EXPLANATION IN MEDICINE

Maël Lemoine

Introduction
The scientific part of medicine seeks explanations. Some think that nowadays evidence-based medicine embodies scientific medicine. Yet evidence-based medicine does not focus on explanation. Moreover, as Thompson puts it, “the results of randomized controlled trials cannot underwrite explanations, notwithstanding the current emphasis on them in clinical medicine” (Thompson 2011).

This chapter aims to determine the nature, the function, and the importance of explanation in medical science, from basic biological studies of diseases to epidemiology. Medicine uses diverse explanatory strategies. After a review of the basics of scientific explanation according to general philosophy of science, as applied to medical cases, the chapter will focus on the specifics of medical explanations, as more recent developments in the philosophy of medicine have established and discussed them. A satisfactory account of explanations in medicine should:

1. encompass all types of medical explanation (molecular, epidemiological, psychiatric, pathophysiological, social. . .) or explicitly define a more limited target;
2. state what is specific to medical explanation, if anything, as opposed to non-medical explanation;
3. account for the function of explanation as opposed to other activities such as providing evidence.

We will successively examine several accounts of explanation in medicine, none of which is likely to be entirely satisfactory.

1. Basics of Scientific Explanation

1.1. The Deductive-Nomological (D-N) Model of Explanation

In 1844–1848, the Austrian physician Ignaz Semmelweis was looking for an explanation of the epidemic of puerperal fever that killed significantly more women in one of the two Maternity Divisions of the Vienna General Hospital than in the other. The ward with the higher mortality was staffed by doctors and medical students, and the other one with midwives. After ruling out many possible explanations that neither matched the facts nor prompted actions that succeeded in controlling the infection, Semmelweis hypothesized that students and doctors, who often arrived in the ward directly from conducting autopsies, transported “cadaveric...
matter” from the dissection tables into the women’s bodies. Midwives, on the other hand, did not perform autopsies. This hypothesis suggested thorough hand washing prior to delivery, which led to a sharp decline in mortality in the ward staffed by doctors and students (Hempel 1966; Gillies 2005).

We can see this short case study as, among other things, an explanation. Semmelweis explained the higher rate of infection among one set of women by the transferring of cadaveric material to them by one set of caregivers but not another. Hempel took this as an example of how scientists seek explanations. His conclusion is that they look for a hypothesis that may “explain” the observed facts and deduce from it further observations that can be checked in turn. This so-called hypotheticodeductive view has been much discussed. For instance, Lipton claims that Semmelweis’s choice of the cadaveric matter hypothesis over other hypotheses was not logical but aesthetic (“loveliness”), as none of Semmelweis’s various hypotheses were, strictly speaking, incompatible with the facts. Thus, Lipton claims, Semmelweis’s reasoning is an example of inference to the best explanation (i.e., of the conclusion that the best explanation is the explanation; Lipton 2003), with best, in this case, being loveliest according to Lipton. Adopting a less demanding view of evidence than Lipton, Bird argues that Semmelweis’s successive contrastive hypotheses were indeed refuted, so that just one remained in the end, which makes this a case of inference to the only explanation (Bird 2007), or eliminative abduction (Bird 2010), a specific form of the inference to the best explanation.

Should an explanation be the loveliest hypothesis or the last one remaining? This may be an important characteristic of what should count as an explanation, but it hardly counts as a definition of explanation. What all three authors above seem to agree on is the fact that hypothesis confirmation implies deduction, because hypotheses are explanations, and explanations imply deduction. In fact, the main classic description of scientific explanation is called the Deductive-Nomological (D-N) view. First put forward precisely by the same Hempel, along with Oppenheim (1948), the D-N view claims that a hypothesis explains an observation by logically implying it. In particular, scientific hypotheses explain by being general statements of the form “All X are Y.” If the phenomenon of interest is indeed X, then it follows that it is Y as well. Moreover, if the general statement is verified, it is considered a law of nature. In that case, the phenomenon of interest is considered explained by this law of nature: we know why the phenomenon of interest is Y; i.e., because it is X (observation) and “all X are Y” (law of nature).

In the case of childbed fever, for the hypothesis to be a possible explanation, it is necessary (and sufficient) that (1) the general statement or hypothesis “childbed fever is blood poisoning produced by cadaveric matter” be true, and that (2) observations (such as “its prevalence is reduced by antiseptic measures”) logically follow from the general statement. However, the observations do not prove the general statement, as Hempel emphasizes.

