INTRODUCTION

Diagnosis is the process, as is well-known, of determining the disease causing illness in a patient with symptoms. It is undoubtedly the most important step in taking care of a patient, as proper treatment and prognosis depend upon accurate diagnosis of a disease. We find practicing physicians to perform diagnosis on a daily basis and diagnose diseases with an overall remarkably high accuracy rate of 85 to 90 percent (1). But the reasoning employed in diagnosis by these physicians, we believe, is not entirely clear. In this paper, we shall study and analyze the process of diagnosis in practice of one particular disease, acute myocardial infarction (MI) to learn about the reasoning employed in diagnosis.

PROCESS OF DIAGNOSIS IN PRACTICE

We find that acute MI occurs, like any other disease, with varying presentations that range from being highly typical to being highly atypical, in different patients. For example, it may occur in a 65 year old man with multiple cardiac risk factors, presenting with highly characteristic chest pain (highly typical presentation) as well as in a healthy 40 year old woman without any cardiac risk factor presenting with highly uncharacteristic chest pain (highly atypical presentation) (2). By a highly typical presentation, we understand that most patients with this presentation would have acute MI while we understand by an atypical presentation that few patients with this presentation would have acute MI. As the goal in practice is to diagnose acute MI accurately in both these patients, an experienced physician suspects acute MI from the presentation in both these patients. As he does not know at this stage if acute MI is present or not in either patient, he formulates the suspected acute MI as a hypothesis in both patients (3,4). This hypothesis is an assumption that acute MI is present in these two patients.

The next step taken by the physician is to evaluate if this hypothesis is correct by performing a test, an EKG, in both patients. Let us suppose, he observes the highly informative test result, acute ST elevation EKG changes with likelihood ratio (LR) of 13 (5) in both patients. He interprets this test result as strong evidence, we suggest, based on its known performance of diagnosing acute MI accurately with the high frequency of 86 percent in a series of patients with varying presentations in whom acute MI is suspected from a presentation (6). Therefore, he diagnoses acute MI from acute ST elevation EKG changes with a high degree of confidence in the high accuracy of this diagnosis in both patients.

We find that any disease which has a test capable of generating a highly informative test result with LR greater than 10 (7) is diagnosed in practice in a similar manner regardless of whether it is suspected from a typical or an atypical presentation. Thus pulmonary embolism is diagnosed from positive chest CT angiogram, LR 20 (8); deep vein thrombosis from positive venous ultrasound study, LR 16 (9) and covid-19 disease from positive covid-19 antigen test, LR 14 (10) in a similar manner in practice. A disease which does not have a test capable of generating a highly informative result with LR greater than 10, is diagnosed in practice, we suggest, from a combination of two or three test results, whose combined LR is greater than 10, but this, we believe, needs to be studied further.

Some important features of the process of diagnosis in practice described above are as follows:

(1) The only role of a presentation, typical or atypical, is to make a physician suspect a disease and formulate it as a hypothesis. It plays no further role in diagnosis once a highly informative test result is observed.
(2) Evidence, from which a disease is diagnosed with a high degree of confidence, is provided only by a highly informative test result based on its performance. Thus this diagnosis is consistent with and validated by our experience.
(3) Therefore there is a clear division between the roles of a presentation and of a highly informative test result, the former making us formulate a hypothesis and the latter providing us evidence to evaluate the hypothesis as being correct.

DIAGNOSTIC REASONING IN PRACTICE IS NOT BAYESIAN

We note that the reasoning employed for diagnosis in practice described above is clearly not Bayesian for the following reasons:

(1) A prior probability based on a presentation is not attached to a hypothesis so that there is no prior degree of belief for or against the disease hypothesis.
(2) A disease is not diagnosed from a posterior probability generated by combining a prior probability and a LR.

