptual usefulness of this. Venous thrombosis is etiologic to pulmonary embolism, and peptic ulcer to peritonitis, but neither one of these relations – between a somatic anomaly and its

complication – exemplifies the etiological standpoint of Koch, to say nothing about the relation – non-causal – of sequelae to their precursor anomalies in diseases or injuries [10]. But that standpoint is exemplified by the etiology of pulmonary edema in failure of the left ventricle of the heart, and that of the latter in aortic stenosis, though without the necessity (and sufficiency) of the cause.

When the concern is to define ‘diseases’ genetically, disease is said to be construed as ‘a state that places individuals at increased risk of adverse consequences’ (emphases deleted) [7]; but again, this is not what Koch’s etiological standpoint was about. That standpoint – and more – would be left behind when defining diseases in terms of, say, the BRCA1 and BRCA2 genes or trisomy 21, as removal of these causes would not ameliorate what patients with these anomalies actually are suffering from, if anything. It is, thus, only outside the etiological standpoint – in Koch’s original meaning – and with disregard for other established patterns of medical thought that ‘essentialists could argue that the essential lesion defining the disease state is a genetic anomaly.’

An agent as the/a cause of an illness

Koch’s ‘proof’ of the tubercle bacillus as being ‘the cause of tuberculosis’ was, in generic terms, a purported demonstration of an agent as being pathogenic – causal to an illness – in the limited meaning of etiology inherent in his etiological standpoint (see above). He indeed wrote about ‘the parasite as the cause of the disease’ [1]; and Wulff and Evans, for example, wrote about microorganisms as causes of infectious diseases (cf. above).

In respect to infectious diseases in the meaning of the likes of tuberculosis and hepatitis – as distinct from infectious diseases such as cervical cancer – there can be no question about the involvement of the/a microorganism in the disease process itself. But as for the causality of this involvement, the question here is whether the microorganism *per se* is a cause of the disease, as commonly is asserted.

In Koch’s ‘proof’ that the mycobacterium is causal to tuberculosis, the final step was – in his own words – to ‘show that animals inoculated with pure culture [of the bacillus] contract the original disease [i.e., the tuberculosis in the source of the cultivated bacillus]’ [1]. But to us this does not show that the bacillus *per se* is causal to tuberculosis. To us this shows, instead, that inoculation with the bacillus is a cause of the disease – or more generally, that *opportunity for the agent’s entry* to the host is a cause of the disease in the host.

Even when saying, in the modern framework of multifactorial etiologic thought about infectious diseases, that a microorganism is among the causes [2, 3], one is oblivious to the now generally understood truth that the cause of an infectious disease always is

a person’s *exposure* to the/a relevant microorganism (rather than the microorganism *per se*) in conjunction with the person’s *susceptibility* thereby to contract the disease – with each of these two major generic causes, each necessary but insufficient, encompassing multiple component causes.

The mycobacterium *per se* thus is not even a component cause of tuberculosis, just as thalidomide in unused pills cannot be etiologic to phocomelia or benzpyrene in uninhaled cigarette smoke to lung cancer. But various component causes in these examples have to do with those agents – exposure or susceptibility to them. It thus is not without reason that exposure is so central a concept in epidemiology (even if universally applicable only to environmental factors, never to constitutional ones, and only selectively to behavioral ones.)

An agent’s activity as the essence of an illness

An illness, like anything else, is properly defined by specifying its essence – that which is present in all instances of it and unique to it.

When Koch redefined tuberculosis as ‘infestation by the tubercle bacillus,’ he implied that all cases of tuberculosis are characterized by infestation with the tubercle bacillus, and not only uniquely but also causally so; and causal in this context he indeed understood to be not the bacillus *per se* (cf. above) but infestation with it – meaning the parasite’s entry into the body followed by its proliferation within the body, this as a necessary and sufficient cause.

These implications of the redefinition can still be viewed as tenable (except, perhaps, the implied non-existence of healthy carriers of the bacillus, inherent in the uniqueness implication). From this it does not follow, however, that the redefinition remains – or ever was – justifiable.

The truth of the matter, pure and simple, is this: Whereas phthisis had been a recognized entity of sickness, and pathologists had recognized (a particular type of) tissue tuberculosis (nodosity) to be causal to it, Koch’s bacteriological work led to the discovery of mycobacteriosis as ‘the’ cause of tuberculosis. And momentous though this bacteriological discovery was, it should not have been taken to have bearing on the definitional essence of tuberculosis as a pathological entity, much less did it call for, or justify, redefinition of the corresponding clinical concept of sickness – phthisis – in bacteriological terms. For, ‘an effect – disease [the anomaly on a more superficial level] – should not be confused with its own cause [in an anomaly on a deeper level].’ Even though etiologically related, the entities on the three levels are mutually distinct, requiring separate definitions, and separate terms also. If pulmonary edema were known always to be caused by congestive heart failure, this would not justify redefining the former as being the latter.

