Differential Diagnosis and the Competing-Hypotheses Heuristic

A Practical Approach to Judgment Under Uncertainty and Bayesian Probability

Fredric M. Wolf, PhD; Larry D. Gruppen, MA; John E. Billi, MD

- Evaluating the same diagnostic information across the plausible competing diagnoses is a practical strategy (ie, heuristic) to guide decision making in the face of uncertainty. The prevalence of use of this competing-hypotheses heuristic by 89 first-year house officers was examined in three simulated patient cases. Results indicated that only a minority (24%) of the house officers selected optimal diagnostic information consistent with this Bayesian heuristic across all three cases. Almost all (97%) of the house officers selecting optimal diagnostic information were able to identify the most probable diagnosis specified by Bayes' theorem, while only a chance number (53%) of house officers selecting nonoptimal information were able to identify the most probable diagnosis. The competing-hypotheses heuristic is discussed within the context of diagnostic problem-solving models derived from the literature on medical decision making and clinicopathological conference case records. It is suggested that the heuristic, which does not necessitate any mathematical calculations, may be useful as a complement to clinical judgment.

(JAMA 1985;253:2858-2862)

THE MEDICAL community is becoming increasingly aware of the growing complexity involved in clinical decision making. The level of understanding and sophistication necessary to diagnose, treat, and manage medical problems increases concomitantly with the increase in new diagnostic and therapeutic procedures. Moreover, it is well recognized that there is a great deal of uncertainty inherent in clinical information, in the interpretation of laboratory data, in the relationship between clinical findings and disease, and in the effects of various therapies. It is also known that physicians, like most people, often do not manage uncertainty very well and may be prone to making certain errors that may affect the quality of clinical care.

Over the past decade, a number of studies have examined the strategies (ie, heuristics) and processes by which physicians solve medical problems. A number of these studies have taken a descriptive approach by trying to clarify the steps and methods that clinicians generally use in problem-solving situations. The picture of clinical problem solving that emerges from these studies is that of a physician with limited capacity for processing information who is trying to make sense of a problem situation typified by a large amount of both complexly interrelated information and uncertainty regarding the correct diagnosis. One model that describes the physician's problem-solving efforts has three components: (1) an early formulation of a small number (three to five) of hypotheses (disease candidates) based on clinical data initially available in conjunction with the physician's prior experience and knowledge, (2) subsequent pursuit of more clinical data to confirm or reject these hypotheses, and (3) selection of a final hypothesis after a critical level of confirmation has been reached.

In contrast to the descriptive approaches taken by some of the studies in medical problem solving, other researchers have focused on a prescriptive approach in an attempt to determine the correct or optimal method that ought to be used in making a particular clinical decision. Normal human decision making often is inconsistent with solutions derived...
from a prescriptive (or normative) approach. Such shortcomings in human judgment include a bias toward positive and confirming evidence, primary effects of initial information, premature closure on a hypothesis, inability to deal properly with probabilities, diagnostic conservatism, and others. These shortcomings have provided the impetus for the development of a wide range of possible corrective strategies.

One prominent prescriptive model is based on Bayes' theorem, which provides a rational, normative means of formulating a differential diagnosis and selecting a most probable diagnosis. Bayes' theorem has been applied to clinical practice and research in a variety of settings. These applications have included guidance on therapeutic decisions in patients with an uncertain diagnosis, assistance in analyzing the cost effectiveness and risk-benefit ratio for medical procedures, and insight into the meaning of such test performance characteristics as false-negative rate. Bayes' theorem can be used in its mathematical form when estimates of the likelihood of a disease for relevant groups of individuals and the sensitivity and specificity of diagnostic tests are available. The sensitivity of a diagnostic test is the ratio of the number of true-positive test results to the total number of patients with the disease [false-negatives plus true-positives]. Specificity of a test is the ratio of true-negative test results to patients without the disease [true-negatives plus false-positives.] While it often is burdensome to perform the mathematical calculations associated with Bayes' theorem, the development of computer programs to manage this task makes it increasingly possible to use the theorem as a supplement to clinical judgment. Even without performing the calculations, Bayes' theorem can aid in patient management by assisting the clinician in deciding the relevance of specific additional data. This can guide test selection, which is increasingly important in light of rising medical costs.