Since its introduction, the D-N account has been subject to many objections. A few of these are particularly relevant to its (potential) use in philosophy of medicine. One such objection is that the D-N view should, but does not, require that the explanans (what explains) be a cause of the explanandum (what is explained). Thus, for example, from mercury rising in a thermometer, fever can be deduced, but the thermometer does not explain the fever. On the other hand, fever explains the rising of mercury because it also causes it. Causality, not only correlation, should be a further requirement of proper explanations, and a version of the D-N view thus amended would account for most explanations in biology and medicine, according to Schaffner (1993: 267).

Yet this does not appear to be an adequate solution to the problem of applying the D-N view to medicine, as it is often the case, in medicine, that we observe the correlation between facts, but do not understand how one causes the other. Many signs of diseases have been known long before we understood what causes them, and some are yet to be properly explained. Schaffner
suggests that such non-causal “explanations” are considered “provisional surrogates for deeper causal explanations” (Schaffner 1993: 274). This raises as many questions as it solves: are they necessary, to whom and why? How long can they remain provisional? In what sense is a surrogate explanation, an explanation?

A second reason to doubt the relevance of the D-N model in medicine is that most medical explanations do not refer to laws of nature. Examining short-term potentiation in neurobiology, inheritance, and AIDS, Schaffner concludes that the general premise in the *explanans* is not a law of nature, but “rather a set of causal sentences of varying degrees of generality” (Schaffner 1993: 286). Some of these sentences are very general and resemble laws of nature. One can think of often implicit principles such as “a phenotypic trait of a heterozygotic individual is determined by the dominant allele.” But others are very specific. For instance, “Gaucher disease is prevalent among Ashkenazi Jews” cannot be thought of as a law of nature. Schaffner attempts to solve this problem for applying the D-N view to medicine by proposing a second amendment to the D-N view: theories in biology and medicine are what he calls “middle range” theories—i.e., not encompassing the whole realm of the living but limited to a species (as *Aplysia californicum*), a particular strain in a species (a mutant *E. Coli*), or even a subpopulation (Ashkenazi Jews). Yet it is questionable whether this should still be considered a nomological view of explanations. Often, explanatory general claims come in numerous versions, depending on the species, or even a particular strain in a species. And many generalizations are required in most medical explanations, not just one law.

1.2. Statistical Models of Explanation

A modified version of the D-N view may fit some explanations in medicine, but it certainly does not account for all of them. Hempel provided other models for scientific explanation in general (1965), some of which are of obvious importance to medical explanations in particular. First, deductive-statistical (D-S) explanations rely on a statistical general premise. An example is:

- 25% of children with heterozygous parents have a disease associated with a recessive allele.
- Sickle-cell anemia is associated with a recessive allele.
- Population X has heterozygous parents carrying the recessive allele of sickle-cell anemia.
- 25% of population X has sickle-cell anemia.

Note that the 25% figure is not an observation but logically follows from the assumption that humans have two alleles of the same gene and that meiosis (i.e., the selection of alleles from parents) is random. Note also that the *explanans* is a general statistical proposition and that the *explanandum* is a less general statistical proposition, derived from the former. Such explanations obviously define a different sort of medical explanations, not all of them.

A further sort of explanation according to Hempel, Inductive-Statistical (I-S) explanation, derives a particular fact, the *explanandum*, from a general statistical fact, the *explanans*. The derivation is not deductive (i.e., logically included in the general claim) but inductive (i.e., consists in an extension of what is observed to a more general claim). Hempel’s (1965) famous example is:

The statistical probability of recovery of patients with streptococcus treated with penicillin is close to 1.
Jones was infected with streptococcus.
Jones is being treated with penicillin which
(made very likely that) he would recover from his infection.

As noted by Schaffner, many explanations in medicine are similar to this one (Schaffner 1993: 268), not only those involving individuals, but populations as well (e.g., replace “Jones” with “n patients in Ohio in 1959” in the previous derivation). This argument is inductive rather than deductive because no particular case follows necessarily from an almost general statement—Jones may have been infected with a streptococcal strain resistant to penicillin.

