

Editorial

Harry B. Burke*

Quantifying diagnostic excellence

<https://doi.org/10.1515/dx-2022-0104>

Keywords: diagnosis; diagnostic education; diagnostic error; diagnostic excellence.

A diagnosis is the name of a medical condition. The Ancient Greeks believed that if something did not have a name, it could not be controlled – so it is with medicine. The goal of the diagnostic process is to give a medical name to a patient’s condition, where names are indexes of sets of necessary and sufficient facts (signs, symptoms, and tests) that identify specific medical conditions and which are, hopefully, linked to effective treatments. Many Ancient Greeks also believed that the truths underlying reality are mathematical. Thus, our naming of diseases may be aided by the formalism of mathematical equations.

Functionally, a diagnosis is a prediction of the presence in a patient of a specific disease at that moment in time. In the real world, many medical conditions are difficult to predict because their diagnostic criteria may overlap with other diagnoses, the diagnostic information may be incomplete, and what information is available may not be accurately communicated to the physician by the patient. The goals of this paper are to quantify the ability of physicians to perform the diagnostic process and to provide a numerical diagnostic excellence score that reflects their performance. A companion independent paper by Shimizu and Graber presents an equation for excellence in diagnostic reasoning [1].

In order to gain a perspective on diagnosis, we examine the ideal diagnostic event. It has three components. (1) Condition: the medical condition has a diagnosis and it is unambiguous. (2) Presentation: sufficient information is present at presentation for a correct diagnosis. (3) Physician: has a comprehensive fund of knowledge (from remembered knowledge and from an electronic knowledge base), follows the correct diagnostic algorithm (e.g.,

differential rule in/rule out process), acquires all the relevant information (e.g., patient, examination, tests, etc.), uses proper reasoning to assess the acquired information (e.g., probabilistic reasoning), doesn’t allow cognitive biases to affect decision-making, is not distracted or interrupted during the diagnostic process, and makes the diagnosis within a reasonable period of time. The amount of time that is reasonable depends on whether the disease is diagnosable at presentation or if it requires additional information and, if it requires additional information, if that information was acquired and utilized within a reasonable period of time.

From the physician component (#3 above) of the ideal diagnostic event, we derive a formula for calculating a physician’s diagnostic excellence score.

$$\text{Diagnostic excellence score (E)} = [K + A + I + R] - [B + Z + T] \quad (1)$$

where K is fund of knowledge, A is following a diagnostic algorithm, I is information acquisition, and R is proper reasoning. B is cognitive bias, Z is distraction, and T is the time interval from presentation to diagnosis. Each variable in the formula requires at least one or more measurement instruments to assess its relationship to diagnostic excellence. For all the formulas in this paper, if a variable is the result of more than one instrument, a sub-formula may be needed to integrate the scores of the multiple instruments. The methods of diagnostic data collection include, but are not limited to: assessing physicians’ documentation (including reporting their rule in/rule out reasoning and rank ordering), observing their interactions with patients, and evaluating their activities in diagnostic simulations.

We can illustrate the scoring of this formula. The score for knowledge, algorithm, information, and reasoning may range from 0 to 25 points each (higher is better). The score for cognitive bias, distraction, and time may range from 0 to 25 points each (lower is better); these scores increase from zero as bias and distraction increase and as the time from presentation to diagnosis crosses a preset threshold. In this illustration, the score can range from +100 to –75. The variable scores possess equal weight in the equation but their weights can be modified based on what is learned about their relationship to diagnostic excellence.

Diseases are not equally easy to diagnose and diagnostic excellence is affected by how difficult the disease is

*Corresponding author: Harry B. Burke, MD, PhD, Professor of Medicine, Chief, Section of Safety, Quality, and Value, Director, DHA-USU Patient Safety and Quality Academic Collaborative, Department of Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814, USA, Phone: 301-295-4162, Mobile: 301-938-2212, E-mail: harry.burke@usuhs.edu. <https://orcid.org/0000-0003-3474-4609>

to diagnose. Therefore, we can adjust the physician's diagnostic excellence score by the difficulty inherent in diagnosing a specific disease. Disease diagnostic difficulty arises from its: (1) frequency of occurrence (how rare the condition is the everyday practice of medicine), (2) diagnostic complexity (the number, sophistication, and interaction of the signs, including tests and procedures, and symptoms), and (3) inherent diagnostic ambiguity (how difficult it is to distinguish from other medical conditions). Using this information, we derive a formula that adjusts for disease diagnostic uncertainty.

$$\text{Disease diagnostic difficulty } (D) = F + X + A \quad (2)$$

where F is the frequency of occurrence, X is complexity, and A is ambiguity. Each variable requires at least one measurement instrument.

