

# Common things are common, but what is common? A foundation for probabilistic diagnosis.

Scott Aberegg<sup>1</sup> and Sean Callahan<sup>1</sup>

<sup>1</sup>University of Utah Health

July 7, 2021

## Abstract

The well-known clinical axiom stating that “common things are common” attests to the pivotal role of probability in diagnosis. Despite the popularity of this and related axioms, there is no operationalized definition of a common disease, and no practicable way of incorporating actual disease frequencies into differential diagnosis. In this expository essay, we aim to reduce the ambiguity surrounding the definition of a common (or rare) disease and show that incidence – not prevalence – is the proper metric of disease frequency for diagnosis. We explore how a numerical estimates of disease frequencies based on incidence can be incorporated into differential diagnosis as well as the inherent limitations of this method. These concepts have important implications for diagnostic decision making and medical education, and hold promise as a method to improve diagnostic accuracy.

## Common things are common, but what is common? A foundation for probabilistic diagnosis.

Scott K Aberegg, MD, MPH

Sean J Callahan, MD

**Funding:** There is no funding for this work.

**Ethical approval:** There is no requirement for ethical approval of this work.

Summary:

The well-known clinical axiom declaring that “common things are common” attests to the pivotal role of probability in diagnosis. Despite the popularity of this and related axioms, there is no operationalized definition of a common disease, and no practicable way of incorporating actual disease frequencies into differential diagnosis. In this expository essay, we aim to reduce the ambiguity surrounding the definition of a common (or rare) disease and show that incidence – not prevalence – is the proper metric of disease frequency for diagnosis. We explore how numerical estimates of disease frequencies based on incidence can be incorporated into differential diagnosis as well as the inherent limitations of this method. These concepts have important implications for diagnostic decision making and medical education, and hold promise as a method to improve diagnostic accuracy.

*“Prevalence is to the diagnostic process as gravity is to the solar system – it has the power of a physical law.”*  
- Clifton K. Meador, *A Little Book of Doctors’ Rules*

Diagnosis is the apotheosis of medical skill<sup>1</sup>. Without an accurate diagnosis, the patients’ foremost questions cannot be answered: what is wrong, what is likely to happen, and what can be done about it? Central to the diagnostic process is the compilation of a list of possible causes of the clinical presentation – a differential diagnosis. Traditionally, diseases comprising the differential diagnosis were rank ordered from highest to lowest in terms of likelihood<sup>2,3</sup>. This method, which is inconsistently used, is loosely probabilistic but can obscure large differences in probability between components of the differential; an ideal method would assign

probabilities to each possibility that sum to 100%<sup>4</sup>. Practical difficulties limit the ability to accurately estimate the probabilities of components of the differential<sup>5-7</sup>, yet probabilistic reasoning is nonetheless an essential part of expert forecasting and diagnosis<sup>8-16</sup>.

Several well-known clinical axioms pay homage to the primacy of probability in diagnosis. The most popular, “common things are common” (CTC), has been part of medical folklore for more than a century<sup>1,17</sup>. It is often expressed as a metaphor: “when you hear hoofbeats, look for horses not zebras”<sup>18-20</sup>. The CTC axiom is rooted in base rates and Bayes’ theorem and is the foundation of other epigrams such as “uncommon presentations of common diseases are more common than common presentations of uncommon diseases.” The CTC axiom and its variants arose to combat what are now known as cognitive biases such as base rate neglect<sup>21,22</sup> and the representativeness heuristic<sup>23,24</sup> but are themselves heuristics – rudimentary rules of thumb that provide general rather than specific guidance for the consideration of probability in diagnosis<sup>1</sup>.

Ironically, the axiom and extant literature are silent on how to determine what is common, what is rare, and related questions. Is sarcoidosis common? Is commonness related to epidemiological metrics such incidence and prevalence, or is it an intuitive or experiential determination, or both? Should diseases be dichotomized as either common or rare, or rated on a frequency continuum? Can the notion of commonness be operationalized as an aid to estimating probabilities of components of a differential diagnosis in a practicable way?

At first blush, discriminating between common and rare diseases is a simple task: pneumonia is common, and lymphangiomyomatosis (LAM) is rare; a physician will encounter many cases of the former and few (if any) of the latter in any given time interval. This self-evident truth is a distillation of the physician’s experience into the coarse dichotomy of common and rare; it obfuscates the magnitude of the difference in the frequency of the two diseases. How much more common is pneumonia than LAM? The dichotomy does not permit an answer to this question. Ideally, we would like to know the relative likelihoods of diseases so that, *ceteris paribus*, their weight in the differential diagnosis could be made proportional to their observed frequencies.