As Hempel points out, I-S explanations are ambiguous in an essential sense. If it is highly likely, but not necessary, that people like Jones, who are infected with streptococcus and treated with penicillin, recover, then there is also a subset of these same people for whom it is highly likely that they will not recover, including people infected with penicillin-resistant streptococcus and octogenarians with weak hearts. There are therefore two equally sound I-S explanations, with true premises, leading to opposite conclusions about people infected with streptococcus and treated with penicillin. We can thus explain both Jones’ recovery and his death equally well. Additionally, I-S explanations are “epistemically relative” (Hempel 1965): they are not true or untrue, but relevant or irrelevant, given the state of knowledge. Indeed, we know that we should not indifferently apply the same probability of recovery to all populations.

But what does “relevance” mean? Since Hempel and Carnap, many have defined it as the choice of the proper reference class to which an individual belong. In this context, a reference class is a group of individuals with causally associated features: for instance, having a weak heart (or being old) and being more likely to die from streptococcus infection. If an individual belongs to several reference classes for which varying I-S explanations can be given, the question of which reference class should be used in the explanation is known as “the problem of relevance.” Consider the following example:

Jones had a cold.
Jones took vitamin C.
Almost all people taking vitamin C recover from a cold within one week.
(makes it very likely that) Jones would recover from his cold within one week.

Salmon objects that this inference conforms to the I-S model of explanation but is not an explanation. The reason it is not is that most people recover from a cold within one week, with or without vitamin C. “People treated with vitamin C” is not a relevant reference class in this situation because this apparent explanans does not make any difference to the outcome. Salmon thus proposes the Statistical Relevance (S-R) model of explanation (Salmon 2006). The basic idea is that the explanans should contain all (and only) statistical information relevant to the explanandum. Jones’ recovery is explained by his being treated by penicillin, because the probability of his recovery given that he has strep infection and has been treated with penicillin is higher than the probability of his recovery given that he has strep infection. On the other hand, vitamin C is not relevant because it does not raise the probability of recovery from common cold. It is also relevant that Jones is host to a penicillin-sensitive bacterium, but irrelevant that Jones has been contaminated by his mother rather than his sister.

Yet Salmon’s proposal still leaves unsolved a problem as to the choice of alternative relevant reference classes. Of course, the probability of recovery given both strep infection and penicillin treatment can be more accurately predicted by referring to narrower subclasses (e.g., including whether the strain of strep is resistant to penicillin). But say Jones, who is 35 years old and has a weak heart, is infected with a non-resistant strain, is given penicillin,
and recovers from strep infection. What is the relevant reference class for the explanation to hold? As a first attempt to resolve this problem, Carnap introduced the “requirement of total evidence,” which demands that all the knowledge we have about a specific situation be used in the assessment of the probability of an outcome (Carnap 1962). Does this mean that probabilities add up, or that the right probability to use is attached to the intersection of all classes? Hempel proposed a different criterion—that we should use the narrowest reference class the individual belongs to—that is, the narrowest class for which we have information relevant in the situation (Hempel 1965). However, these requirements may conflict. The reason is that no matter how narrow the reference class for which evidence of probability exists, there still is further information relevant to the probability of the outcome for a particular patient—for instance, that Jones may not be observant because he is afraid of antibiotics, changes the probability of recovery. Even if the probability that he does not take the treatment given that he is afraid was known, the narrowest reference class requirement states that we use the conditional probability given everything else as well, which is different (Fuller and Flores 2015). In addition to the questions just raised, other problems with identifying reference classes in medical explanations have been raised (Clarke 2011).

In the end, just as Schaffner does in the case of the D-N model, Salmon argues that causation should be involved between the explanans and the explanandum in the S-R model. In the explanation of GI Joe’s death by leukemia, Salmon says, many facts are relevant, such as the atomic blast he was exposed to, his position relative to it, his shelter, etc. All of them are causally relevant to his death, and that is why they are explanatory. Because of the apparent universality of the requirement of causality for explanations, and because of the difficulties of the D-N, D-S, I-S, and S-R models of explanations, many now prefer to start all over again from the simpler idea that explanations are causal, and more specifically in the life sciences, that they are mechanistic.

