The Bayesian method has been prescribed as the normatively correct method of diagnosis (1,2,3) but as we point out in this paper, it does not appear to be employed for diagnosis in practice. We shall argue in this paper that this discrepancy between prescription and practice arises from a profound difference between the reason for which the Bayesian method has been prescribed and the stated goal of diagnosis in practice.

In the Bayesian method (1), the prior probability of a disease which is suspected from a presentation is estimated from its prevalence in a population and interpreted as a prior degree of belief for or against it in a patient. It is then combined with the likelihood ratio (LR) of a test result to generate a posterior probability, which is interpreted as a total degree of belief from which the disease is inferred in the patient.

The Bayesian method appears to have been prescribed for diagnosis in early 1960s (4) on grounds of its rationality due to its coherence in inference based on a betting argument. In this argument, a bet is placed on a Bayesian inference with odds based on the posterior probability from which the inference is made. This bet and therefore the Bayesian inference is considered coherent as it prevents a Dutch book, which is a series of bets which ensures loss, from being created against the inferring agent. What is of interest in this argument from the point of view of a practicing physician is that there is no mention of inferential (diagnostic) accuracy in this prescription. The comments of the well-known Bayesian statistician, Dennis Lindley (5), in this context are of interest. He writes, “I am often asked if the (Bayesian) method gives the right answer, or more particularly, how do you know if you have the right prior. My reply is that I don’t know what is meant by ‘right’ in this context. The Bayesian theory is about coherence, not about right or wrong”. We shall see Lindley’s comments offer an important clue to why the Bayesian method is prescribed, but not employed for diagnosis in practice.

The only goal in diagnosis in practice, as is well known, is to determine a disease accurately in a patient in a patient with symptoms, so that proper treatment and prognosis can be given for it. The achievement of this goal is made challenging by the well- known fact that practically every disease occurs with varying presentations and therefore with varying prior probabilities in different patients. For example, acute myocardial infarction (MI) is known to occur in a 65 year old man with multiple cardiac risk factors presenting with highly characteristic chest pain, in whom its prior probability is very high, say at 86 percent. And it is known to occur as well in a healthy 40 year old woman without any cardiac risk factor presenting with highly uncharacteristic chest pain in whom its prior probability is very low at 7 percent (6).The challenge in practice is to diagnose acute MI accurately in both these patients. We shall now examine how this challenge is met by first examining how acute MI is diagnosed in practice in the 40 year old woman, who is a real patient who was presented and discussed in a clinical problem solving exercise.

Acute MI is suspected from the presentation and an EKG performed to evaluate it, which reveals acute ST elevation EKG changes with LR of 13 (7) in this patient. The discussing physician in this exercise diagnoses acute MI conclusively and accurately from this test result alone for as he says, "the ECG is quite specific for coronary ischemia and should outweigh the absence of risk factors, including older age " (6). We suggest, he is basing this diagnosis on the known high frequency of accurate diagnosis of about 86 percent from acute ST elevation EKG changes on repeated testing in other patients (8). If an EKG reveals acute ST elevation EKG changes in the 65 year old man mentioned above, we believe, acute MI would be diagnosed in this patient as well from this test result alone with a similar high degree of accuracy. We note that in the method employed for diagnosis in these two patients, the prior probability of acute MI does not play any role indicating this method is not Bayesian.

We shall now point out the problems that arise if the Bayesian method were to be employed for diagnosis of acute MI in these two patients. First of all, the very low prior probability of acute MI of 7 percent in the 40 year old woman would be interpreted as very strong prior degree of belief against acute MI which poses the risk of not suspecting or testing acute MI, leading to a diagnostic error as this patient actually had acute MI. Even if an EKG is performed and acute ST elevation EKG changes observed in this patient, the prior probability of acute MI of 7 percent would be combined with the LR of 13 for acute ST elevation EKG changes to generate a posterior probability of 50 percent (Appendix 1), which would be interpreted as total degree of belief from which acute MI would be inferred to be indeterminate in this patient. This diagnosis (inference) is coherent as it is equivalent to a bet with odds of 1 to 1, based on the posterior probability of 50 percent, but this coherence does not help us in any way in managing this patient. What we wish to know is about the accuracy of this diagnosis, which is unknown and therefore this Bayesian diagnosis is not made in this exercise.