Fortunately, dichotomization is unnecessary. The actual frequencies of diseases can be used as the metric for comparison of their commonness. First, it is necessary to determine which measure of epidemiological disease frequency - incidence or prevalence - should be used. Incidence is the number of *new* cases diagnosed per person per year, whereas prevalence is the number of *existing* cases already diagnosed per person at a given time point. (Incidence is customarily expressed as cases/100,000 person-years, and prevalence as cases per 100,000 persons; prevalence is, roughly, the product of incidence and the average duration of the disease.) Incidence and prevalence are similar when a disease has a high recovery or short-term mortality rate (e.g., pneumonia), since death or recovery removes cases from the numerator. Prevalence is higher than incidence when the short-term mortality and recovery rates are low (e.g., emphysema) because chronic cases accumulate in the population, growing the numerator. Because incidence relates to *new* or *previously undiagnosed* cases and prevalence to *existing* or *already diagnosed* cases, it is incidence that is germane to the diagnostician. Therefore, the epigraph from *A Little Book of Doctors’ Rules* requires modification – incidence, not prevalence, has the power of a physical law<sup>25</sup>.

This distinction, until now neglected in the vast literature on diagnosis and clinical reasoning, is paramount because the prevalence of many diseases is higher than their incidence, sometimes markedly so. The probability that a patient seen tomorrow will present with previously undiagnosed symptomatic hypothyroidism is related to the incidence of hypothyroidism and most practitioners will go a month or longer without diagnosing a new case of (incident) hypothyroidism. By contrast, the probability that a patient with established (prevalent) hypothyroidism will be seen on an average day is high. The physician does not “diagnose” these cases of prevalent disease – they have already been diagnosed. For already diagnosed chronic diseases prone to complications or flare-ups, such as systemic lupus erythematosus (SLE), the prevalence of lupus will affect the probability of seeing the complications, as the latter are conditional upon the former<sup>26</sup>. However, when diagnosing complications of SLE, it is already known that SLE is present; therefore, the prevalence of SLE is immaterial. It is the *incidence* of the complication that relates the probability of the complication, given SLE.

If, as seems likely, notions of disease commonness are based on how frequently patients with the disease are encountered without regard to whether they represent new or existing diagnoses, the resulting amalgam of incident and prevalent cases will bias intuitions about what is common, making many diseases appear to be more common than they are. Referral bias and clinicopathological conferences may similarly skew intuitions about disease incidence, since patients with rare diseases are concentrated in these samples compared to unselected patients<sup>15,27,28</sup>. Indeed, clinicopathological conferences and grand rounds customarily select the rarest diseases for presentation, turning the natural order of disease frequency topsy-turvy<sup>28</sup>.

Fortunately, commonness need not be based upon intuitions: because of the proliferation of epidemiological cohort data in recent decades, the incidence of most diseases can now be readily found in epidemiological cohort series. Similarly, online resources such as [www.uptodate.com](http://www.uptodate.com) commonly report disease incidences under a subsection on epidemiology. Estimates from these sources are not always in agreement, but the problems posed by variability are not as serious as they may seem; precise incidences are not necessary. Worthwhile comparisons can be made based on order of magnitude differences in incidence estimates for different diseases.

For example, Table 1 shows that the incidence of pneumonia is approximately 650 cases per 100,000 persons-years. By comparison, that of segmental pulmonary embolism is on the order of 60 per 100,000 person-years. (UpToDate was used as the default source for incidence data for the sake of simplicity and ease of use and to limit the size of the bibliography.) Suppose we were to make a differential diagnosis for dyspnea in a patient presenting to the emergency department before any individuating information about the illness was known that would allow us to differentiate between pneumonia and pulmonary embolism. (A scenario such as this occurs countless times each day as emergency room physicians approach a patient's room with nothing more than age, gender, and chief complaint recorded by the triage nurse on the intake sheet.) Based on incidence alone, we could say that pneumonia is an order of magnitude more likely than pulmonary embolism; it is more than five orders of magnitude more likely than lymphangiomyomatosis. Indeed, diseases with incidences of less than 1/100,000 person-years are so rare that most clinicians outside of specialized referral centers will diagnose new (incident) cases on average no more than once or twice during their entire career<sup>14</sup>.