Thus Bayesian reasoning is not employed for diagnosis in practice even though it has been prescribed (11) in theory. We note it has been prescribed due to its coherence defined in terms of not losing a bet placed on a Bayesian diagnosis (inference) (12) and not due to its diagnostic accuracy. It is not employed in practice, we suggest, because it is likely to lead to diagnostic errors as discussed below if it were to be employed:

(a) The Bayesian notion of interpreting a prior probability of a disease as prior degree of belief for or against it is likely to encourage failure to suspect or test a disease with a highly atypical presentation such as in the 40 year old woman above by interpreting its low prior probability as prior degree of belief against it. Thus it would encourage such errors which have been reported in several studies (13,14).
(b) A Bayesian diagnosis from a posterior probability may lead to a diagnostic error in a patient with a highly atypical presentation (very low prior probability) in whom a high informative test result with LR greater than 10 is observed. For example, in the 40 year old woman above, who was discussed in a clinical problem solving exercise (2), the Bayesian diagnosis would have been of acute MI being indeterminate from a posterior probability of 50 percent generated by combining a prior probability of 7 percent and LR of 13 (Appendix). But the discussing physician in this exercise did not make this Bayesian diagnosis, instead he diagnosed acute MI conclusively and accurately from acute ST elevation EKG changes alone.

DIAGNOSTIC REASONING IN PRACTICE IS FREQUENTIST

The reasoning employed for diagnosis in these two patients in practice, we propose, is frequentist due to presence of the two following key features of frequentist reasoning (15):

(1) A presentation, that is initially available information, is employed only to generate a hypothesis. A prior probability based on the presentation is not attached to this hypothesis.
(2) A disease hypothesis is evaluated and a disease diagnosed only from evidence provided by known performance of a highly informative test result in the form of a frequency.

We shall now discuss how a frequency, which represents performance of a highly informative test result, is generated in practice. In a series of patients with varying presentations and therefore with varying prior probabilities in whom we suspect acute MI over a period of time, we do not know in advance about prior probability in the next patient in whom we suspect acute MI. In addition, the prior probability in one patient is independent of prior probability in another patient in this series, therefore, we propose, the prior probability can be looked upon as being a random variable (16) and the series as being a random series. This series can also be looked upon, we suggest, as being a random sample that is drawn from a population of patients with varying prior probabilities in whom acute MI is suspected. We believe, any series of patients with varying prior probabilities in whom acute MI is suspected anywhere in the world can be looked upon as being a random sample which is drawn from this population.

In one such random sample, the frequency of acute MI in patients with acute ST elevation EKG changes has been observed to be 86+/-2 percent with confidence level 95 percent (6). This means this frequency will be between 84 and 86 percent in 95 percent random samples drawn from the population by the Central Limit Theorem (17). Therefore, when we observe acute ST elevation EKG changes in a patient suspected to have acute MI, we are 95 percent confident this patient has been drawn from a random sample in which the frequency of acute MI in presence of acute ST elevation EKG changes is between 84 and 86 percent. This enables us to diagnose acute MI, we suggest, with a high degree of confidence (95 percent) in the high accuracy (84 to 86 percent) of this diagnosis in this patient.

The above described frequentist reasoning is employed for diagnosis in practice for a number of reasons as follows:

(1) A very high degree of accuracy of 85 percent or greater is achieved in diagnosis of a disease in patients with varying prior probabilities that are invariably encountered in practice by frequentist reasoning. Thus this reasoning leads to an accurate diagnosis in 8 to 9 out of 10 patients and so helps achieve the primary goal in diagnosis of very high diagnostic accuracy.

(2) It is especially noteworthy that this reasoning leads to accurate diagnosis of a disease with an atypical presentation, which has a low prior probability as we see in published diagnostic exercises in real patients such as in clinical-pathologic conferences (CPCs) and clinical problem solving exercises (18,19) in which such diseases are routinely diagnosed accurately with frequentist reasoning. This is made possible primarily by the fact, we suggest, that a suspected disease with an atypical presentation is merely a hypothesis without any prior probability attached to it so that it does not have any prior degree of belief against it in frequentist reasoning.

(3) A diagnosis in a given patient, made by frequentist reasoning, is experience based as it is made on the basis of an observed frequency and is thus highly reliable.

(4) A diagnosis of a disease, for example, of acute MI from acute ST elevation EKG changes in a patient in whom it is suspected is made by frequentist reasoning with the same high degree of confidence in the high accuracy of this diagnosis all over the world, whether it is in USA (20) or in Europe (20) or in India (21) or in Africa (22). This uniformity in diagnosis, we suggest, is due to the random nature everywhere of the series of patients with varying prior probabilities in whom acute MI is suspected.