Let us suppose non-specific T wave EKG changes with LR of 1 are observed when an EKG is performed to evaluate acute MI which is suspected from the presentation in the 65 year old man mentioned above. In the Bayesian method, the prior probability of acute MI of 86 percent in this patient would be combined with the LR of 1 of non-specific T wave EKG changes to generate a posterior probability of 86 percent (Appendix 2), which would be interpreted as very strong total degree of belief from which acute MI would be diagnosed to be present with a high degree of certainty in this patient. We doubt if this Bayesian diagnosis would ever be made in practice, as the test result, non-specific T wave EKG changes with LR of 1 is known to be totally non-informative and thus worthless in diagnosing anything.

The method employed for diagnosis of acute MI in practice in these two patients and in any other patients in whom we suspect acute MI, we propose, is to formulate acute MI as a diagnostic hypothesis without any prior probability attached to it so that it does not have any prior degree of belief for or against it in a patient. This hypothesis is evaluated by performing a test, an EKG, and inferred to be correct if the highly informative test result, acute ST elevation EKG changes with LR 13, is observed. This diagnosis is made in all patients regardless of prior probability of acute MI and is highly accurate with a frequency of accurate diagnosis of around 86 percent. This method of diagnosis in practice, we propose, is the frequentist, confidence method (9), which is the other (other than the Bayesian) main method of statistical inference.

We shall now describe the confidence method, which is employed for diagnosis in practice, with reference to diagnosis of acute MI. We suggest, the application of this method in practice is made possible by the fact that the prior probability of acute MI in a patient in whom we suspect it is a random variable (10) as it is independent of prior probability in another patient in whom we suspect it and we do not know in advance about its prior probability in the next patient in whom we suspect it. Therefore the series of patients with varying prior probabilities in whom we suspect acute MI can be looked upon as being a random sample which is drawn from a population of patients in whom we suspect acute MI. In one such series as random sample (8), the frequency of acute MI in presence of acute ST elevation EKG changes has been observed to be 86+/-2 percent with confidence level of 95 percent. This means this frequency is 86+/-2 percent with confidence level 95 percent in the population from which this random sample has been drawn. Therefore in all other random samples drawn from this population, the frequency of acute MI in presence of acute ST elevation EKG changes will be between 84 and 88 percent in 95 percent of these samples by the Central Limit Theorem (11). This means that if we observe acute ST elevation EKG changes in a patient in whom we suspect acute MI, we are 95 percent confident this patient belongs to a sample in which this frequency is between 84 and 88 percent which enables us to infer acute MI with a high degree of confidence (95 percent) that this inference is highly accurate (84 to 88 percent) (9). The limits, 84 and 88 percent, of the confidence interval, function, so to speak, as Cox (12) points out, as a measuring technique in this patient, which is calibrated like other measuring instruments, indirectly by the hypothetical consequences of its repeated use.

The confidence method is employed for diagnosis in practice due to its many features which promote diagnostic accuracy as follows:

(a) A confidence inference (diagnosis) in a given patient is highly reliable as it is based on an observed high frequency of this inference being accurate on repeated testing in other patients. For example, the inference of acute MI from acute ST elevation EKG changes in a given patient, such as in the 40 year old woman is highly reliable as it is based on this inference being accurate in 84 to 86 percent patients. This enables us to start treating this patient right away, for example, by transferring her to a cardiac catheterization laboratory. A Bayesian inference, on the other hand, is often unreliable as its accuracy is unknown. For example, the Bayesian diagnosis of acute MI being indeterminate from the posterior probability of 50 percent in the 40 year old woman is unreliable as the accuracy of this diagnosis is unknown. It is due to this unreliability, this Bayesian diagnosis is not made in this patient as we mentioned above.

(b) An extremely important feature of the confidence method is that a disease suspected from a presentation is formulated as a diagnostic hypothesis without any prior probability attached to it so that it does not have any prior degree of belief for or against it. This enables every suspected disease in a differential diagnosis to be tested and diagnosed accurately if it is present regardless of its prior probability. This is seen in all published diagnostic exercises in real patients such as in clinical-pathologic conferences (CPCs) and clinical problem solving exercises (13,14), in which every disease in a differential diagnosis is formulated as a diagnostic hypothesis only. This is the reason, we believe, a disease with a low prior probability, such as one which is rare or has an atypical presentation is diagnosed accurately with relative ease in these exercises. In the Bayesian method, on the other hand, the interpretation of low prior probability of a disease as prior degree of belief against it encourages diagnostic errors, which have been reported in several studies (15,16), due to failure to suspect a disease with an atypical presentation.

