Applying Medical Knowledge: Diagnosing Disease

William E. Stempsey*
Department of Philosophy, College of the Holy Cross, Worcester, MA, USA

Abstract

The term “diagnosis” can refer to the name of a disease that afflicts a person or to the process of determining a diagnosis in the first sense. Complex philosophical, metaphysical, epistemological, normative, and logical issues permeate all aspects of diagnosis, beginning with the question of what is being diagnosed. The nature of disease is philosophically controverted. Both ontological and physiological conceptions of disease continue to influence thinking about disease and hence how to diagnose it. The question of whether the concept of disease is essentially value laden remains open. Diagnosis presupposes some classification of diseases, known as a nosology, in order to distinguish one disease from another. There are many possible ways to classify diseases, and nosologies can have different goals, e.g., providing a basis for rational treatment and prognosis, enabling statistical reporting, fostering research, and for administrative aspects of health care. The major elements of the diagnostic process include the history of the illness, the physical examination, and various kinds of laboratory and clinical testing. Each of these elements requires interpretation and is influenced by philosophical presuppositions. Diagnostic reasoning makes use of probabilistic, causal, and deterministic models. It is fundamentally a process of hypothesis formation about possible diagnoses (differential diagnosis) and the systematic confirmation and ruling out of possibilities until one diagnosis is judged best to explain all the data. Diagnosing disease is important for a broad array of medical and social reasons.

Introduction

The term “diagnosis” can refer to the name of a disease or biomedical problem afflicting a person or to the process of determining the presence and nature of the disease. Although the process of diagnosis has been extensively studied and there is general agreement on basic aspects, the complexity of many real-life situations makes it difficult to formulate a universal description of the diagnostic process. One reason for this is that complex philosophical issues permeate all facets of diagnosing disease, including the question of just what disease itself is. This chapter addresses metaphysical, epistemological, normative, and logical issues in three aspects of diagnosing disease: the nature of disease, nosology or the classification of diseases, and the diagnostic process and logical models of diagnostic reasoning.

Disease

If diagnosis is a search for the presence and nature of disease, the nature of diagnosis will depend upon conceptions of disease and how diseases are classified. This section deals with the concept of disease and the following one with classification. The word “disease” can refer to the class of all diseases or to various subsets of that class, such as pneumonia, diabetes mellitus, arthritis, etc., and further subsets, such as bacterial pneumonia, viral pneumonia, rheumatoid arthritis, septic arthritis, etc. “Disease” can also be

*Email: wstempse@holycross.edu
used to refer to a single instance of one of the above. Other related terms, such as illness and sickness, are often used in common discourse as synonyms but can also have precise meanings and relationships that vary in different theories of health and disease. This chapter focuses on disease as distinguished from related terms such as illness and sickness in the standard biomedical model. Disease refers to a set of biological phenomena that are said to be the cause of a person’s experience of illness, which is feeling unwell. The social ramifications of disease and illness are described by the term “sickness.”

The metaphysics of disease remains controversial. Historically, there have been two fundamental conceptions of the nature of disease qua disease: physiological and ontological (Temkin 1963, p. 631; Engelhardt 1975, pp. 125–141). The classic physiological conception of disease is the humoral theory that goes back to the time of Hippocrates and dominated for centuries. It presupposes a biologically conceived teleology. When the body is functioning in accord with its nature, there is a proper balance of the four humors: blood, phlegm, black bile, and yellow bile. Disease is deviation from the normal healthy state, that is, an imbalance of the humors (Temkin 1973, p. 398). Even though the classical humoral theory has been long abandoned, the physiological theory of disease has remained influential in concepts such as homeostasis and the idea of disease as deviation from the state of normality called health.

While physiological conceptions take diseases to be deviations from some normal state, ontological conceptions take diseases to be things in themselves. This does not necessarily mean that diseases are concrete things. For example, Thomas Sydenham (1624–1689) held that diseases are observable clusters of signs and symptoms but that they cannot be localized to any particular organ in the body. Still, these clusters display regularity from individual to individual and were conceived as unchanging abstract objects similar to species of plants (Nordenfelt 1995, pp. 152–153). Other ontologists take disease to be actual physical entities. With the discovery of the association of bacteria, parasites, and the like, with certain clusters of symptoms, some ontologists identified these invading organisms as the disease. Another ontological view is that of Rudolf Virchow (1821–1902), the early champion of cellular pathology. Virchow at first repudiated the ontological conception of disease in favor of a physiological one. In 1847, he wrote that diseases were only physiological phenomena under altered conditions, but by 1895 he was calling himself a “thoroughgoing ontologist,” regarding pathological cells within the body as the disease itself and not merely the cause of the disease (Virchow 1958, pp. 26, 192; Stempsey 2000, p. 72).