1.3. Mechanistic Explanations in Medicine

In accounting for scientific explanations in general, philosophers of science have increasingly moved their attention away from laws and toward models of phenomena. In medicine, mechanistic models in particular have been the focus of much attention when discussing explanations. Mechanistic models are the kind of schemas you find in biology or physiology textbooks: they represent entities (e.g., cell receptors and neurotransmitters), the interaction of which account for the general outcome, that is, regular behaviors of the system (Machamer, Darden, and Craver 2000). In medicine, mechanisms are deeply connected to physiology. Physiology has been a basic medical science since Hippocratic medicine, and one could say that it is the science of the mechanisms of organisms. Another way to put it is that physiology is the science of biological functions. This means that our understanding of mechanism and function within medicine are closely connected. Therefore, an important debate in philosophy of biology about the definition of “function” is important to understanding mechanistic explanations in medicine. One approach to functions in biology is to understand them as a causal role in a system (see Ariew, Cummins, and Perlman 2002). A causal role is the contribution a part of the system is expected to provide for a general effect to obtain: for instance, the heart pumps blood for circulation, so pumping blood is the heart’s causal role in the body. In a different understanding, a function is an effect that is currently present in a population because it has been selected during evolution (Millikan 1989): at a certain time, a certain contractile structure improved circulation by pumping fluids, and this ancestor of hearts was selected because it improved fitness. This general debate leads to two complementary understandings of what a mechanistic explanation in medicine should look like: “functional explanations”
EXPLANATION IN MEDICINE

and “evolutionary explanations.” Whereas the former focuses on what parts do and what their functions are—explaining *based on* functional effects—the latter focuses on why the parts are here in the first place (i.e., because their effects were selected)—explaining *why* effects such and such, rather than anything else, *are* functional effects. There may, for instance, be vomiting in the course of a bacterial infection. One can either explain the vomiting as a dysfunctional mechanism, some part of the system not fulfilling its function, or explain that vomiting may have been selected during evolution as a defense mechanism (Nesse, Williams, and Brown 1996). Let us examine each of these approaches.

1.3.1. Functional Explanations in Medicine

Explaining an observed effect by the causal contributions of biological functions is the essential task of physiology, thereby playing a major role in medical explanations. Organisms can be thought of as systems within which interacting parts tend to maintain or attain certain states, such as keeping the level of blood glucose fairly constant or minimizing pain (Sommerhoff 1950; Nagel 1961: 411–9). These phenomena can be explained more specifically through “functional analysis.”

The most famous description of functional analysis is Cummins’ (1975, 1985). According to Cummins, functional analysis is a strategy used to explain how a system can exert a capacity. The functional analysis consists in decomposing this capacity into an organized series of sub-capacities: for instance, blood circulates (capacity), because heart pumps (sub-capacity 1), artery walls keep the system close (sub-capacity 2), etc. Each of these capacities may in turn be explained in the same way, until it bottoms out of biological science at the molecular level. *Functional analysis* is the analysis of one capacity of a system into a more or less complex concatenation of functions.

There is widespread acknowledgment that functional analysis accounts for a very large set of medical explanations, roughly, physiological explanations, such as the example of blood circulation above. However, this view cannot account for a very important, related sort of medical explanations. Take the following example:

The function of the adrenal glands is to produce cortisol.

Jones’ adrenal glands do not produce cortisol.

Jones’ adrenal glands are dysfunctional.

Explanations of this sort, which are important in medicine, say why an observed fact should be judged normal or abnormal. The mechanistic approach, however, which does not make any difference between mechanisms of diseases and mechanisms of healthy processes (Craver 2001: 67), does not account for such explanations: the outcome of the mechanism of Jones’ adrenal glands is what they do, not that they are dysfunctional. Yet, doctors do consider functional statements such as the one above about the adrenal gland to be explanatory. Moreover, they seem to consider them to be laws of nature from which judgments are deduced.

The D-N model of explanation, thus, seems to give a better account than the mechanistic approach for this specific class of medical explanations. This is a superficial understanding of these types of explanations, however. For D-N to apply here, we would need to understand “the function of adrenal glands is to produce cortisol” as a law of nature. But surely, “the function of adrenal glands is to produce cortisol” does not mean “they always produce cortisol” (Jones’ don’t). It is not a statistical law of nature either (i.e., it does not mean “the frequency of adrenal glands producing cortisol is close to 1”). In fact, something being functional or dysfunctional cannot be derived from frequency or rarity: the frequency of people dying before they turn 100
is close to 1, which does not entail that Jones surviving over 100 is dysfunctional; being left-handed or red-headed is not dysfunctional because it is rare; dental caries are not functional because they are frequent, etc. Functional statements are important as *explanans*, yet they are not reducible to general or statistical statements about reference classes.