| Incident Disease                         | Annual Incidence* | Source+                       |
|--|-------------------|-------------------------------|
| Community Acquired Pneumonia             | 650               | UpToDate                      |
| Asthma requiring hospitalization         | 119               | UpToDate                      |
| Pulmonary Embolism after advent of CTPA  | 112               | UpToDate                      |
| Pulmonary Embolism before advent of CTPA | 62                | UpToDate                      |
| Lung Cancer                              | 59                | Dela Cruz <sup>29</sup>       |
| Idiopathic Pulmonary Fibrosis            | 10                | UpToDate                      |
| Sarcoidosis                              | 10                | UpToDate                      |
| Systemic Lupus Erythematosus             | 5                 | Stojan <sup>30</sup>          |
| Tuberculosis                             | 3                 | UpToDate                      |
| Hypersensitivity Pneumonitis             | 2                 | Fernandez-Perez <sup>31</sup> |
| GPA (Wegener's granulomatosis)           | 1                 | Mohammad <sup>32</sup>        |
| Idiopathic Pulmonary Hypertension        | 0.7               | Prins <sup>33</sup>           |
| Behcet's Syndrome                        | 0.38              | Calamia <sup>34</sup>         |
| Lymphangiomyomatosis                     | 0.03              | Harknett <sup>35</sup>        |

**Table 1.** Annual incidences of selected pulmonary diseases. \*Incidence per 100,000 person-years. +UpToDate ([www.uptodate.com](http://www.uptodate.com)) was accessed on January 13, 2020 and was used as the default source for incidence data. Additional sources were consulted (as referenced) if UpToDate did not provide an estimate of incidence for the disease. CTPA, Computed Tomography Pulmonary Angiography

As is evident from Table 1, there is no clear dividing line between common and rare diseases, which exist on a continuum. Likewise, there is no precise method or formula for taking the raw incidence rate in an

unselected member of the population from Table 1 and transforming it into a probability of that disease for a specific patient<sup>14,15,26</sup>. However, some back-of-the-envelope calculations are illuminating. Suppose that a patient is being admitted through the emergency department with symptoms compatible with pneumonia and the physician estimates, based on experience, that the probability of pneumonia is on the order of 65%; two out of three times he admits a similar patient, the diagnosis is ultimately confirmed to be pneumonia. The remaining 35% of the diagnostic probability space<sup>4</sup> is shared among several less likely possibilities including lung cancer, atelectasis, infarction, etc. Since, as seen in Table 1, the annual incidence of lung cancer is approximately one tenth that of CAP, we could estimate that the probability of lung cancer is on the order of 6.5%, or one tenth of the 65% probability of CAP. This crude approximation will pass muster with physicians who regularly admits patients with findings compatible with pneumonia – a handful of them are ultimately diagnosed with lung cancer rather than (or in addition to) pneumonia. Similarly, if a medical student included tuberculosis on the differential diagnosis for this patient, and she were pressed on how likely it is, she could respond that the incidence of tuberculosis is just 1/200<sup>th</sup> that of CAP making the probability in this patient, *ceteris paribus*, 65%/200 or about 0.33%. This number also has face validity for experienced clinicians: the authors, pulmonologists at an academic medical center, admit on the order of one hundred or more cases of community acquired pneumonia for every case of tuberculosis they ultimately diagnose.

With this background, we may now attempt to answer the questions posed at the outset. What diseases are common? Diseases with the highest incidences, (e.g., community acquired pneumonia) on the order of hundreds of cases per 100,000 person-years; practitioners are likely to encounter these diseases commonly – daily or weekly – in general medical practice. What diseases are rare? Diseases such as pheochromocytoma, with an incidence on the order of one case or fewer per 100,000 person-years; diagnosticians are likely to encounter new cases of such diseases on the order of once during their entire career<sup>14,36</sup>. Is sarcoidosis common? This question is more difficult since sarcoidosis does not fall on the extremes of the incidence continuum. What we can say is that community acquired pneumonia is 65 times more common; diagnosticians are likely to encounter something like 65 new cases of pneumonia for each newly diagnosed case of sarcoidosis they encounter. Should commonness be assessed according to incidence or prevalence? For the purposes of diagnosis, it is related to the incidence of new, previously undiagnosed disease. (For the purposes of healthcare expenditures or burden of disease, it is better assessed by prevalence.) Can the notion of commonness be operationalized in a practicable way to assist in the assignment of diagnostic probabilities? We hope to have shown that it can.