Notwithstanding these contributions, the use of Bayes' theorem as a normative guide for decision making has come under attack from a variety of directions. Some of the most frequent criticisms have focused on the large amount of specific information required for its use, the difficulty in generating hypotheses from signs and symptoms (since medical knowledge is taught primarily by disease), the documented inability of human beings to manipulate probabilities in their heads, and the unavailability of prior probability estimates (such as prevalence or incidence rates for relevant patient subgroups). Harris has pointed out the additional difficulty created by the paucity of available information concerning the sensitivities and specificities of commonly used laboratory tests, information that is needed to use Bayesian or any other decision-analysis approach. Bayes' theorem also has been criticized for its inability to conveniently allow for the hypothesis of multiple diseases in the same patient or for the nonindependence of symptoms within one disease. Although attempts to extend Bayes' theorem to avoid these oversimplifications are currently under investigation, they are not as yet ready for clinical use.

In spite of these criticisms, the underlying logic of Bayes' theorem is still valid and useful even when knowledge of the precise probabilities is not available. Interestingly, apparently simple and efficient heuristics suggested by Eddy and Clanton and Elstein et al. can be incorporated within a Bayesian framework. An understanding of the logic underlying the Bayesian framework per se may be more helpful in improving systematic diagnostic problem solving than the manipulation of probabilities. In order to explain this reasoning, it is necessary first to review briefly Bayes' theorem. In its simplest form with only two disease possibilities, the theorem may be written as follows:

\[
P(D_i|S) = \frac{P(S|D_i)P(D_i)}{ P(S)}
\]

The subscripts indicate whether we are examining the first or second disease. If \(P(D_i|S) > 0.50\), then the probability that the patient has disease 1 is greater than the probability of his having disease 2. If \(P(D_i|S) < 0.50\), then disease 2 is more likely. If \(P(D_i|S) = 0.50\), then both diseases are equally likely. While not in use in this way in the present study, disease 2 could be considered to be "not disease 1," that is, it could be any other disease or no disease at all. As can be readily seen, using Bayes' theorem becomes a complex task even in this simplest form with only two disease possibilities. Most medical cases are not limited by even this simplicity. While it is understandable that Eddy and Clanton, Harris, and others are skeptical of the practical value of a Bayesian approach to medical decision making, there is an inherent element of logical reasoning built into Bayes' theorem that may prove helpful even without the necessity of calculating any probability values. If we examine the denominator on the right side of formula 2, we see that any given symptom \(S\) in this illustration) must be examined across all the diseases being entertained in the differential diagnosis \(D_i\) and \(D_j\) in this illustration). Elstein et al. have defined the competing-hypotheses heuristic as the consideration of each piece of information (eg, symptom) with respect to all hypotheses under consideration before a diagnostic judgment is made. Wolf pointed out the relationship between the competing-hypotheses heuristic and Bayes' theorem, as the information neces-

Bayesian Heuristics—Wolf et al 2859
ecessary to use the competing-hypotheses heuristic is also the information required to solve the denominator in Bayes' theorem in formula 2 above. Individuals are unable to compare the diagnosticity of the datum for each disease if they do not select additional diagnostic information (data) pertaining to the likelihood of the first datum, given each of the other plausible diseases in the differential diagnosis. Other data selections, eg, additional symptoms pertaining to the first plausible disease category under consideration, cannot be used to solve Bayes' theorem to arrive at the correct probability of having the disease given the symptom. In the context of Eddy and Clanton's decision-making model, this heuristic could be used by evaluating the one or two pivotal findings of a patient's case across each disease in the "possible cause list." By evaluating each competing-disease hypothesis in light of the pivotal finding, the physician could make more effective use of the available information and would be less likely to close prematurely on a single diagnostic hypothesis. The physician also would be directed toward the systematic collection of more useful information, be it sign, symptom, or laboratory result.

In spite of the value of the competing-hypotheses heuristic, Kern and Doherty and Wolf have shown that most medical students prefer to select symptomatic information that is incompatible with the competing-hypotheses heuristic and the logic of Bayes' theorem. Instead of evaluating one symptom across multiple diseases, many medical students seek information about several symptoms for a single disease. The degree to which physicians employ the competing-hypotheses heuristic currently is not known and is addressed for the first time in this study, to the best of our knowledge. The implications of selecting diagnostic information consistent with the heuristic also is examined by analyzing the effects of optimal datum selection on diagnostic accuracy.