Thus, D-N explanation does not in fact work well here. On the other hand, the mechanistic approach is blind to the functional/dysfunctional difference. Some have proposed that this normative nature of functional statements, determining what a part’s contribution ought to be, is not statistical, but rather theoretical (Wachbroit 1994). What is meant is that functional analysis results in an explanatory and predictive model, and that what is called dysfunction is simply departure from this model. In this approach, the question remains of determining how to choose the factual basis of the model.

The upshot is that functional analysis accounts for *how*-explanations of physiological processes, either healthy or pathological, but does not account for *which*-explanations (i.e., explanation of which processes are healthy and which are pathological). For this type of explanation, some turn to evolutionary explanations.

### 1.3.2. Evolutionary Explanations

Just as the theory of evolution has translated the notion of an organ’s “goal” into an explanation of how it came to being, some have hoped that it would account for which observed effects are functions and which are not, and maybe, which are healthy and which are pathological. In their view, ascribing functions can only be justified in light of evolution. Jones’ adrenal glands are dysfunctional because they ought to do what adrenal glands were selected for. In this way, the evolutionary account makes function relevant to medical explanations of dysfunction without entailing that functional statements are laws of nature (Millikan 1989; Neander 1991).

Whether evolution alone can provide *which*-explanations or not, the evolutionary sense of “function” is also relevant to medical explanations in a different sense. Philosophers of biology have long emphasized the difference between *how*- and *why*-questions. Medical explanations have been more concerned with *how*-questions than with *why*-questions about diseases. Explaining how diseases act is indeed more important, but explaining why they exist in the first place is valuable as well. The latter question has been addressed through evolutionary means by Darwinian medicine, whose agenda has been expressed in Nesse and Williams (1995). In this book, a case is made for the explanation of many diseases as resulting from a mismatch between the optimal, evolution-selected hunter-gatherer of the Pleistocene era we genetically still are and modern life. Moreover, some symptoms, such as vomiting in bacterial infection, are actually naturally selected defense mechanisms. Finally, they explain why there still are diseases—why total resistance to disease has not been the result of evolution. This paradigm has also been popular in evolutionary psychiatry, as is illustrated by Cosmides and Tooby’s work mainly since the 1990s (Barkow, Cosmides, and Tooby 1995), and discussed by philosophers of psychiatry (Wakefield 2001; Faucher 2012).

Méthot (2011) has proposed a distinction between Darwinian medicine and evolutionary medicine. Darwinian medicine is “backward-looking” and seeks for an explanation of why we get sick in the way we do at present, given what our evolutionary history (presumably) is. It is thus purely hypothetical. Evolutionary medicine, on the contrary, does not provide any theoretical framework for medicine but uses evolutionary thinking as a tool for solving specific problems. Méthot proposes the example of bacterial resistance to antibiotics, one of the few observable processes of evolution through selection, to illustrate his distinction. The theory of evolution explains and even predicts why and how resistance occurs, making resistance an example of Darwinian medicine. But evolution also suggests how resistance could
be minimized and prevented (e.g., by alternating antibiotics), making this case amenable to problem solving by evolutionary medicine.

So far, there does not seem to be any general picture of explanations in medicine. Yet most of the accounts we have discussed are general accounts of explanation that were applied to medicine rather than accounts designed for medicine. Given the deficiencies in these general accounts when applied to medicine, philosophers of medicine have developed approaches that deal with problems of explanation specific to medicine.

### 2. Specifics of Medical Explanation

Specific problems in explanation for philosophy of medicine are the role, variety, and structure of explanations in medical science. First, a focus on explanation in medicine, as opposed to many other less action-oriented sciences, may lead to the conclusion that they may be of a lesser importance, leading to the view that medicine is an engineering knowledge rather than a science. Second, some have doubted that all medical explanations are or should be mechanistic: the originality of medicine might be to resort to types of explanation as different as mechanistic and statistical explanations. Eventually, the variety of medical sciences has led some to develop the view that what is explanatory in medicine is what unifies knowledge, a view here labeled as “coherentism.”