Koch's bacterial redefinition of tuberculosis and other infectious diseases thus was unjustifiable as such, and it therefore unjustifiably became a paradigm for etiologic definitions of illnesses in general – and, misguidedly, even in respect to etiologies that do not fit into the etiological standpoint which concerns, to say it again, synchronous anomalies one of which sustains the other or its progression.

Etiology as the essence of an illness

The concept of etiology has undergone an evolution in the context of infectious diseases, and quite generally also. Tuberculosis as a somatic anomaly is no longer thought of as being 'etiologic' to the sickness entity phthisis, nor is coronary stenosis viewed as 'etiologic' to angina pectoris. Sickness in an illness is now thought of as the patient-relevant manifestation of the illness, with the latter construed as a somatic anomaly with potential for such manifestation, and etiologic attention now focuses on the genesis of the illness-defining somatic anomaly. Descriptively the anomaly defining a disease or a defect is the result of its pathogenesis, and etiology now is taken to be the etiogenesis of that anomaly – the aggregate of *causal influences initiating and/or advancing pathogenesis* [11].

When etiology E, in this meaning, is taken/defined to be the essence of illness I, definitional to I is its E-etiology alone and, thus, not at all the particular nature of the resulting somatic anomaly. For example, when an illness is defined by infestation with mycobacteria in – mycobacteriosis of – the lungs, and this not as the illness but as the etiology of it, the essence/definition of the illness itself is: the (further) somatic anomaly (or anomalies) caused by (the primary anomaly of) pulmonary mycobacteriosis. That this illness is generally characterized by tuberculosis/nodosity of the lung parenchyma is not at all definitional to this etiologically defined illness. Different from this example, syphilis – 'the great imitator' – really needs to be defined by the infestation (spirochetosis) specific to it, this as the etiology of the multitudinous anomalies resulting from this cause.

Etiological definition is natural for a multitude of infectious diseases so long as the etiology is taken to be the infestation rather than the etiology of this (in exposure and susceptibility to the microorganism(s) involved), and for many other illnesses besides. Generic examples of those other illnesses are conditions of poisonings and vitamin deficiencies, both of which involve an anomalous somatic state as the etiology, akin to the infestation in an infectious disease. Various metabolic anomalies can be defined by their endocrinological etiology instead of the nature of the metabolic anomaly itself; and the etiological definition of the anomaly can also be genetic, as an alternative to inability to metabolize phenylalanine, for example.

Etiology in the essence of an illness

Some illnesses, while defined as a somatic anomaly of a given type as a first approximation, gain a more specific definition of the nature of this by reference to a particular one of its various etiologies in a somatic anomaly of a different kind. Thus, while pneumonia is any inflammation of the lungs, it is useful to bring etiology into the concept and thus to have the more specific concepts of anthrax pneumonia and aspiration pneumonia, for example. And while myocardial infarction is the formation of ischemic necrosis in the myocardium, it is useful to distinguish subtypes for this according to etiology in coronary thrombosis and other somatic causes.

On the other hand, bringing trisomy (or translocation) of chromosome 21 into the definition of Down's syndrome provides no further specificity to this syndrome. There thus are these two options: definition of the syndrome by its particulars and without any reference to its etiology, or definition of it as the syndrome caused by trisomy (or translocation) of chromosome 21 and without reference to the syndrome's particulars. For when it is said that a concept's definition specifies that which is present in all instances of it and unique to it, implicitly meant is that a proper definition involves nothing but this.

Etiology in a taxonomy of illnesses

Once illnesses have been defined, a taxonomy of illnesses is the classification of illnesses resulting from this – so long as the defined illnesses/anomalies constitute a set of mutually exclusive and all-inclusive categories/taxa.

Defects and injuries naturally are generally defined directly, by the nature of the anomaly, ventricular septal defect and hip fracture, for example; but some may be more advantageously defined indirectly, by their etiology, trisomy 21 syndrome and cerebral concussion, for example. As for diseases, some need to be defined directly, malignant and atherosclerotic anomalies, for example; but etiologic definition is natural for others, infectious diseases and vitamin deficiencies, for example.

We presume it to be abundantly clear from the foregoing that a taxonomy based solely on etiology was an unrealistic dream in the context of 'the bacteriological revolution,' and that it also is unrealistic in the present context of 'the genetic revolution.'