METHODS
Description of the Instrument
The instrument used in the present study was adapted from those developed by Kern and Doherty and modified by Wolf. The questions were designed to capture the simplest application of Bayes' theorem to medical diagnosis. Subjects were asked to decide which of two diagnoses was more likely, given information concerning one symptom. The optimal solution based on the competing-hypotheses heuristic is to consider the probability of each disease, given the same symptom, with the disease with the higher probability being the more likely diagnosis. In order to assess house officers' behavior in selecting the optimal information (datum), a 2X2 matrix yielding four pieces of information was provided for each of three separate simulated patient situations. An illustration of this type of problem is provided in the Figure. In each case, house officers were told that the patient had two symptoms, and they were initially provided with one datum regarding the percentage (ie, probability) of people with disease A that have a particular symptom. In the first situation, information was provided that 66% of those people who have contracted disease A have had a fever. Subjects were then given the choice of selecting either the percentage of patients with disease B that have the same symptom (fever, the optimal choice), or data regarding the prevalence of a different symptom (rash), in disease A or disease B. These were selected by removing one of the three opaque stickers that cover the alternative data choices. Subjects were not given the correct answers to any of the problems. In cases 1 and 2, subjects were informed that about an equal number of people suffer from disease A as from disease B (that is, the prior probabilities or base rates for diseases A and B are equal). In case 3, subjects were informed that about twice as many people suffer from disease A as from disease B (ie, unequal base rates). In all cases, the competing-hypotheses heuristic dictates that the same symptom must be evaluated in relation to the rival plausible disease(s) before the most probable diagnosis can be determined.

The Table summarizes the actual datum given for each case and the probability values available for selection by the subjects. In case 1, the symptom rash is not diagnostically helpful because it does not provide the information needed to solve the equation for Bayes' theorem. Moreover, the probability of having a rash is the same (62%) for both diseases A and B. The probability of having a fever and disease B (84%), however, is greater than the probability of having a fever and disease A (66%). Thus, the most likely diagnosis in this example is disease B. It is likely that subjects selecting information regarding a rash for either diseases A or B would not believe that disease B is the more probable of the two diseases; only those subjects selecting information regarding fever for disease B would be likely to correctly select it as the more probable disease. As can be seen in the Table, the location of the initial datum provided to subjects was varied for the three cases. This minimized the possibility of a response bias effect.

Subjects and Data Collection
The 89 subjects in the study were new first-year house officers at the University of Michigan Medical Center, Ann Arbor, who participated in an optional orientation course. The house officers represented a variety of fields of planned medical specialization and accurately reflected the distribution of medical specialties at the university. Approximately 92% of all the first-year house officers in the course completed the questionnaire.
RESULTS

Individual Cases

An analysis of the distribution of choices over the three possible pieces of data for each of the three problems indicated that these subjects are not making random choices on any one case ($\chi^2=14.03, 25.21, \text{and } 39.15$, respectively, all $P<.01$). Most choices for each case clustered either around the datum that, relative to the provided datum, corresponded to the same symptom and the alternative disease (the optimal choice) or around the datum that corresponded to the same disease and the alternative symptom. The subjects tended not to select the datum on the diagonal from the given information, that is, the datum for the alternative disease and the alternative symptom. Among the 1,554 optimal data, only 94% (92%) selected the optimal datum for the alternative disease and corresponding alternative disease. A nonspecific symptom frequently may be present in more than one disease. Failing to evaluate a symptom across all potential diagnostic hypotheses may result in a false-negative (or positive) error because the frequency of the symptom in a second disease actually may be greater (or lower) than in the first disease. This may be true even when the frequency of the symptom is extremely high (or low) in the first disease. The failure to request and evaluate symptomatic data across competing diseases has been termed "pseudodiagnostics." 24

Because medical education is presented in a disease-oriented manner, it is likely that much of actual clinical diagnostic reasoning is approached from a "within" hypothesis perspective in which the clinician determines the most likely diagnosis by how well the actual case fits a prototypical case for the disease under consideration. Clinicians who make diagnostic judgments solely based on the clusters of clinical findings associated with a particular disease most likely would not make comparisons consistent with the competing-hypotheses heuristic and Bayes' theorem. Interestingly, most house officers (62%) selected the optimal datum in only one (33%) or two (29%) of the three cases. This might suggest a case-specific pattern of reasoning rather than a preset strategy used across all cases. Asking subjects to explain why they select certain data and not others might provide insight into the mental strategies being used. The pattern of results for each of the three cases was in general similar even though different base rates and conditional probabilities (in the cells of the matrices) were provided to house officers before they made their datum selections. A more systematic study of the effect on datum selection of (1) differences in magnitude of the conditional probabilities provided in the matrix, (2) differences between

Relation Between Data Selection and Choice of Diagnosis

Across all three cases, approximately 97% of the subjects selecting optimum data also chose the most likely diagnosis according to Bayes' theorem; only 53% of the subjects selecting nonoptimal data chose the most likely diagnosis. Because the diagnostic choice was limited to only two diseases in each of these cases, disease A or B, it was expected that 50% of the subjects would select the most likely diagnosis on the basis of chance alone.