What are the implications of these answers? Despite decades of articles and dozens of books on clinical problem solving claiming that rational diagnosis and therapeutics require formal probabilities of disease<sup>14-16,37-41</sup>, controversy about the role of probability in diagnosis is ongoing<sup>6,10,11,42-48</sup>. This may stem from the fact that the formal systems proposed for probabilistic problem solving are too complicated and cumbersome for day-to-day use in the hustle and bustle of medical practice<sup>15,26,40,41,49</sup>. Thus, experts typically arrive at a diagnosis by pattern recognition, and any use of probabilistic reasoning is intuitive rather than explicit, rarified exceptions (including one of the authors) notwithstanding<sup>22,50,51</sup>. But perhaps the baby has been discarded with the bathwater: the unsuitability of complex decision trees for everyday use does not mean that probability is irrelevant to diagnosis, rather that its use must be simplified to be practicable.

The enduring popularity of the CTC axiom and its metaphorical variants is a tacit acknowledgment of the importance of probability in diagnosis. However, CTC is often invoked after the fact as a corrective (as when thrombotic thrombocytopenic purpura is disproved by positive blood cultures) rather than as a general guide for estimating pre-test probability of diseases. Resolving ambiguity about how to determine what is common may make the axiom more practicable and the notion of probability more tractable for diagnosticians. We propose that using incidence to compare the relative frequencies of competing diseases early in the differential diagnosis may avert base rate neglect and related biases from the first. This advice may prove especially useful for students and trainees who do not yet have an intuitive sense of disease frequencies. For them, knowing that the incidence of pneumonia is an order of magnitude greater than that of pulmonary embolism will represent a pedagogical leap over the tautological admonition that “common things are common.”

There are several limitations to the use of incidence as a starting point for estimating the probability of disease. As noted, Table 1 shows raw incidences in unselected persons in the population. For many diseases, the incidence varies markedly in subsets of patients stratified by age, gender, race, geography, and other factors; diagnosticians must attend to these differences to find an incidence rate appropriate for a specific (as opposed to unselected) patient. For diseases such as HIV where screening programs are in place, the incidence will reflect asymptomatic cases detected by screening in addition to those diagnosed because of symptoms, comingling incidence and prevalence. For specialists and those practicing at specialized centers, patients referred from other physicians or facilities will likely have enriched probability of rare disease, sometimes remarkably so. For many diseases, unique combinations of presenting features or pathognomonic signs and symptoms will give large probability boosts to otherwise rare diseases – i.e., the rare disease is not rare in the specific clinical scenario. Some rare diseases, e.g., vibrio vulnificus sepsis, are not rare in the presence of strong risk factors such as hemochromatosis and consumption of raw oysters. Lastly, incidence estimates are susceptible to a host of epidemiological biases including over- and under-diagnosis and their internal and external validity can be uncertain. Most of these issues, once acknowledged, can be accounted for, or serve as an injunction against relying too heavily upon incidence in specific cases.

Because of its immanent probabilistic and stochastic nature, diagnosis is a special type of forecasting. Expert forecasters of all stripes use base rates and probabilistic reasoning explicitly in the development of their forecasts<sup>9,13,52</sup>. In addition to improving accuracy, doing so enables feedback to be more easily brought to bear on the process, improving subsequent forecasts<sup>8,13,52</sup>. Failure to make explicit probabilistic predictions hinders calibration by fostering the “I knew it all along” effect of hindsight bias<sup>52-55</sup>. Physicians, when they have been studied, have shown unimpressive proficiency at forecasting and calibration<sup>7,10,53</sup>. Whether explicit consideration of numerical disease incidences can improve diagnostic accuracy must await empirical research dedicated to that question. Such research will require new paradigms of physician evaluation that depart from the traditional use of material focused on rare diseases and recondite knowledge, untethered to base rates, to assess medical competence and clinical acumen. Meanwhile, we see little downside to considering incidence as a starting point for gauging what diseases are common for the purpose of diagnosis. “Common diseases have the highest incidences” merely revises the tautology, and admittedly makes for a less euphonious axiom. Nonetheless, being more explicit and concrete, it holds the promise of providing practicable guidance for how to integrate probability into diagnosis.