Etiology in epidemiological objects of study

Etiology in the meaning of Koch's etiological standpoint has not been seen as a natural concern in epidemiologists' etiologic research; and thus, etiology in that meaning has in all essence been the province of

laboratory and clinical researchers only. As for genetic etiology, epidemiologists used to distinguish themselves from geneticists by a focus on nongenetic etiology with the idea that relevance to community-level preventive medicine requires this. These distinctions have, however, become rather blurred upon the advent of, notably, ‘molecular epidemiology’ and ‘genetic epidemiology,’ and ‘clinical epidemiology’ besides.

To the extent that epidemiologic researchers continue to see themselves in the service of public health in the particular meaning of developing the knowledge base for community-level preventive medicine through etiologic research, it does not matter to them whether illness anomalies are defined indirectly by clinical syndromes such as idiopathic low-back pain, directly as somatic anomalies, or indirectly by somatic etiology of these, genetically perhaps. The pragmatic preventive-medicine standpoint does, however, continue to call for focus on potential etiologic factors subject to community-level intervention, factors environmental and behavioral in the main, but constitutional also, as in the need for immunization. Even genetic anomalies can have prevention-relevant environmental or behavioral etiology, and so can acquired constitutional susceptibility to illness, that of AIDS, for example.

Etiologic research of this type can have the side-effect of misguidedly obfuscating the taxonomy of illnesses, leading to concepts such as ‘occupational lung cancer.’ Properly, lung cancer is construed as that which this term implies, as cancer originating in the lungs. It has no ‘occupational’ subtype(s); only its etiology does.

Specialties of epidemiological practice of community-level preventive medicine each have their particular sets of illnesses of concern, different among occupational epidemiology, nutritional epidemiology, infectious-disease epidemiology, etc. The set of these sets, however, does not constitute an etiologic taxonomy of illnesses on an aggregative level, as the categories in this are neither mutually exclusive nor all-inclusive.

Finally, a point about the operational definition of the illness outcome in epidemiological research on etiology. The present routine is to include in the case series all cases of the illness identified – by rule-in diagnosis in routine practice – in the study base. But a serious validity problem can arise from the etiologic history being a criterion in many of the diagnoses, as

when studying the etiology of pulmonary embolism in oral-contraceptive use. This problem can be obviated, largely or even totally perhaps, by restricting the case series to instances of the illness that are severe and typical [12]. For these cases come to medical attention irrespective of the history at issue, and they are prone to be diagnosed without that history having a role in it.

References

1. Carter KC. The Rise of Causal Concepts of Disease. Case Histories. Hants (UK): Ashgate Publishing Company, 2003, pp. vii, 1, 38, 52, 135, 147, 152, 158.
2. Wulff HR. The causal basis of the current disease classification. In: Nordenfeld L and Lindahl BIB (eds): Health, Disease, and Causal Explanations in Medicine. Dordrecht: D Reidel Publishing Company, 1984.
3. Evans AS. Causation and disease: The Henle–Koch postulates revisited. *Yale J Biol Med* 1976; 49: 175–195.
4. Duffin J. Lovers and Livers. Disease Concepts in History. Toronto: University of Toronto Press, 2005, pp. 85, 87, 97.
5. Scadding JG. Essentialism and nominalism in medicine: Logic of diagnosis in disease terminology. *Lancet* 1996; 348: 594–596.
6. Scull A. Madhouse. A Tragic Tale of Megalomania and Modern Medicine. New Haven (CT): Yale University Press, 2005, 28 ppff.
7. Temple LKF, McLeod RS, Gallinger S, Wright JG. Defining disease in the genomics era. *Science* 2001; 293: 807–808.
8. Miettinen OS, Flegel KM. Elementary concepts of medicine: III. Illness somatic anomaly with... *J Eval Clin Pract* 2003; 9: 315–317.
9. Miettinen OS, Flegel KM. Elementary concepts of medicine: IV. Sickness from illness and in health. *J Eval Clin Pract* 2003; 9: 319–320.
10. Miettinen OS, Flegel KM. Elementary concepts of medicine: VII. Course of illness: manifestations, complications, outcome. *J Eval Clin Pract* 2003; 9: 329–331.
11. Miettinen OS, Flegel KM. Elementary concepts of medicine: VI. Genesis of illness: pathogenesis, aetio-genesis. *J Eval Clin Pract* 2003; 9: 325–327.
12. Miettinen OS. Etiologic research: needed revisions of concepts and principles. *Scand J Work Environ Health* 1999; 25: 484–490.

Address for correspondence: Johann Steurer, Horten Zentrum für praxisorientierte Forschung und Wissenstransfer, Universitätsspital Zürich, CH-8091, Zürich, Schweiz
Phone: +41-1-255-31-98; Fax: +41-1-255-397-20;
E-mail: johann.steurer@usz.ch