While the physiological and ontological conceptions of disease continue to influence certain aspects of contemporary diagnosis, current thinking is more focused on historical development of concepts of disease. Disease today is more likely to be seen as a process, influenced by current scientific thinking and cultural influences, emphasizing cause, the bearer of the disease, and the set of manifestations (Nordenfelt 1995, pp. 172–173). For example, since the completion of the mapping of the human genome, the concept of genetic disease has seen great emphasis.

One of the major issues philosophers of medicine have explored is the question of whether the concept of disease is essentially value laden. Two opposing positions are often referred to as naturalism and normativism. Naturalists, most prominently Christopher Boorse (1997) in his influential biostatistical theory, believe that the concept of disease is purely descriptive; disease is a value-free scientific concept. Normativists, who have proposed a diverse set of theories, take various kinds of values to be essential components of the concept of disease (Engelhardt 1975; Nordenfelt 1995; Stempsey 2000).

In short, contemporary thinking about the concept of disease can only be described as complex (Hofmann 2001). It is being reframed contextually, based on metaphysical, epistemological, and axiological commitments, to reflect the pathological processes that afflict human beings, bring them to the attention of medical professionals, and serve as warrants for treatment (Cutter 2003).
Nosology

If diagnosis is about explaining a set of symptoms (subjective experiences of feeling unwell) and signs (objectively observable phenomena) in terms of some particular disease, then the question of how one disease is differentiated from another is foundational for diagnosis. The classification of diseases is known as nosology. A fundamental philosophical question for nosology is whether disease classifications mirror independently existing diseases in the realist sense or whether classifications are constructed for various purposes and values. Whether the class of diseases constitutes a natural kind is still debated; Reznex (1987) has denied this. However, even if disease were a natural kind, it is clear from an examination of various existing nosological systems that disease classifications aim at much more than simply trying to map entities that exist in a realist sense.

There are several desirable characteristics for any nosology (Murphy 1997, pp. 122–126): (1) Disease categories should correspond to naturally occurring sets of characteristics seen in particular diseases. (2) A classification ought to be exhaustive, i.e., it should include all the conditions for which people seek medical help. (3) Categories should be disjoint. That is, no particular case should fall into more than one category. (4) The classification should be useful for understanding disease mechanisms, categorizing descriptive features, fostering effective treatment or management, and determining a prognosis and for purposes of such matters as administration, law, and education. (5) The classification should be as simple as possible and still achieve its goals. (6) The classification ought to be constructible, in the sense of allowing both exhaustiveness and disjointness. It is clear that these are ideal characteristics; they cannot all be achieved simultaneously. For instance, complete exhaustiveness and disjointness are in practice impossible to achieve. Most diseases can be classified in several ways, e.g., according to etiology, symptoms, or anatomical location. Exhaustive classifications will necessarily fail to be disjoint. Decisions about which ideals should take priority will have to be made according to usefulness for the particular intended purposes of the nosology.

Thomas Sydenham (1979), in the seventeenth century, and François Boissier de Sauvages (1768), in the eighteenth, constructed nosologies that sought explicitly to be exhaustive in scope. Assuming that diseases were ontological kinds, they relied on empirical observation over rationalistic systems, believing that this would provide a basis for rational treatment. René Laennec (1982), in the late eighteenth century, advocated an anatomico-clinical approach, showing that a nosology based purely on symptoms would not be exhaustive. He conceived of the human organism as consisting of three parts: solids, liquids, and the principe vital, an animating force. All diseases were classified as lesions in one of these parts. (Stempsey 2000, pp. 109–110).

Such nosologies are of interest today primarily for historical reasons, but they do reflect the ongoing interest in classifying diseases according to a set of ideals and also emphasizing the practical usefulness of a nosology for rational treatment. An examination of contemporary disease classifications shows a commitment to the same sorts of ideals but with a greatly expanded set of particular purposes.

The International Statistical Classification of Diseases and Related Health Problems (ICD) is one of a family of international classifications of health, disease, disability, and health interventions produced by the World Health Organization. The current, tenth revision (ICD-10), is intended to assist in the “systematic recording, analysis and interpretation and comparison of morbidity and mortality data” across time and place (WHO 2010, v.2, sec. 2.1). Thus, its primary purpose is to facilitate statistical reporting of morbidity and mortality. The classification is divided into 21 chapters. Chapters I–XVII cover local diseases, arranged by the main systems of the body. Chapter XVIII includes symptoms, signs, and abnormal clinical and laboratory findings that are not elsewhere classified. Chapter XIX covers injuries, poisoning, and other consequences of certain external causes. Chapter XX includes other external causes.
of morbidity and mortality. Chapter XXI classifies data explaining reasons for contact with health services by a person not currently sick, or circumstances in which a person is receiving care.