In addition to diagnostic uncertainty, other adjustments are required related to the diagnostic conditions. Because all the necessary information may not be available, the score must be adjusted for the available information. Furthermore, because not all physicians have the same resources (e.g., time to assess the patient, availability of laboratory and imaging tests, etc.), there must be an adjustment to the score that reflects the available resources. Using this information, we derive a formula that adjusts for disease difficulty.

$$\text{Disease diagnostic conditions adjustment } (C) = D + L + S \quad (3)$$

where D is the disease diagnostic difficulty score, L is lack of diagnostic information, and S is few system resources. The larger the C , the more difficult the diagnosis. Each variable requires at least one measurement instrument.

Combining the physician component of the diagnostic excellence score (Equation (1)) and the disease difficulty adjustment (Equation (3)), allows us to derive a formula for calculating a physician's disease-adjusted diagnostic excellence score.

$$\begin{aligned} \text{Disease diagnostic conditions excellence score } (Dc) \\ = C\{[K + A + I + R] - [B + Z + T]\} \end{aligned} \quad (4)$$

where C is the disease difficulty adjustment (Equation (3)) and the remainder of the equation is the disease excellence score (Equation (1)).

It is important to note that the disease diagnostic conditions excellence score does not replace the diagnostic excellence score. The diagnostic excellence score provides information regarding how well the physician applies the diagnostic process, while the disease diagnostic conditions excellence score provides information regarding how well the physician handles varying degrees of disease difficulty

and conditions. For example, a physician may perform well on easy diagnoses but poorly on difficult diagnoses or adverse conditions.

For the diseases diagnosed by a physician, the disease diagnostic conditions excellence score (Dc) can be averaged across the physician's panel to obtain the physician's clinical disease diagnostic conditions excellence score.

$$\begin{aligned} \text{Clinical disease diagnostic conditions excellence score } (cDc) \\ = \frac{\sum_{i=1}^n i = D1 + D2 + \dots + Dn}{N} \end{aligned} \quad (5)$$

where Dc is the disease diagnostic conditions excellence score for each patient and N is the number of patients.

Formulas 1 through 5 take a normative approach to diagnostic excellence; each variable has its own standard for assessing its relationship to the physician's performance and diagnostic excellence is the sum of their scores. However, one can ask; why not train a multivariate regression model on diagnostic errors and calculate the diagnostic accuracy score? We can create a logistic regression formula for calculating diagnostic accuracy.

Inferential diagnostic accuracy score

$$= \zeta(\beta_1 K + \beta_2 A + \beta_3 I + \beta_4 R + \beta_5 B + \beta_6 Z + \beta_7 T) \quad (6)$$

where ζ is the logit, and the remaining variables are the weighted diagnostic excellence factors. (Note that for simplicity, conditions was not included in this formula).

There are several reasons why we cannot, at the current time, use diagnostic error as a dependent variable. (1) Many errors go undetected and unreported, and the reported errors are biased by how they are detected and reported. In other words, the errors we currently observe, that is, the errors that we would use to train the models, are not representative of the universe of errors. This means that the trained models will not properly reflect diagnostic accuracy, rather, they will reflect a biased error detection process. (2) There are many diagnostic categories of errors and degrees of harm. Unfortunately, the error categories and degrees of harm are poorly defined (there are no necessary and sufficient criteria for each error, each error category, and each harm) which means that some errors will not fit into any error category and many errors will fit into more than one category of error (this also applies to harm), resulting in category ambiguity; which degrades the utility of error as a dependent variable. (3) Each category and degree of harm may have its own etiology and each etiology may require its own variables and formula. Furthermore, because of their different etiologies, error categories and harms cannot be combined into a single dependent variable – a model has to

be created for each error category and harm, resulting in a proliferation of models. In other words, there will be uncertainty in diagnostic errors, the error categories are ambiguous, and there will be many error category and harm models.

In addition, there are problems related to training the regression models. The problems relate to the number of models, the number of independent variables in each model, and the number of diagnostic error events used to train the models. As noted above, many models are required to represent the various diagnostic errors, error categories, and harms. In order to capture the error mechanism, each model must have the correct independent variables and many independent variables may be required for a model. This means that for each error category and harm model, there must be enough errors to stabilize the variables coefficients and to create a correct model. Unfortunately, because errors are relatively rare there will probably not be enough events to train the inferential models. Finally, there is the issue of missing data. The independent variables must be provided with variable values, but it is difficult to find all the factors related to a diagnostic error and many of the variables, and their associated values, will not be found. In other words, there will be a great deal of missing data and the missing data mechanism will usually be unknown. To conclude, insufficient number of cases and of events will result in incorrect and unstable variable weightings and inaccurate predictions.

