Probability and uncertainty in clinical and forensic medicine

Joseph M. Guileyardo, MD

In medicine, and in life in general, uncertainty is unavoidable. Throughout history, man has tried to bring order to a chaotic world by predicting future events based on past experiences. However, physicians in training are often surprised by the subjectivity and lack of certainty involved in the practice of medicine. Helpful advice is available—and Sir William Osler had much to say about this problem. This article reviews some of Osler's writings regarding medical uncertainty, provides an outline of the classification currently used by the courts to evaluate medical testimony, and offers some historical notes on probability theory and evidence-based medicine.

In medicine, absolute diagnostic certainty is the exception rather than the rule (1). Nevertheless, many physicians in training are uncomfortable with the degree of uncertainty involved in therapy and diagnosis, although Sir William Osler wrote, “Medical education is, to a large extent, education for uncertainty” (2). Osler also set forth some reasons for diagnostic uncertainty and suggested ways to deal with it. The forensic pathologist soon learns that medical uncertainty is not surprising to the courts, who have established more or less formal categories to deal with levels of uncertainty in medical testimony. Furthermore, the entire structure of evidence-based medicine rests largely on statistical associations, which themselves define a certain degree of uncertainty. Therefore, uncertainty is an integral and unavoidable part of medical practice (and life in general).

OSLER ON UNCERTAINTY

Sir William told medical students:

A distressing feature in the life you are about to enter, a feature which will press hardly upon the finer spirits among you and ruffle their equanimity, is the uncertainty which pertains, not alone to our science and art, but to the very hopes and fears which make us men. In seeking absolute truth, we aim at the unattainable, and must be content with finding broken portions (3).

The practice of medicine is an art based on science, but Osler believed that this was one of the most difficult arts in the world to acquire. He further stated the extremely critical point that “errors in judgment are inevitable in the practicing of an

art which consists largely of balancing probabilities” (2). This balancing of probabilities is integral to the day-to-day activities of the physician who produces a differential diagnosis and then systematically confirms or eliminates the various hypotheses. However, as Osler pointed out, “Variability is the law of life and no two bodies are exactly alike, and no two individuals react alike under the abnormal conditions which we know as disease. This is the fundamental difficulty in the education of the physician” (2).

Such a statement is surprisingly timely in our era of personalized or precision medicine, in which targeted therapies are tailored to the individual genetics of a patient at the molecular level. This variability among patients also is why experienced physicians, like William C. Roberts, MD, approach every case with the question: “What is unusual about this case?”

Osler also stated, “Probability guides us when certainty fails,” but he emphasized that uncertainty is unavoidable. He further stated, “There is no discredit, though much discomfort at times in the everlasting *perhaps* which must preface so much connected with the practice of our art.” Therefore, physicians in training should not be surprised that *opinion* (not full knowledge) “must be their stay and prop” (2).

UNCERTAINTY AND THE COURTS

Expressing medical opinions under oath is a common activity of the forensic pathologist, and the degree of certainty required depends on the actions and claims that rest on them. The following discussion provides a general outline of the various categories and levels of certainty as used by the courts (1).

Proof beyond the shadow of a doubt is a concept derived from written and televised fiction dealing with the law. Nevertheless, “absolute diagnostic certainty” can be considered to represent the highest level of probability. This is certainty beyond a possible doubt where the likelihood is 100%. One must hold that every other imaginable contingency is impossible. In reality, the law requires lesser degrees of proof, depending on the type of proceeding.

From the Department of Pathology, Baylor University Medical Center at Dallas.

Corresponding author: Joseph M. Guileyardo, MD, Department of Pathology, Baylor University Medical Center at Dallas, 3500 Gaston Avenue, Dallas, TX 75246 (e-mail: guileyardo@sbcglobal.net).

A “reasonable degree of medical *certainty*” is the level of proof required in criminal proceedings. In this category, the conclusion is beyond any “reasonable” doubt. This category has no formal definition of numerical probability, but the opinion is generally considered as *significantly exceeding 50% likelihood*.

A “reasonable degree of medical *probability*” is the next lower level of proof as used by the courts. In this category, the conclusion is defined as *having a probability greater, but not significantly higher, than 50%*. This is the degree of probability or burden of proof required in most civil (tort) proceedings, such as medical malpractice cases. This is also the level of certainty required for opinions given on death certificates. In this category, any contingency having a likelihood exceeding 50% may be considered *probable or more likely than not*. For example, in a medical malpractice case, it must usually be established by expert testimony that a deviation from the standard of care, *more likely than not*, occurred and led to the patient’s injury. (Adherence to the “standard of care” simply means the physician did what a “reasonably prudent” physician would have done under similar circumstances.) In such cases, the opposing experts often disagree, and the jury must decide which expert or argument is most credible.