(c) As a confidence inference of acute MI from acute ST elevation EKG changes is made solely from the performance of this test result, this inference is made with a high degree of confidence in the high accuracy of this inference in any patient regardless of its prior probability, as we saw above in the 40 year old woman (low prior probability) and in the 65 year old man (high prior probability). This simplifies diagnosis of acute MI immensely. By contrast, a Bayesian inference is made from a posterior probability which may be low due to a low prior probability even if the test result has a high LR or may be high due to a high prior probability even if the LR is 1 as we saw earlier. This introduces ambiguity and confusion in diagnosis of acute MI, we believe, which may lead to diagnostic errors.

(d) The series of patients with varying prior probabilities in whom acute MI is suspected is a random sample everywhere in the world. Therefore acute MI is inferred (diagnosed) with a high level of confidence (95 percent) in the high accuracy (around mid-eighties) from acute ST elevation EKG changes everywhere, whether in USA (17) or Europe (17) or India (18) or Kenya (19). This inference does not differ from country to country due to varying prevalence of acute MI in these regions as we would expect in the Bayesian method.

In addition to confidence diagnosis of acute MI as discussed above, any disease which has a test capable of generating a result with LR greater than 10 (20) is inferred (diagnosed) in practice in a similar manner by the confidence method. For example, pulmonary embolism is inferred from positive chest CT angiogram, LR 20 (21), deep vein thrombosis from positive venous ultrasound study, LR 16 (22) and covid-19 disease from positive covid-19 PCR test (23) in any patient regardless of prior probability of disease. The great impact of availability of a test, capable of generating a result with LR greater than 10 on diagnostic accuracy is seen most clearly in the case of pulmonary embolism. The diagnostic accuracy of this disease increased dramatically, as is well known, with the development and availability of perfusion lung scan and chest CT angiogram which can generate such results. A disease which does not have such a test yet is inferred in practice, we believe, from a combination of two or three test results whose combined LR is greater than 10, but this needs to be investigated further.

We note from the above discussion that it is the hypothesis testing confidence method which is employed for diagnosis in practice instead of the Bayesian method in which a prior degree of belief in a disease, represented by its prior probability is updated to a total degree of belief by a test result. The reason this is done, we suggest, is due to our goal in diagnosis and the circumstances in which this goal is to be achieved.

The goal in diagnosis, as we mentioned earlier is to diagnose a disease accurately in every patient. The real life circumstance which makes the achievement of this goal challenging in practice is that practically every disease occurs in different patients with different prior probabilities that vary over a wide range. This challenge is overcome in the confidence method by formulating a suspected disease as a hypothesis only without any prior probability attached to it so that there is no prior degree of belief for or against it in a patient regardless of its prior probability. The function of a presentation in the confidence method is not to provide any prior evidence but only to provide information from which we can suspect a disease and formulate it as a hypothesis. The prior probability of a disease in the confidence method represents only chance of a disease in a patient and its only role in diagnosis in practice, we suggest is in prioritizing testing of various diseases in a differential diagnosis in a non-urgent diagnostic situation. The disease with the highest prior probability is tested first as it has the greatest chance of being present.

In the confidence method, a disease is inferred (diagnosed) in a given patient by a highly reliable procedure, which is known to infer the disease with a high degree of accuracy. The overall diagnostic accuracy of the confidence method in practice is 85 to 90 percent (24), while it is 98 percent in CPCs (13). This procedure consists of inferring it from a highly informative test result with LR greater than 10, which has a high frequency of around 85 percent or greater of accurate inference of the disease on repeated testing in patients. This process is facilitated by the fact that the series of patients with varying prior probabilities in whom we suspect a given disease is a random sample. We illustrated the confidence method of diagnosis in practice with the example of diagnosis of acute MI and enumerated its many features above which make it highly desirable for diagnosis in practice.

The Bayesian method, on the other hand, has been prescribed on grounds of its rationality based on its coherence. The greatest shortcoming of this prescription, in our view, is that the Bayesian method does not lead to diagnostic accuracy which is the only goal in diagnosis. In fact, this method often leads to diagnostic errors, as we have discussed above. We cannot think of a single reason from the point of view of diagnostic accuracy for employing the Bayesian method for diagnosis in practice, therefore it is not surprising it is not employed in practice. In our view, the situation regarding the Bayesian method in diagnosis in practice has been summarized best by the eminent clinical investigator, Alvan Feinstein (25) who writes: "I know of no published work in which the initial claims of a Bayesian enthusiast have been confirmed by the results found in clinical reality. 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. I know of no specific constructive practical diagnostic decisions-involving real patients, data and doctors-in which the Bayesian methods have made a prominent contribution that could not have been achieved as easily without Bayes' formula. (If readers know of any, I hope they will tell me)".