One purpose of a final diagnosis is to determine if the diagnosis predicted by a physician, or by a formula, is correct. When attempting to compare a predicted diagnosis to a final diagnosis several related issues must be considered. One issue is how to select one predicted and one final diagnosis. To perform this selection, a single threshold is set that applies to all diagnoses, i.e., that selects only one-and-only-one diagnosis from among the many possible diagnoses. The threshold must be capable of rejecting all the “maybe” and “it could be” diagnoses. Alternatively, one can try to determine a threshold for each possible diagnosis. It is not altogether clear how to predetermine predicted and final diagnosis thresholds for each disease. Another issue is the nature of the relationship between a predicted diagnosis and a final diagnosis. Their association is inexact for at least three reasons: (1) Both a predicted diagnosis and a final diagnosis possesses a degree of uncertainty, thus there isn’t a fixed connection between a predicted diagnosis and a final diagnosis, (2) they approach diagnosis from different temporal directions, the predicted diagnosis is prospective and final diagnosis is retrospective, and (3) they do not acquire and use exactly the same information. In other words, although there is an

association between a predicted diagnosis and a final diagnosis, there is not an invariant relationship between the two. This means that there will be some degree of error in determining diagnostic error. To summarize, diagnostic excellence differs from diagnostic accuracy in that excellence is a process and accuracy is an outcome. Diagnostic accuracy faces a host of challenges that are not present for diagnostic excellence.

Our diagnostic excellence goals are to learn how to define and score the variables in the diagnostic process formulas. In addition, we would like to improve physicians’ diagnostic excellence scores. To do so, it is useful to view diagnostic excellence as a skill. The formula for diagnostic skill is:

$$\text{Diagnostic skill} = \text{training} + \text{experience} + \text{talent} \quad (7)$$

Medical educators have observed that in a group of medical students and residents with the same training, experience, and level of effort, a few students usually stand out. We say that those students have a talent for medicine. From a societal perspective, in addition to students being trained in diagnosis in medical school and residency, and physicians participating in lifelong diagnostic learning activities, medical schools should select for students with a talent for medicine. Unfortunately, at the current time, we do not know how to select for these students but, perhaps in the future, we will develop methods that can identify them.

Diagnostic excellence refers to how well the physician performs the diagnostic process. Some would like to add administrative requirements to diagnostic excellence. For example, they would like diagnostic excellence to include the idea of diagnostic efficiency, e.g., that the diagnostic process is quick, uses the fewest resources, and is cost-effective. Unfortunately, the desire to obtain a quick and inexpensive diagnosis may conflict with the need to achieve a correct diagnosis. This is because, in the service of efficiency, physicians may not acquire, or have sufficient time to understand, all the information that is necessary for a proper diagnosis. In other words, quick, resource-limited diagnoses may not enhance diagnostic excellence, rather, they may degrade it.

As a practical matter, taking into account all the factors required to calculate the approximate probability of each possible disease in a differential diagnosis list is very complex and difficult. Therefore, physicians should use, as part of the diagnostic process, an electronic clinical decision support system to assist them in calculating the approximate probability of each potential diagnosis.

Historically, tests of medical knowledge have been used as a surrogate for diagnostic excellence. But this

approach is divorced from the practical reality of medical care; it does not acknowledge how well the information is being acquired and used in the clinical encounter. Although a fund of knowledge is important, no physician has comprehensive knowledge. Therefore physicians should use, as part of the diagnostic process, an electronic knowledge database.

In conclusion, it is possible to quantify a physician's diagnostic excellence for a disease and for their panel of patients. This information will allow us to better understand the diagnostic process, to enhance physician diagnostic performance and, as a direct result, to improve patient care.

Research funding: Funded by the DHA-USU Patient Safety and Quality Academic Collaborative.

Author contributions: Single author contribution.

Competing interests: Author states no conflict of interest.

Informed consent: Not applicable.

Ethical approval: Not applicable.

Reference

1. Shimizu T, Graber ML. An equation for excellence in diagnostic reasoning. *Diagnosis* 2022, in print. <https://doi.org/10.1515/dx-2022-0007>.