1. Hutchison R. An Address on THE PRINCIPLES OF DIAGNOSIS. *British Medical Journal*. 1928;1(3504):335-337.
2. Tumulty PA. *The effective clinician: His methods and approach to diagnosis and care*. Philadelphia: Saunders; 1973.
3. Harvey AM, Bordley J., Barondess, J.A. *Differential Diagnosis: The Interpretation of Clinical Evidence*. Philadelphia: W.B. Saunders; 1979.
4. Aberegg SK, Johnson SA. When Alternative Diagnoses Are More Likely Than Pulmonary Embolism: A Paradox. *Ann Am Thorac Soc*. 2020.
5. Dolan JG, Bordley DR, Mushlin AI. An evaluation of clinicians’ subjective prior probability estimates. *Medical decision making : an international journal of the Society for Medical Decision Making*. 1986;6(4):216-223.
6. Puhan MA, Steurer J, Bachmann LM, ter Riet G. Variability in diagnostic probability estimates. *Ann Intern Med*. 2004;141(7):578-579.
7. Bushyhead JB, Christensen-Szalanski JJ. Feedback and the illusion of validity in a medical clinic. *Medical decision making : an international journal of the Society for Medical Decision Making*. 1981;1(2):115-123.
8. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge University Press; 2006.

9. Tetlock PaG, D. Superforecasting: The art and science of prediction. 2015.
10. Morgan DJ, Pineles L, Owczarzak J, et al. Accuracy of Practitioner Estimates of Probability of Diagnosis Before and After Testing. *JAMA internal medicine*. 2021.
11. Cahan A. Diagnosis is driven by probabilistic reasoning: counter-point. *Diagnosis*. 2016;3(3):99-101.
12. Cahan A, Gilon D, Manor O, Paltiel O. Probabilistic reasoning and clinical decision-making: do doctors overestimate diagnostic probabilities? *QJM : monthly journal of the Association of Physicians*. 2003;96(10):763-769.
13. Dhaliwal G, Detsky AS. The evolution of the master diagnostician. *Jama*. 2013;310(6):579-580.
14. Kassirer JP, Wong, J., Kopelman, R. *Learning CLinical Reasoning*. Second ed: Lippincott Williams Wilkins; 2009.
15. Ledley RS, Lusted LB. Reasoning Foundations of Medical Diagnosis. *Symbolic logic, probability, and value theory aid our understanding of how physicians reason*. 1959;130(3366):9-21.
16. O'Connor GT, Sox HC. Bayesian Reasoning in Medicine: The Contributions of Lee B. Lusted, MD. *Medical Decision Making*. 1991;11(2):107-111.
17. BURCHELL HB. Unusual Forms of Heart Disease. *Circulation*. 1954;10(4):574-579.
18. Montgomery K. *How doctors think: Clinical judgment and the practice of medicine*. Oxford: Oxford University Press; 2006.
19. Imperato PJ. *Medical Detective*. New York: R Marek; 1979.
20. Sotos JG. *Zebra Cards: An Aid to Obscure Diagnosis*. Third ed. USA: Mt Vernon Book Systems; 1991.
21. Bar-Hillel M. The base-rate fallacy in probability judgments. *Acta Psychologica*. 1980;44(3):211-233.
22. Weber EU, Böckenholt U, Hilton DJ, Wallace B. Determinants of diagnostic hypothesis generation: Effects of information, base rates, and experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1993;19(5):1151-1164.
23. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. *Science*. 1974;185(4157):1124-1131.
24. Elstein AS. Heuristics and biases: selected errors in clinical reasoning. *Academic medicine : journal of the Association of American Medical Colleges*. 1999;74(7):791-794.
25. Meador CK. *A Little Book of Doctors' Rules I*. Hanley & Belfus; 1992.
26. Balla JJ, Iannsek R, Elstein A. Bayesian diagnosis in presence of pre-existing disease. *Lancet*. 1985;1(8424):326-329.
27. Salive ME. Referral Bias in Tertiary Care: The Utility of Clinical Epidemiology. *Mayo Clinic Proceedings*. 1994;69(8):808-809.
28. Eddy DM, Clanton CH. The art of diagnosis: solving the clinicopathological exercise. *N Engl J Med*. 1982;306(21):1263-1268.
29. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin. Chest Med*. 2011;32(4):605-644.
30. Stojan G, Petri M. Epidemiology of systemic lupus erythematosus: an update. *Curr. Opin. Rheumatol*. 2018;30(2):144-150.