We have argued in this paper that the correct method of diagnosis from the point of view of diagnostic accuracy is the confidence method which is already employed in practice and not the prescribed Bayesian method. This argument has important implications for teaching diagnosis to medical students. For example, it removes the current inconsistency in teaching creation of a broad differential diagnosis which includes diseases with low prior probabilities and teaching the Bayesian method in which a low prior probability is interpreted as prior degree of belief against a disease. In the confidence method, a suspected disease is merely a hypothesis, as we have discussed above, without any prior probability attached to it, so that it does not have any prior degree of belief for or against it. In teaching the Bayesian method, emphasis is placed on accurate estimation of a prior probability as it plays a key role in diagnosis of a disease in this method. This emphasis is not required in teaching the confidence method, as a prior probability plays no role in diagnosis apart from prioritizing testing of various diseases in a differential diagnosis.

Appendix 1

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

In odds form of Bayes' theorem,

Prior odds x Likelihood ratio = Posterior odds

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

Appendix 2

Prior probability of 86 percent = Prior odds of 86/14

In odds form of Bayes' theorem,

Prior odds x Likelihood ratio = Posterior odds

Thus 86/14 x 1 = 86/14 = Posterior probability of 86 percent

## References


1. Weinstein MC, Fineberg HV. Clinical Decision Analysis. Philadelphia: WB Saunders, 1980.
2. Kassirer JP, Wong JB, Kopelman RI. Learning Clinical Reasoning. Baltimore: Lippincott Williams & Wilkins Health, 2010.
3. Sox HC, Higgins MC, Owens DK. Medical Decision-making. 2nd ed. West Sussex, UK: Wiley-Blackwell, 2013.
4. Edwards W, Lindman H, Savage LJ. Bayesian statistical inference for psychological research. Psych Review 1963; 70: 193-242.
5. Lindley D. Bayesian statistics in Harper W and Hooker C (Eds). Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science. Vol 2, Boston MA: D Reidel: 1976 pp 353-62.
6. Pauker SG, Kopelman RI. How sure is sure enough? N Engl J Med 1992; 326: 688-91.
7. 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 analysis of 3,697 patients. Am J Card 1983; 52: 936-42.
8. 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-elevation myocardial infarction. JAMA 2007; 298: 2754-60.
9. Mayo DG. Statistical Inference as Severe Testing: How to get beyond the Statistics Wars: Cambridge UK: Cambridge University Press 2018 pp 190-201.
10. Wikipedia. Random variable.
11. Adams WJ. The life and times of the Central Limit theorem. Providence RI: American Mathematical Society 2009.
12. Cox DR. Principles of Statistical Inference. Cambridge UK: Cambridge University Press 2006 p 197.
13. Jain BP. An investigation into method of diagnosis in clinical-pathologic conferences (CPCs). Diagnosis 2016; 3: 61-64.
14. Jain BP. Why is diagnosis not probabilistic in clinical-pathologic conferences (CPCs): Point. Diagnosis 2016; 3: 95-97.
15. Ely JW, Kaldjian LC, D'Alessandro DM. Diagnostic errors in primary care: Lessons learnt. J Am Board Fam Med 2012; 25: 85-97.
16. 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.
17. 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): 1502-13.
18. Guha S, Sethi R, Saumitra R, Behl VK, Shanmugasundaram S, Kerkar P et al. Cardiological Society of India: Position statement for the management of ST elevation myocardial infarction in India. India Heart J 2017 April 69 (Suppl 1): S63-S97.
19. 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; 23(6): 308-21.
20. 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.

21. Stein PD, Fowler SE, Goodman LR, Gottschalk A, Hales C, Hull RD et al. Multi-detector computed tomography for pulmonary embolism. N Engl J Med 2006; 354: 2317-27.
22. Zierler BK. Ultrasonography and diagnosis of venous thromboembolism. N Engl J Med. Circulation 2004; 109: 1-9-1-14.
23. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 result. BMJ 2020. 369m1808doi10.1136/bmj.m1808 (Published online on 12 May 2020).
24. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med 2008; 121: S2-S23.
25. Feinstein AR. The haze of Bayes', the aerial palaces of decision analysis and the computerized Ouija board. Clin Pharmacol Ther 1977; 21: 482-95.