#### 2.1. Does Medical Science Model Mechanisms or Establish Evidence?

One possible response to the problem of providing an account of explanation in medicine is to see medical science as an engineering practice, to which explanations are of minor importance. In this view, evidence for action would be of primary interest for medicine. There has indeed been a trend in recent years to think that medical science should focus on establishing evidence for effective treatments and diagnostic and prognostic tools. The so-called evidence-based medicine movement is largely responsible for this trend. To an engineering practice, explanation is deemed less essential or even largely dispensable. Establishing *that* something works does not involve establishing *how* it works. Whereas the former is done through establishing difference-making, the latter is done through the investigation of mechanisms, according to a distinction now common among philosophers of medicine. Indeed, some studies just establish *that* a factor makes a difference to an outcome; others establish *how* a mechanism links this factor to the outcome. There does not seem to be any alternative. Knowledge of difference-makers is necessary and sufficient for prediction and control.

However, even if one admits that prediction and control, not explanation, is all medicine looks for, it remains that causal claims must be established to that end: there is no prediction, nor control, of an effect E by a cause C, without knowing that C causes E. In particular, the Russo-Williamson Thesis claims that establishing evidence of causality between C and E implies establishing both *that* C makes a difference to E and *how* a mechanism may link C to E (Russo and Williamson 2007). Establishing a mechanism is *ipso facto* explaining. Some even admit that it also is the main, if not the only, way to explain in medicine:

> [C]ausal explanations are only explanatory to the extent that they can be viewed as providing a glimpse of the structure of a corresponding mechanistic explanation.

(Clarke et al. 2014)

So even if explanation is not the primary goal of medicine, the Russo-Williamson Thesis shows that it is central to medical science.
A further argument against ignoring explanation in medicine goes that causal explanation, in the form of models of mechanisms, provides a stronger and broader basis to clinical practice. Paul Thompson, who also strongly opposes the view of medicine as mere engineering, acknowledges the difference between “clinical medicine” and “basic medical science” (Thompson 2011: 115). Medical science, he insists, requires explanations, in the form of quantitative models of both basic physiological processes and epidemics, and not just evidence gathered from randomized controlled trials (RCTs). Moreover, the view that medical science would only consist in the knowledge that environmental factor X causes disease Z or that treatment Y improves it weakens medical science: isolated causal claims are stronger when they are based on integrated, multicausal models. Such explanatory models are precisely what basic medical science is interested in. Only clinical medicine is to medical science what engineering is to physical science (Thompson 2010: 273). Medicine itself, however, is not merely engineering for Thompson.

2.2. Alternatives to Mechanistic Explanations in Medicine

Whatever their position on the importance of explanation in medical science, some philosophers of medicine have raised doubts about whether mechanisms matter much in medical explanation. The great complexity of organism-environment interactions often makes a satisfactorily exhaustive mechanistic model of diseases unattainable (Howick 2011). Fortunately, mechanisms are not all there is to explanations of causal relationships, as recent work in the philosophy of epidemiology demonstrates.

A traditional view of epidemiology is that it is both a descriptive science of population-related aspects of disease and an empirical inquiry into their potential causes, among which is the environment, that does not need to establish mechanisms. “Epidemiology is the study of the distribution and determinants of disease frequency in human populations” (Rothman, Lash, and Greenland 2012). It contains no theoretical framework, only methods. Therefore, the discipline is often rather seen as heuristic, descriptive, and predictive at best, rather than as explanatory.

Recently, however, Broadbent has developed an approach to understanding epidemiological explanations as truly explanatory (Broadbent 2013). Epidemiology is an inquiry into what associations signal causation. According to Broadbent, in epidemiology, a cause should be interpreted as an explanatory factor. In particular, it explains a difference in observed effects in two contrasted groups of people, but it does not consist in the disclosure of the complete causal network that led to the presence of an effect. Exposure may thus explain the difference between exposed and unexposed samples, say, the presence or absence of an infectious syndrome when exposed, or not exposed, to an infectious agent. It does not, however, explain why exposure, rather than non-exposure, produces the effect. This latter task would be much more demanding. To prove that exposure, rather than non-exposure, produces the syndrome, one would have to know what else could possibly cause it, whereas establishing that exposure produced the syndrome rather than the absence of the syndrome, one only has to show a difference in outcome when exposed and unexposed. Epidemiological explanation is also less demanding than mechanistic explanation, as epidemiological explanations admit of so-called black-box approaches to causation (i.e., the claim that to establish causality, one does not have to identify the mechanisms).