DIFFERENCES BETWEEN BAYESIAN AND FREQUENTIST REASONING

The basic difference between Bayesian and frequentist reasoning is that the former employs a logic of confirmation while the latter uses a logic of testing (23). In Bayesian reasoning, a posterior probability represents total degree of confirmation generated by combining prior probability (prior degree of confirmation) and LR (evidence due to test result). A disease is diagnosed in Bayesian reasoning from a high posterior probability which represents strong degree of total confirmation, but this approach causes problems in diagnosis in practice. For example, in the above 65 year old man with very high prior probability of acute MI, let us suppose, non-specific T wave EKG changes with LR of 1 are observed when an EKG is performed to evaluate the hypothesis of acute MI. A high posterior probability would be generated by combining the high prior probability and LR of 1, which would represent strong degree of total confirmation from which acute MI would be diagnosed to be present with a high degree of certainty in Bayesian reasoning. But this Bayesian diagnosis would not be made in practice, we believe, because the test result, non-specific T wave EKG changes with LR of 1, is known to be worthless in frequentist reasoning as we discuss below.

In frequentist reasoning, a disease hypothesis is diagnosed to be correct only if it passes a stringent test by observing a highly informative test result. For example, the disease hypothesis of acute MI in a patient is considered correct only if the highly informative test result, acute ST elevation EKG changes with LR of 13 is observed. This test result is considered highly informative based on its performance of diagnosing acute MI accurately with the high frequency of 86 percent (8 to 9 out of 10) in patients with varying prior probabilities. The test result, non-specific T wave EKG changes with LR 1, on the other hand, is considered worthless, based on its performance of diagnosing acute MI accurately, perhaps, with a frequency of around 50 percent in patients with varying prior probabilities.

Another important difference between Bayesian and frequentist reasoning in diagnosis relates to the role of probability in these two approaches. In the former, a probability is attached to a disease and is interpreted as a subjective degree of belief in it while in the latter it is attached to a procedure and is interpreted as a frequency with which the procedure leads to an accurate diagnosis (23). The goal in diagnosis in practice to achieve highly reliable accurate diagnosis of a disease, based on experience, in patients with varying prior probabilities, is achieved, as we have discussed in this paper only by frequentist reasoning in which a probability is attached to a procedure.

NO EVIDENCE FOR BAYESIAN DIAGNOSTIC REASONING IN PRACTICE

We do not find any published study or report in which the prescribed Bayesian reasoning has been employed for diagnosis in a real patient. Our experience is similar to that of the eminent clinical investigator, Alvan Feinstein (24), who noted, "I know of no clinical setting or institution in which the Bayesian diagnostic methods are being regularly used for practical diagnostic purposes in a routine or specialized manner". Wherever we look in practice, for example, in published diagnostic exercises in real patients such as in CPCs and in clinical problem solving exercises (18,19), we do not find Bayesian reasoning to be employed for diagnosis.

CONCLUSION

In conclusion, we have argued in this paper that diagnostic reasoning in practice consists of suspecting a disease from a presentation and formulating it as a hypothesis without attaching any prior probability to it. The disease hypothesis is evaluated by performing a test and inferred to be correct if a highly informative test result with LR greater than 10 is observed, based on the performance of this test result in diagnosing the disease accurately with a very high frequency.. This reasoning is frequentist in nature and leads to highly accurate diagnosis of a disease in patients with varying presentations (prior probabilities).

We would like to point out that frequentist diagnostic reasoning in practice has been developed by experienced physicians on their own to meet the challenge of diagnosing a disease accurately with a high degree of reliability in patients with varying presentations (prior probabilities). The fact that a series of patients with varying prior probabilities in which a disease is suspected is a random sample makes frequentist diagnostic reasoning about this disease highly accurate wherever this disease is suspected. It is important, we believe, to recognize that it is frequentist reasoning and not the prescribed Bayesian reasoning which is employed for diagnosis in practice as this has important implications, we believe, for teaching diagnosis to medical students and for reducing diagnostic errors.