The Systematized Nomenclature of Medicine (SNOMED) is a comprehensive, multilingual collection of clinical health-care terminology. It is owned and maintained by the International Health Terminology Standards Development Organisation (IHTSDO), a not-for-profit association governed by its national members, twenty-seven countries as of April 2014, and headquartered in Denmark (IHTSDO 2014a). The current version, SNOMED CT (Clinical Terms) is a computer-based terminological system. Terminological systems provide terms denoting concepts and their relations from a specific domain and can be used to describe information in a structured and standardized way. SNOMED CT enables consistency in indexing, storing, retrieving, and aggregating clinical data across specialties and health-care venues. It enables computerizing medical records, thus providing consistency in the way data is stored, encoded, and used for clinical care and research (Cornet and deKeiser 2008). Thus, it serves primarily a clinical purpose: it brings consistency to the way patient data is stored.

SNOMED grew out of the Systematized Nomenclature of Pathology (SNOP), developed by the College of American Pathologists (CAP) in 1965 and later extended to other medical fields as SNOMED. In 2007, the intellectual property rights for SNOMED CT and previous versions of SNOMED and SNOP were transferred from the CAP to the IHTSDO (IHTSDO 2014b).

SNOP was intended to assist pathologists in standardizing terminology in the cataloguing of specimens. SNOP presumes an anatomico-clinical model of disease, in which all diseases can be described by some anatomical change. Although SNOP did not claim to be clinically oriented, it did provide an exhaustive classification of disease, if disease is understood in that particular anatomico-clinical model. Diseases are described with respect to four fields: topography (the part of the body affected), morphology (the structural changes produced in the disease), etiology (the etiologic agent responsible for the disease), and function (the manifestations of the disease). Within each field, terms are given a number up to four digits, the number of digits reflecting increasing specificity. For example, if a small bowel specimen shows ulceration (M4003) in the ileum (T65) with recovery of Salmonella typhi (E1361), and the patient shows clinical manifestations of disease (F9497), the specimen is coded as T65-M4003-E1361-F9497 (Stempsey 2000, pp. 112–114). SNOMED in 1979 expanded the “SNOP concept” with three additional fields: a disease field was added to record discharge diagnoses for statistical reporting purposes; a procedure field to record administrative, diagnostic, and therapeutic and preventative procedures; and an occupation field to formally report a patient’s work for purposes of specialties like industrial medicine (Stempsey 2000, pp. 114–115). Thus, even early versions of SNOMED have already moved beyond simple classification.

The current version, SNOMED CT, refines the primary purpose of this family of classifications with its emphasis on making health records accessible electronically and meaningful for clinical and administrative uses. Content is presented using three components: concepts, descriptions, and relationships. Concepts are arranged into hierarchies (clinical findings, procedures, body structures, organisms, physical objects, physical forces, events, social context, etc.) from the general to the more detailed; each concept has a unique numeric identifier. Descriptions link other appropriate terms to concepts. A concept can have several associated terms that describe the same clinical concept. Every description has a unique numeric description identifier. Relationships link concepts to other concepts. One example is the “is-a” relationship, which can be used to relate a concept to more general concepts. For instance, the concept “infective pneumonia” bears the “is-a” relationship to the more general concept “pneumonia,” and both “bacterial pneumonia” and “viral pneumonia” bear the “is-a” relationship to the more general “infective pneumonia.” SNOMED CT thus allows retrieval of information about many elements of disease and its clinical management with a high degree of complexity. This enables a wide range of clinical meanings to be
captured in a record, without requiring the terminology to include a separate concept for every detailed combination of ideas that may be relevant to a particular case (IHTSDO 2014c).