COMMENT

Only a minority (24%) of house officers selected optimal symptomatic data consistent with the competing-hypotheses heuristic and Bayes' theorem across all three patient cases. Evaluating symptoms in relation to only one disease does not by itself address the specificity of those symptoms for that disease. A nonspecific symptom frequently may be present in more than one disease. Failing to evaluate a symptom across all potential diagnostic hypotheses may result in a false-negative (or positive) error because the frequency of the symptom in a second disease actually may be greater (or lower) than in the first disease. This may be true even when the frequency of the symptom is extremely high (or low) in the first disease. The failure to request and evaluate symptomatic data across competing diseases has been termed "pseudodiagnostics." 24

Because medical education is presented in a disease-oriented manner, it is likely that much of actual clinical diagnostic reasoning is approached from a "within" hypothesis perspective in which the clinician determines the most likely diagnosis by how well the actual case fits a prototypical case for the disease under consideration. Clinicians who make diagnostic judgments solely based on the clusters of clinical findings associated with a particular disease most likely would not make comparisons consistent with the competing-hypotheses heuristic and Bayes' theorem. Interestingly, most house officers (62%) selected the optimal datum in only one (33%) or two (29%) of the three cases. This might suggest a case-specific pattern of reasoning rather than a preset strategy used across all cases. Asking subjects to explain why they select certain data and not others might provide insight into the mental strategies being used. The pattern of results for each of the three cases was in general similar even though different base rates and conditional probabilities (in the cells of the matrices) were provided to house officers before they made their datum selections. A more systematic study of the effect on datum selection of (1) differences in magnitude of the conditional probabilities provided in the matrix, (2) differences between

Bayesian Heuristics—Wolf et al 2861

Downloaded From:  by a NHS Royal Devon & Exeter Foundation Trust User on 11/22/2017
diagnostic problems of 50 clinicopathological conferences published in the New England Journal of Medicine. These steps included (1) aggregation of groups of findings into patterns, (2) selection of a “pivot” or key finding, (3) generation of a cause list, (4) pruning of the cause list, (5) selection of a diagnosis, and (6) validation of the diagnosis. In cases where they were unable to identify these elements, they were unable to identify any other pattern. While they pointed out that Bayes’ theorem theoretically should govern the formulation of a differential diagnosis and the most probable diagnosis, they recognized that a variety of “obstacles make it exceedingly unlikely that the reasoning process used by physicians to perform complicated diagnoses resembles Bayes’ theorem.” We suggest that clinical decision making might profit by evaluating the pivotal finding (Eddy and Clanton’s pivot heuristic) across all competing plausible diseases (Elstein and colleagues’ competing-hypotheses heuristic) to arrive at the most likely diagnosis. Furthermore, Eddy and Clanton suggested a practical approach in which the clinician performs this comparative evaluation by beginning with the two most probable diseases and proceeding in a sequential pairwise fashion down the list of other possible diagnoses. In this way, the competing diseases can be eliminated one at a time. Because Eddy and Clanton identified this strategy but did not give it a name, we suggest labeling this the “pairwise” heuristic.

The first and most useful step in the application of Bayes’ theorem to clinical situations is to learn to differentiate optimal (useful) from nonoptimal (useless or misleading) data. Selection of optimal data may in itself assist the clinician in arriving at the most probable diagnosis even without performing the mathematical calculations involved in the use of Bayes’ theorem. It is in this sense that probability and decision analysis may find a useful place in clinical reasoning. While this study and others have provided evidence of the less than optimal use of the competing-hypotheses heuristic in simulated cases, the degree to which practicing physicians select optimal information in real patient encounters remains as yet undetermined.

This investigation was supported in part by a grant from the Spencer Foundation to Dr. We have Steven A. Eraker and J. Frank Yates for their helpful comments on an earlier version of this article.

References


2862 JAMA, May 17, 1985—Vol 253, No. 19

Bayesian Heuristics—Wolf et al