Appendix

Prior probability of 7 percent = Prior odds of 7/93

In odds form of Bayes' theorem,

Prior odds x Likelihood ratio = Posterior odds

Therefore,

7/93 x 13 = 1/1 = Posterior probability of 50 percent

References

1. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med 2008; 121: S2-S23.
2. Pauker SG, Kopelman RI. How sure is sure enough? N Engl J Med 1992; 326: 688-691.
3. Elstein AS, Shulman LS, Sprafka SA. Medical Problem Solving: An analysis of Clinical Reasoning. Cambridge MA: Harvard University Press 1978.
4. Kassirer JP, Gorry GA. Clinical Problem Solving: a behavioral analysis. Ann Intern Med 1978; 89(2): 245-255.
5. Rude RE, Poole WK, Muller JE, Turi Z, Rutherford J, Parker C et al. Electrocardiographic and clinical criteria for recognition of acute myocardial infarction based on an analysis of 3,697 patients. Am J Card 1983; 52: 936-942.
6. Larson DM, Menssen KM, Sharkey SW, Duval S, Schwartz RS, Harris J et al. "False positive" cardiac catheterization laboratory activation among patients with suspected ST-segment elevation myocardial infarction. JAMA 2007; 298: 2754-60.
7. Guyatt G, Rennie D, Meade MO, Cook DJ. Users' guide to the medical literature: A manual for evidence-based clinical practice. New York: The McGraw-Hill Companies 2008 p 428.
8. Stein PD, Fowler SE, Goodman LR, Gottschalk A, Hales C, Hull RD et al. Multidetector computed tomography for pulmonary embolism. N Engl J Med 2006; 353: 2317-27.
9. Zierler BK. Ultrasonography and diagnosis of venous thromboembolism Circulation 2004; 109: 1-9-1-14.

10. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. BMJ 2020.369M1808DO110.1136/bmj.m1808 (Published online on 12 Mat.
11. Weinstein MC, Fineberg HV. Clinical Decision Analysis. Philadelphia: WB Saunders Company 1980.
12. Edwards W, Likndman H, Savage LJ. Bayesian statistical inference for psychological research. Psych Review 1963; 70: 193-242.
13. Ely JW, Kaldjian LC, D'Alessandro DM. Diagnostic errors in primary care: Lessons learnt. J Am Board Fam Med 2012; 25: 85-97.
14. Singh H, Giardina TD, Meyer AN, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings, JAMA Intern Med 2013; 173: 418-25.
15. Cox DR. Principles of Statistical Inference. Cambridge UK: Cambridge University Press 2006.
16. Blitzstein JK, Hwang J. Introduction to Probability. London : Chapman and Hall 2019 p 103.
17. Adams WJ. The life and times of the Central Limit Theorem. Providence RI: American Mathematical Society 2009.
18. Jain BP. An investigation into method of diagnosis in clinical-pathologic conferences (CPCs). Diagnosis 2016; 3: 61-64.
19. Jain BP. Why is diagnosis not ptobabilistic in clinical-pathologic conferences (CPCs): Point. Diagnosis 2016; 3: 95-97.
20. Myocardial infarction redefined-a consensus document of the Joint European Society of Cardiology/American College of Cardiology Committee for the Redefinition of Myocardial Infarction. Eur Heart J. 2000; 21(18): 1502-1513.
21. Guha S, Sethi R, Saumitra R, Behl VK, Shanmugasundaram S, Kerker P et al.Cardiological Society of India: Position statement for the management of ST elevation myocardial infarction in India. Indian Heart J 2017 April 69 (Suppl 1) S63-S97.
22. Shavadia J, Yarga G, Otieno H. A prospective review of acute coronary syndromes in an urban hospital in sub-Saharan Africa. Cardiovasc J of Africa 2012; (6): 318-21.

23. Mayo DG. Statistical Inference as Severe Testing: How to get beyond\d the Statistics Wars. Cambridge UK. Cambridge university Press 2018.
24. Feinstein AR. The haze of Bayes': the aerial palaces of decision analysis and the computerized Quija board. Clin Pharmacol Ther 1977; 21: 482-495.

;