The entire question of disease and nosology in psychiatry is more controverted and will only briefly be considered here. The fifth chapter of ICD-10 is devoted to mental disorders and serves as a worldwide standard, but the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders (DSM) is perhaps more influential and has received more notoriety for its sometimes very significant revisions. The current, fifth major revision (DSM-5) was published in 2013. Successive revisions of the DSM have increasingly moved away from Freudian understandings of psychiatry to a more scientific, or at least evidence based, approach to understanding psychiatric disease and its classification. A major problem, however, is that the etiology of a great number of mental disorders is inadequately understood and so cannot serve as the basis for nosology as is increasingly the case with somatic disease. Furthermore, the obvious social dimensions of mental disorders and the difficulty in standardizing such dimensions create problems in trying to establish precise disease categories for mental disorders. Social norms change, and this raises the question of whether psychiatric diagnostic categories simply reflect current norms or are naturally occurring entities more like somatic disease entities. The inclusion and then removal of homosexuality as a mental disorder, for example, has led some to suggest that psychiatric diseases are simply reflections of current social mores and defined by committee fiat. The situation is far more complex, however, and will not be further considered here. It is simply noted that the purpose of psychiatric nosology remains oriented toward the same practical goals as other types of nosologies (Jablensky 2012, pp. 77–94).

Current nosology thus is oriented toward very different goals than were the first attempts to categorize all diseases centuries ago. Current classifications are primarily reporting mechanisms. Although they still recognize the practical import of nosology as facilitating effective care of those who suffer, they make no claims at offering frameworks for choosing treatments. Rather, they serve primarily as means for categorizing disease for purposes of research that will foster public health and for administrative assistance in health-care management and payments.

Diagnosis

The process of diagnosis includes three basic elements: the history of the illness, the physical examination, and laboratory and other sorts of clinical tests. In a typical case, a physician “takes a history” from the patient; develops hypotheses about the diagnosis; performs a physical examination; generates a differential diagnosis, a list of possible diagnoses that fit the hypotheses; tests the hypotheses by laboratory and other sorts of clinical tests such as x-ray examinations; modifies the differential diagnosis; and repeats these various steps until a diagnosis is arrived at (LeBlond et al. 2009, pp. 11–12). This section first considers the three elements and then turns to the logic of diagnosis and the process of reasoning that physicians follow.

History of the Illness

“Taking the history” refers to the initial conversation of physician and patient in which the physician listens to the patient’s relating of why he or she is seeking medical help and asks questions to help clarify the information provided by the patient. Some have recommended that the term “taking a history” is misleading and should be replaced because it puts the patient in a merely passive state (Lazare et al. 1995, p. 18), but “taking the history” is still most widely used.

Patient assessments, which include the history and other elements of diagnosis, can be comprehensive or focused. Comprehensive assessments are appropriate for new patients, whether in the hospital or in an
outpatient setting. A comprehensive health history can provide fundamental and personalized knowledge about the patient and strengthen the clinician-patient relationship. Comprehensive initial assessments can also provide baselines for future assessments. Focused assessments are more appropriate for established patients, especially during routine or urgent care visits. They address particular concerns or symptoms, often restricted to a specific body system (Bickley and Szilagyi 2013, pp. 4–6).

The comprehensive adult health history seeks both subjective information, such as the patient’s experience of pain, and objective information, such as age and family history. The history should include identifying data such as age, gender, occupation, marital status, and an assessment of reliability, which may vary according to the patient’s memory, trust of the physician, and other social factors. The initial focus is on the chief complaint or complaints, what it is that brought the patient to seek medical care. Several aspects of the patient’s history are then sought. The history of the present illness includes the symptoms, how and when each symptom developed, and the patient’s thoughts and feelings about the illness. The physician’s questioning may seek to identify pertinent positives and negatives to aid hypothesis formation and to find out about the patient’s medications, allergies, and history of smoking and alcohol use. The past history includes childhood illnesses; adult illnesses, whether medical, surgical, obstetric/gynecological, or psychiatric; and health maintenance practices such as immunizations, screening tests, and lifestyle issues. Family history includes age and health or cause of death of siblings, parents, and grandparents and the presence of any genetic diseases or other diseases that are disposed to run in families. Personal and social history includes educational level, family of origin, current household, and personal interests and lifestyle. These elements can provide important clues in the diagnostic process. The final part is the review of systems, in which the physician systematically asks about the presence or absence of common symptoms related to each major body system (Bickley and Szilagyi 2013, pp. 6–13).

The medical interview has three functions (Lazare et al. 1995, pp. 3–19). The first is determining and monitoring the nature of the patient’s problem with the objective of enabling the clinician to establish a diagnosis, recommend further diagnostic procedures, suggest a course of treatments, and predict the nature of the illness. The medical interview enables a physician to generate multiple hypotheses during the course of the interview and elicit additional data to refute or confirm them. The history is estimated to contribute 60–80% of data for diagnosis. The second function is focused on developing, maintaining, and concluding the therapeutic relationship. It helps to define the nature of the physician-patient relationship; communicate professional expertise, interest, respect, support, and empathy; and elicit the patient’s perspective on the problem and the methods and goals of treatment. This may indirectly help the physician to glean effective diagnostic data from the interview. The third function is patient education and implementation of treatment plans with the objective of fostering consensus, patient satisfaction, cooperation, and improved treatment outcome. The importance of each function of the interview varies according to the nature of the interview, but the three functions are often interdependent. This functional analysis calls attention to the dynamics and complexity of the medical interview. The medical interview, or taking the history, has multiple purposes that go beyond mere diagnosis. The diagnostic process is ultimately carried out as a means for providing the most effective treatment for the patient and enabling the physician to offer a prognosis. Treatment involves providing drugs, surgery, and the like, but also developing a healing relationship between doctor and patient.