31. Pérez ERF, Kong AM, Raimundo K, Koelsch TL, Kulkarni R, Cole AL. Epidemiology of Hypersensitivity Pneumonitis among an Insured Population in the United States: A Claims-based Cohort Analysis. *Annals of the American Thoracic Society*. 2018;15(4):460-469.
32. Mohammad AJ, Jacobsson LTH, Westman KWA, Sturfelt G, Segelmark M. Incidence and survival rates in Wegener's granulomatosis, microscopic polyangiitis, Churg–Strauss syndrome and polyarteritis nodosa. *Rheumatology*. 2009;48(12):1560-1565.
33. Prins KW, Thenappan T. World Health Organization Group I Pulmonary Hypertension: Epidemiology and Pathophysiology. *Cardiol. Clin.* 2016;34(3):363-374.
34. Calamia KT, Wilson FC, Icen M, Crowson CS, Gabriel SE, Kremers HM. Epidemiology and clinical characteristics of Behcet's disease in the US: a population-based study. *Arthritis Rheum.* 2009;61(5):600-604.
35. Harknett EC, Chang WY, Byrnes S, et al. Use of variability in national and regional data to estimate the prevalence of lymphangioleiomyomatosis. *QJM*. 2011;104(11):971-979.
36. Lau ES, Scirica B, Schaefer I-M, Miller AL, Loscalzo J. Hypertensive Heartbreak. *New England Journal of Medicine*. 2021;384(22):2145-2152.
37. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med*. 1980;302(20):1109-1117.
38. Pauker SG, Kassirer JP. Therapeutic Decision Making: A Cost-Benefit Analysis. *New England Journal of Medicine*. 1975;293(5):229-234.
39. Djulbegovic B, van den Ende J, Hamm RM, Mayrhofer T, Hozo I, Pauker SG. When is rational to order a diagnostic test, or prescribe treatment: the threshold model as an explanation of practice variation. *Eur J Clin Invest*. 2015;45(5):485-493.
40. Weinstein MC, Fineberg, H. V. *Clinical Decision Analysis* Philadelphia: W.B. Saunders; 1980.
41. Sox HC, Higgins, M.C., Owens, D.K. *Medical Decision Making*. second ed. UK: John Wiley & Sons, Ltd.; 2013.
42. Jain BP. Why is diagnosis not probabilistic in clinical-pathological conference (CPCs): Point. *Diagnosis (Berlin, Germany)*. 2016;3(3):95-97.
43. Reid MC, Lane DA, Feinstein AR. Academic calculations versus clinical judgments: practicing physicians' use of quantitative measures of test accuracy. *Am J Med*. 1998;104(4):374-380.
44. Manrai AK, Bhatia G, Strymish J, Kohane IS, Jain SH. Medicine's uncomfortable relationship with math: calculating positive predictive value. *JAMA internal medicine*. 2014;174(6):991-993.
45. Casscells W, Schoenberger A, Graboys TB. Interpretation by physicians of clinical laboratory results. *N Engl J Med*. 1978;299(18):999-1001.
46. Richardson WS. Five uneasy pieces about pre-test probability. *J Gen Intern Med*. 2002;17(11):882-883.
47. Sanders S, Doust J, Glasziou P. A systematic review of studies comparing diagnostic clinical prediction rules with clinical judgment. *PLoS One*. 2015;10(6):e0128233.
48. Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. *Ann Intern Med*. 1999;130(12):995-1004.
49. Berwick DM, Fineberg HV, Weinstein MC. When doctors meet numbers. *Am J Med*. 1981;71(6):991-998.
50. Norman G. Building on experience—the development of clinical reasoning. *N Engl J Med*. 2006;355(21):2251-2252.

51. Brush JE, Jr, Brophy JM. Sharing the Process of Diagnostic Decision Making. *JAMA internal medicine*. 2017;177(9):1245-1246.
52. Arkes HR. Overconfidence in judgmental forecasting. In: Armstrong JS, ed. *Principles of Forecasting: A Handbook for Researchers and Practitioners* . Massachusetts: Kluwer Academic Publishers; 2001.
53. Dawson NV, Connors AF, Jr., Speroff T, Kemka A, Shaw P, Arkes HR. Hemodynamic assessment in managing the critically ill: is physician confidence warranted? *Medical decision making : an international journal of the Society for Medical Decision Making*. 1993;13(3):258-266.
54. Dawson NV, Arkes HR, Siciliano C, Blinkhorn R, Lakshmanan M, Petrelli M. Hindsight bias: an impediment to accurate probability estimation in clinicopathologic conferences. *Medical decision making : an international journal of the Society for Medical Decision Making*. 1988;8(4):259-264.
55. Arkes HR, Wortmann RL, Saville PD, Harkness AR. Hindsight bias among physicians weighing the likelihood of diagnoses. *J Appl Psychol*.1981;66(2):252-254.