This suggests that epidemiological explanations are an alternative to mechanistic explanations. The same sort of alternative to simple mechanistic explanation is provided by quantitative genetics and genome-wide association studies, that is, systematic investigation of parts of the genome that may be associated with a given disease. Infectious diseases again
provide interesting examples: for instance, only one out of ten people infected with influenza virus develops the flu. Darrason (2013) points out that multiple factors play a role in explaining such differences, one of which being genetic variability. Genetic variability is expressed by the notion of heritability. Heritability is a ratio of phenotypic variance to genotypic variance: it expresses how much of the former is explained by the latter, but as a statistical concept, it bypasses the individual genes and the various molecular pathways by which the variation of a trait such as susceptibility to flu explains. Yet, heritability provides an explanation of which patients are susceptible to the flu. It is a non-mechanistic explanation, as both Darrason (2013) and Schaffner (2016) have emphasized. Indeed, explanation by heritability is only contrastive, of the sort Broadbent emphasized in epidemiology: it expresses a proportion of differences in populations. The same could be said of recent approaches based on Big Data Biology: they seem to provide a statistical sort of explanations as compared to mechanistic explanations. However, it remains an open discussion whether epidemiology and data-driven approaches really provide explanations, and not just descriptions or predictions (Ratti 2015).

2.3. Coherence Approaches to Explanation in Medicine

Contemporary analyses of medical explanations have led to the observation that explanations are somewhat subordinate to therapeutic action and that there may be at least two sorts of explanations in medicine, statistical and mechanistic. This suggests that medical knowledge might be more pragmatic and eclectic than other sciences such as physics. In fact, just as unification of knowledge has been proposed in the general philosophy of science as defining what explanation consists in (Kitcher 1981), coherence of apparently pluralistic knowledge may reveal that what is explanatory, is a unificatory schema that provides coherence to various beliefs held by medical scientists.

Thagard articulates such a view about the example of the science of peptic ulcer (1999). Until the 1980s, peptic ulcer was understood as the result of high acid concentration in the stomach or duodenum, which in turn was explained by stress. These hypotheses were supposed to be robustly established by converging observations and meshed well with traditional representations of the “bilious type” in medicine and their consequences in popular cultures. This was coherent and held as explanatory. In the early 1980s, Warren and Marshall discovered the existence of a bacterium in the stomach, *Helicobacter Pylori*, that survives the low pH of this environment. They then observed that almost all patients with peptic ulcer are hosts to this bacterium, although some of those who host it are not affected. Ultimately, they showed that antibiotic treatment was efficient on almost all cases of peptic ulcer. Their conclusion was that *H. Pylori* infection explains peptic ulcer.

To be explanatory, the coherentist view goes, this belief had to provide a coherent picture of all other beliefs held about peptic ulcer; it could not be explanatory if it remained isolated from all other widely held beliefs about peptic ulcer. According to Thagard’s view, “coherence of beliefs” is defined as the “positive constraint” (Thagard 1999: 67) that two hypotheses exert on each other, so that adopting either one both forces one to adopt the other and to reinforce it. For example, before Warren and Marshall’s discoveries, the accepted hypothesis that bacteria could not survive in environments as acidic as the stomach was not coherent with the hypothesis that *H. Pylori* colonizes the stomach; after their discoveries, the coherent picture is that infection to *H. Pylori* causes most peptic ulcers, that treatment against *H. Pylori* cures most cases, and that stress plays a minor part in the disease, if any (Thagard 1999: 135–47). Progress in explanation has thus been made, in that this wider set of consistent beliefs can now be held.

Thagard also notes that “unified understanding does not come from the availability of a general overarching theory but from the availability of a system of explanation schemas” (Thagard...
1999: 34–5). For Thagard, explanations are based on a disease explanation schema consisting of an explanation target (i.e., an answer to the question “why does a patient have a disease with associated symptoms?”) and an explanatory pattern (i.e., the demonstration that the patient “is or has been subject to causal factors” and that “the causal factors produce the disease and the symptoms”) (Thagard 1999: 20). This explanation schema may be fulfilled by many different approaches. One example is the Germ Theory Explanation Schema:

“Explanation target
Why does a patient have a disease with symptoms such as fever?