Kathryn Montgomery Hunter (1991, p. 5) sees all of medicine as fundamentally narrative, but most important are the opening stories that patients tell their physicians. She is representative of the school of thought that holds the nature of medicine to be better explained by the methodological analogue of literature than by the natural or social sciences. Even if one does not accept this narrative analysis for all aspects of medicine, it does seem apt for certain activities, especially the medical interview. In this type of literary analysis, patients are seen as texts. While patients are ordinary readers of their texts, physicians are more like sophisticated interpreters. Physicians make sense of signs using a “diagnostic circle” very much
like the hermeneutic circle, whereby parts of a text can be understood only with reference to the whole, and the whole can be understood only with reference to the parts (Hunter 1991, p. 9). Thus, with both patient and physician as readers, there is “one illness, two stories” (Hunter 1991, pp. 13–15). The patient’s account of illness and the medical version of that account are fundamentally, irreducibly different narratives, and this difference is essential to the work of medical care. The “medical plot,” the narrative organization of the case, is fundamentally shaped by the search for a diagnosis and an answer to the question of what is the best treatment for this particular patient (Hunter 1991, p. 65). The patient’s story is itself given a sort of medical treatment, being retold in light of the physical examination and clinical tests; in the process, it becomes a fundamentally different story (Hunter 1991, pp. 128–130). The rewriting of the patient’s story becomes part of the healing process (Hunter 1991, pp. 138–141). This sort of analysis reinforces the importance of seeing diagnosis in the context of a larger therapeutic relationship. It also highlights the importance that interpretation has in all aspects of diagnosis.

Physical Examination
The physical examination is the second major element that contributes data to the diagnostic process. A comprehensive presentation of the goals and processes of carrying out the physical examination can be found in well-established texts on physical diagnosis such as LeBlond et al. (2009) and Bickley and Szilagyi (2013); a detailed description of them will not be given here. The physical examination is often considered to be a purely scientific matter of observation and a source of objective data, opposed to the subjective data that patients report during the medical interview. However, there are several conceptual issues underlying the physical examination that influence how physical findings are perceived and interpreted. First, the observations made by clinicians during a physical examination are first of all perceptions of the examiner. As such, they are inherently subjective. As has been generally recognized by philosophers of science, all observations are fashioned by the underlying theories that are brought to the observations. To take one classic example, Koplik’s spots are considered to be pathognomonic for measles. These are small, bluish-white specks on an irregular red background in the buccal mucosa and occur early in the course of measles. Whether particular bluish-white specks should be considered Koplik’s spots is a judgment that depends on the subjective perceptions and judgments of the diagnostician. Such judgments already depend on conceptual commitments about the nosological entity of measles (Engelhardt 1981, pp. 305–306). Psychological factors, such as assumed probabilities and distortions due to expectations, are also known to influence perceptions of what is and what is not observed (Kahneman and Tversky 1982, pp. 144–146).

Diagnostic Testing
The third element of diagnosis consists of the many sorts of laboratory and other clinical tests carried out for the purpose of confirming or ruling out the hypotheses arrived at during the history and physical examination. In light of the ambiguity that can result from the perceptual issues inherent in the physical examination, laboratory and other clinical test results are often preferred and considered to be more objective. Many of the same conceptual issues, and even more complex ones, enter into the interpretation of test results, however.

Clinical tests can include laboratory testing of blood and body fluids such as urine, surgical biopsies of tissues, and clinical examinations such as x-ray studies, computerized tomography, echocardiography, sonography, and other types of procedures.

There are several purposes for which such testing might be done. First, laboratory and clinical tests can be a way to mitigate problems of interobserver variability in both the medical interview and the physical examination. Second, testing allows physicians to extend their medical examinations below the surface of the body. Third, quantitative assessment of levels of bodily constituents can give insight into the
physiological processes occurring within the body. Fourth, testing such as bacterial and fungal culturing or toxicological screening can identify the etiological agents responsible for disease. Fifth, laboratory testing of body components can identify markers that might indicate risk for future development of disease. With the rapid development of biotechnology focused on the gene, this last purpose is assuming increasing importance (Stempsey 2000, pp. 149–150).