Explanatory pattern
The patient has been infected by a microbe.
The microbe produces the disease and the symptoms.”

(Thagard 1999: 25)

These statements are not interpreted by Thagard as stating (mechanistically related) facts, but as expressing beliefs. Other medical explanations may rely on different entities than microbes, for instance, presence or absence of a nutrient, genes-RNA-protein and mutations, etc. The form of the explanation schema remains the same; that is, it always relies on the same sort of cognitive arrangement of beliefs. These beliefs, however, may evolve with time and the succession of dominant theories; that is, they are also under the influence of social processes that spread them. In other words, the fact that beliefs are widely held is also important to their being considered as explanatory.

Thagard's approach to explanations is cognitive—it accounts for why arguments are accepted as explanations, rather than accounting for their logical structures. It is also social, in that what makes a causal reasoning an explanation are social beliefs shared by a community. Of particular interest here is the case of the adoption of a new theory, as illustrated by the explanation of peptic ulcer by infection. Adoption of any theory is explained following the same explanation schema:

“Integrated Cognitive-Social Explanation Schema
Explanation target:
Why did a group of scientists adopt a particular set of beliefs?

Explanatory pattern:
The scientists had a set of mental representations that include a set of previous beliefs and a set of interests.
The scientists’ cognitive mechanisms included a set of mental procedures.
The scientists had social connections and power relations.
When applied to the mental representations and previous beliefs in the context of social connections and power relations, the procedures produce a set of acquired beliefs.
The scientists adopted the acquired beliefs.”

(Thagard 1999: 9)

Thagard’s cognitive view also emphasizes that part of the difficulty of medical explanation comes from the existence of multiple types of explanations to medicine. Heterogeneous beliefs
about various entities involved are generally held about the same disease: recall the various beliefs about peptic ulcer before the H. Pylori revolution.

Lemoine (2011) also endorses a cognitive, coherentist view of explanations. However, he emphasizes the fact that all beliefs about the same disease do not have to, or even cannot, be coherent. Lemoine suggests that only subsets of beliefs about the same disease have to be coherent with one another. They are defined by different, incommensurable scientific views of explanation, which can converge or diverge about a specific disease. These scientific views of explanation, Lemoine calls “explanatory values.” They state what counts as a sufficient explanans (i.e., set of facts) for the explanation to be satisfactory in their perspective. For instance, the clinical explanation that “patient p presents signs x, y, and z because he suffers from disease entity D” is satisfactory only if all relevant signs in the exhaustive list of semiology provide, have been observed, and if diagnostics has been correctly conducted. According to a different explanatory value, the explanation is considered satisfactory only if a continuous mechanism from etiology to symptomatology has been established. Lemoine distinguishes several other types of explanation—pharmacological, evolutionary, infectious, etc.—and shows that each states the specific set of conditions accepted as sufficient for a satisfactory explanation of one sort.

Lemoine also explores the various possibilities that explanations of various types converge, diverge, or just coexist, which he dubs the “disunity of medicine.” He considers the plurality of medical explanations to play an important role in medicine, as strengthening or weakening evidence.

**Conclusion**

A satisfactory account of medical explanations should be explicit about what is specifically medical (if anything), whether all medical explanations are of the same sort, and what the role of explanations is in medical science. Philosophical accounts of scientific explanations in general cast light on the logical structure of medical explanations in particular, and also reveal something of their diversity: D-N, D-S, I-S, and S-R explanations are resorted to in medicine, albeit in a modified version. Mechanistic explanations focus on causal relations and come in two forms: functional and evolutionary. These investigate respectively how a process works and why it is there in the first place. Despite the relevance of these accounts of explanations in medicine, they gloss over the specific problems medical explanations raise. These pertain to the role of explanation in medicine—is it prominent or not, is medicine a search for knowledge, or is it imbued with pragmatism?—and to the question whether all explanations should ultimately be mechanistic in nature. In the end, this chapter suggests that from a cognitive point of view, how distinct types of explanations relate to each other is critical to understanding the specificity of medical science when it comes to explaining. Thagard and Lemoine both think that there is no unitary theory of diseases from which explanations are derived. However, Thagard emphasizes the need for coherence of different beliefs, whereas Lemoine emphasizes the usefulness of convergence of differently grounded explanations.

**References**


**Further Reading**