Two aspects of philosophical importance in laboratory and clinical testing are the choice of tests and the interpretation of tests. Choice of whether to do a test at all and choice of the particular tests to be done involve normative elements, which can be understood in terms of a complex cost-benefit analysis. First, the value of the test for confirming or excluding the particular diagnosis must be considered. Second, the consequences for the patient of including or excluding the diagnosis must be considered. Some diagnoses may be of little importance for the patient, but others, such as being HIV positive, bring serious social consequences. Reducing uncertainty or identifying risk may be important to some patients but not to others. Third, the risk of the diagnostic procedure itself to the patient has moral import. Tests bring discomfort and risk of harm to the patients. The importance of gaining information must be balanced against the potential harm of the test. Many routinely done tests expose patients to radiation, for example. Even a routine blood draw exposes a patient to a very small chance of bruising and infection. To take a more serious case, it is hard to see how one could justify exposing a patient to the dangers of a brain biopsy if the information that would be obtained could have no influence on decisions about treatment or prognosis. Fourth, testing incurs economic costs including wages to health-care personnel, cost of instruments and equipment, and possibly costs of hospitalization and loss of income to the patients. Fifth, if enough tests are performed, it is likely that conflicting results will be obtained. This is usually resolved by getting more tests, which might resolve the conflict but might also exacerbate it (Wulff 1976, pp. 109–110).

The second important philosophical aspect of laboratory and clinical testing is the interpretation of test results. As already mentioned, test results can conflict with one another. This problem is made worse by the easy availability of technology, such as automated machines that can produce many test results from one very small blood sample. Even when there is no conflict of data, results must still be interpreted for significance. Quantified data present particular problems. Precise numbers can suggest a degree of certainty that should not be presumed. In addition, establishing a range of values that is considered to be normal presents the difficulty of correlating population data and drawing inferences from that data for individual patients. These issues are further addressed in the next section.

**Diagnostic Reasoning**

The data obtained from the three elements of history, physical examination, and diagnostic testing must be gathered and interpreted to formulate a diagnosis, i.e., a name of a disease that best explains the data. This process is diagnostic reasoning. Diagnostic reasoning proceeds from effect to cause, the opposite logical direction used in explaining pathogenesis. Its process cannot be depicted by any one simple logical scheme for several reasons: because nosologies differ and change over time, because diagnosis has different end goals that vary in different clinical situations, and because many patients have multiple diseases rather than just one (Feinstein 1973a, pp. 212–232). For simplicity, “data” here refers to the evidence gleaned from the history, physical examination, and various clinical tests. The diagnostician must authenticate data, decide whether the data deviate from some designated state of normality, and consider the pertinence of the data for the goal of the particular diagnosis. Inferential reasoning then proceeds toward the goal of a diagnostic category and ends when the ultimate end goal, whether it is the ability to make a prognosis or to render a rational treatment, is reached (Feinstein 1973b, pp. 264–283).

It has been shown that physicians use many different strategies and make extensive use of heuristics in order to reach the end goal (Elstein et al. 1978, pp. 252–272; Tversky and Kahneman 1982, pp. 3–20). The
process of diagnosis focuses on one diagnostic hypothesis or perhaps more than one. Hypotheses might be general or specific. Reasoning proceeds by progressive modification and refinement of the hypotheses. The arrived at diagnosis is then assessed for coherency, adequacy, and parsimony (Kassirer 1989, p. 894).

Psychologists have carried out extensive studies of how diagnosticians reason. Expert diagnosticians generate diagnostic hypotheses early. The hypotheses they consider are limited in number, rarely exceeding five. Physicians vary considerably in diagnostic effectiveness, depending on the nature of the problem at hand. Hence, diagnostic competence is not simply a characteristic of an individual diagnostician but is case dependent. Experience is also found to be a basic element of competence. Expert diagnosticians have knowledge of how findings relate to diseases or conditions, the relative frequency of the possible conditions in the population, and the particular characteristics of those conditions that carry severe risk, even if their occurrence is low. Effective diagnosticians are able to retain all this information and use it correctly when needed. This capacity has been shown to be an outcome of repetitive practice (Elstein et al. 1978).

Although it is probably impossible to give an exact description of a precise reasoning process for every expert diagnostician in every case, three types of diagnostic reasoning are commonly used (Kassirer 1989, p. 894). First, probabilistic reasoning focuses on the probability of a particular diagnosis given the evidence. Assessment using only terms such as “likely,” “common,” and “rare” is problematic because such terms are vague and have no standard meaning. Probabilistic diagnostic reasoning turns to quantitative methods as more satisfactory. An ideal test would give an unequivocal answer that confirms or rules out a diagnostic hypothesis, but this is a rare occurrence except in cases where a disease is defined by a test result. Ascertaining the probability of a diagnosis given a positive or negative test result is covered by a rule formulated by Thomas Bayes in the eighteenth century. According to Bayes’s theorem, the probability of a diagnosis in a particular patient depends on other probabilities: the prevalence of the diagnosis in the population from which the patient comes, the sensitivity of the test (the proportion of positive test results in people who have the disease), and the specificity of the test (the proportion of negative test results in people who do not have the disease). Thus, test results can be interpreted according to mathematical principles that are easily derived from the most fundamental laws of probability. This approach can be valuable for the diagnostic process, but its limitations must also be recognized. For example, suppose a test to detect a particular cancer has a false-positive rate of just 1 % (specificity = 0.99). Suppose, further, that the prevalence of the cancer in the adult population is known to be 100 per 100,000. Thus, of 100,000 people, 99,900 do not have cancer. If the test were administered to those without cancer, 1 % or 999 would have a false-positive test result (Bradley 1993, pp. 70–90). This alone shows the problem of relying on test results at face value and the need for understanding their statistical basis to allow proper interpretation. But there are also other limitations in this sort of probabilistic reasoning. The prevalence of a disease in a population is not always known; when this is the case, it must be subjectively estimated. In addition, many results cannot be described simply as positive or negative; continuous variables must be broken into discrete intervals to use in calculations. Bayesian calculation also depends on the assumption that diseases are mutually exclusive, which is problematic in many cases. In addition, certain diseases manifest themselves in stages and cannot be considered simply as present or absent (Kassirer 1989, p. 895).

The second type of diagnostic reasoning is causal reasoning, which is especially valuable for its explanatory power. Causal reasoning relies on the common sense notion of cause-and-effect relations between variables. In diagnostic reasoning, it focuses on describing anatomical, physiological, and biochemical mechanisms of the normally functioning human body, the body’s pathophysiological behavior in disease, and idiosyncrasies of individual patients. Causal models (e.g., fluid-electrolyte equilibrium) are generated, usually relating stimuli and responses. The process of testing, verifying, and falsifying hypothetical causal connections is a fundamental aspect of diagnosis. Causal reasoning can
also be useful in setting the context for future data gathering in the diagnostic process. It can help in verifying a diagnosis and assessing its coherency. Major benefits of causal reasoning in diagnosis are its explanatory power and its ability to provide a rational basis for therapeutic interventions (Kassirer 1989, pp. 896–897).

The third type of reasoning is deterministic or categorical reasoning. Deterministic reasoning uses predominantly compiled knowledge that may arise from probabilistic or causal associations between clinical findings. It requires the identification of rules that describe routine practices. The rules can have many purposes, such as describing therapeutic approaches or making prognoses. They might also recommend further diagnostic tests given certain already ascertained data. The rules are in the form of conditionals: If \( x \) obtains, then do \( y \); if \( x \) does not obtain, then do \( z \). They are represented by branching algorithms, ordered sets of instructions in a flow chart. The flow chart contains a diagram of graphic symbols for each act of reasoning. Two main types of “boxes,” often referred to as “nodes,” are used to indicate logical activities. A decision box contains a statement of a question to be answered; an execution box contains a statement of a procedure to be performed. A decision box is followed by a branching pathway of at least two possibilities; the reasoning pathway takes a direction indicated by the answer to the question. Arrows are used to indicate the exits and pathways leading from one decision or execution box to the next (Feinstein 1974, pp. 6–7). Typically, each nonterminal node requires unequivocal answers, which then serve as the matter of the branches leaving that node. The terminal nodes represent precise outcomes, answering to the questions the algorithm was designed to answer.

Algorithms are particularly useful in relatively straightforward cases where the logic of the diagnostic process can be precisely defined. Another advantage is that with a well-defined and explicit procedure, it is difficult to omit important questions or tests. Deterministic reasoning, however, depends on the quality of the data that serve as input and does not deal effectively with uncertainty. It may also yield bad answers if the algorithm is applied in a context sufficiently different from the one for which it was designed. Finally, the need to formulate all the rules necessary for even moderately complex diagnostic tasks is a challenge. In complex cases, the branching algorithm can become unwieldy (Kassirer 1989, pp. 897–898).

Each of these three approaches has benefits and limitations, but the limitations can sometimes be ameliorated by the concurrent use of other approaches. Hence, the three approaches are complementary. Probabilistic models can be useful for triggering hypotheses but are dependent on knowing the prevalence of disease in the population from which the patient comes. Causal models, on the other hand, are specific to disease entities and independent of the patient population. They are dependent on fundamental knowledge of physiological function and dysfunction. Once a hypothesis proposes a particular cause, causal reasoning is useful for verification of the cause and for explaining the observations. Causal reasoning can also identify circumstances in which the assumption of independence between diseases required by probabilistic models does not hold. When a knowledge base is built from these models, deterministic models may be constructed to aid in future diagnosis and even serve as bases for computer-assisted diagnosis (Kassirer 1989, p. 898).

**Diagnostic Goals and Context**

From a semantic standpoint, there are several different kinds of diagnoses. Nosological diagnosis purports to identify a disease or diseases from which a patient suffers. This, however, cannot completely describe diagnosis. It would require that every aspect of the diagnosis is the name of some disease and this does not reflect actual medical practice. Some diagnoses, for instance, are abnormality diagnoses; they include disease but also other disorders, injuries, wounds, lesions, defects, deformities, disabilities, etc. Other
Diagnoses are causal diagnoses; they give accounts or explanations of the data obtained in the diagnostic process. The category of diagnosis itself may not have clear boundaries, but may be “fuzzy” (Sadegh-Zadeh 2012, pp. 328–335).

The diagnostic context includes the patient, the physician, the physician’s practice, the hospital, the patient’s family, medical knowledge, and other factors; it produces a diagnosis as one of its outputs. Although a diagnosis is commonly purported to be a statement of some truth, it is perhaps better described as a performative utterance, a speech act, which generates truth and triggers individual, group, and even organizational behavior. A diagnosis imposes a social status on a person. It can exempt people from normal obligations, provide special financial compensation, and cause people who have committed crimes to be found non-culpable by reason of insanity. Thus, diagnosis is also essentially a social act. Diagnosis is, in this sense, a social construct (Sadegh-Zadeh 2012, pp. 335–339). While it may be the case that a diagnosis is constructed from facts, taken in a realist sense, the process of diagnosis depends on conceptual commitments and value judgments at every stage (Stempsey 2000). Diagnosing disease is important not only as a basis for effective treatment, but for a much broader array of medical and social reasons.

Definitions of Key Terms

Disease: (1) a set of biological phenomena that are said to be the cause of a person’s experience of illness, which is feeling unwell; (2) the class of all diseases or various subsets of that class; and (3) a single instance of (2)
Diagnosis: (1) the name of a disease that afflicts a person; (2) the process of determining (1)
Differential diagnosis: (1) a set of diagnostic hypotheses that fit the data obtained from the diagnostic process; (2) the process of formulating (1)
Nosology: classification of diseases
Probabilistic model of diagnostic reasoning: ascertaining the probability of a diagnosis given particular data using standard mathematical models of probability such as Bayes’s theorem
Causal model of diagnostic reasoning: the process of testing, verifying, and falsifying hypothetical cause-effect relationships that explain anatomical, physiological, and biochemical mechanisms of the body’s pathophysiological behavior in disease
Deterministic model of diagnostic reasoning: formulating rules that describe routine diagnostic practices based on compiled knowledge that may arise from probabilistic or causal associations. Rules are typically represented by branching algorithms in a flow chart containing decision points about possible ways to proceed given some determined answer

Summary Points

“Diagnosis” can refer to the name of a disease or the process of determining a disease present in an individual.
The concept of disease is philosophically controverted, and it influences judgments about diagnosis.
Diagnosis presupposes nosology, which can take many forms depending on the goal or goals judged to be most important.
The history of a patient’s illness, the physical examination, and various kinds of laboratory and clinical tests all provide data for diagnosis.
The process of diagnosis requires interpretation of all elements that disclose data and depends on conceptual and value commitments. Diagnostic reasoning includes various strategies and uses probabilistic, causal, and deterministic models in a complementary way. Diagnosis is essentially a social act and carries important social implications.

References

Bradley GW (1993) Disease, diagnosis and decisions. Wiley, Chichester
de la Croix B, de Sauvages F (1768) Nosologia methodica sistens morborum classes juxta Sydenhami mentem et botanicorum ordinem. Fratrum de Tournes, Amsterdam


Temkin O (1963) The scientific approach to disease: specific entity and individual sickness. In: Crombie AC (ed) Scientific change: historical studies in the intellectual, social and technical conditions for scientific discovery and technical invention, from antiquity to the present. Basic Books, New York, pp 629–647