Interestingly, this level of certainty is essentially the same burden of proof used by grand juries who determine whether *probable cause* exists that a crime has been committed and whether an indictment (true bill) should be returned against a defendant. Probable cause simply means “having more evidence for than against.” If a true bill of indictment is returned by the grand jury, the case enters criminal proceedings where a higher burden of proof (beyond a reasonable doubt) is necessary for conviction.

A “reasonable *possibility*” involves an opinion or conclusion whose probability is less than 50%. This generally means that the hypothesis does not defy the laws of nature. Such a standard has negligible evidential importance in court; however, a medical witness may be asked if some scenario is “possible” or “conceivable” in order to establish that he or she is not entirely excluding the opinion given by the opposing expert, even though he believes that it is unlikely. Quasi-judicial proceedings such as workmen’s compensation boards may be satisfied with this degree of certainty in a case where the presumption ordinarily favors the worker.

“Speculation” involves a possibility that is so remote that it has no basis in reason. In court, this is usually considered to represent “a *guess*.” Generally, speculation is not allowed in civil or criminal testimony. Furthermore, if a witness wants to generate some excitement in court, he may preface an opinion with “I would speculate that . . .” This is guaranteed to generate a forceful objection by one of the attorneys, and often an admonition from the judge to refrain from speculation.

One caveat regarding the above categories is that medical opinions are based on facts as they are known at the time the opinion is rendered. If the facts change (such as when new testimony is given by witnesses), the opinions are likely to change as well. Some experts state in their reports: “These opinions are based on a standard of reasonable medical probability, using the currently available information; I reserve the right to

supplement or amend this report if warranted by additional information.” Although somewhat obvious, this statement may cause a judge to allow a second report to be submitted if the facts change.

STATISTICS AND CERTAINTY

Physicians must also understand that the statistical definition of “probable” used in medical research or evidence-based medicine applies only to data numerous enough to permit statistical analysis. A *P* value less than .05 merely indicates that the likelihood of a difference between two data sets being due to chance is less than 5% (1 chance in 20). This is an incorrect and inappropriate standard to require when rendering a medicolegal opinion in a single case, because a single datum is not amenable to statistical analysis (1). However, the use of statistical probability in general and the *P* value in particular are so important to conclusions regarding evidence-based medicine that a brief look at their origin and history is warranted.

Much of the mathematical foundations of probability theory and combinatorial algebra began with questions relating to certain gambling problems posed to French mathematicians Blaise Pascal and Pierre de Fermat in 1654 by Antoine Gombaud, the Chevalier de Méré. Subsequent developments, including the production of gambling manuals by other workers, led to mathematical formulas and methods that generate the bell-shaped curve, or normal distribution (4). Subsequently, in 1894, Karl Pearson coined the term *standard deviation*, and in 1925 Sir Ronald Fisher, in his book *Statistical Methods for Research Workers*, appeared to be the first to mention *P* = .05 as a level determining statistical significance (4).

Sir Ronald stated:

It is *convenient* to take this point as a limit in judging whether a deviation is to be considered significant or not. Deviations exceeding twice the standard deviation are thus formally regarded as significant. If one in twenty does not seem high enough odds, we may, if we prefer it, draw the line at one in fifty (the 2 percent point), or one in a hundred (the 1 per cent point). *Personally*, the writer [Fisher] prefers to set a low standard of significance at the 5 percent point, and ignore entirely all results which fail to reach this level. A scientific fact should be regarded as experimentally established only if a properly designed experiment *rarely fails* to give this level of significance (4).

Therefore, the selection of a cut-off for significance is somewhat arbitrary and raises philosophical questions such as whether people (scientists and nonscientists alike) generally feel that an event that occurs 5% of the time or less (when multiple trials are executed) is a rare event. If the answer is yes, then the adoption of this level as a criterion for judging outcomes is justifiable. Therefore, selection of a threshold depends on subjective interpretations, and, as a formal statement, the level has a more complex history than is generally appreciated (4).

The theories of probability that originated in a gambler’s dispute are now at the base of many enterprises that we consider more important than gambling, including all kinds of insurance, mathematical statistics, and their application to biology, medi-

cine, educational measurements, and much of modern theoretical physics. We no longer think of an electron being “at” a given place at a given instant, but we do calculate its probability of being in a given region. A little reflection will show that even the simplest measurements we make (blood pressure, etc.) are statistical in nature (5).

Definitions of probability are most frequently based on formal mathematical theory, but what often eludes precise definition (the level of “acceptable” uncertainty) is the subjective probability involved in the personal cognition of individuals whereby past experience aids in the formation of expectations for future events. Furthermore, our need to use mathematical probability theory, in a sense, reflects a need to bridge the reality of events in everyday life and the philosophy of logic (4).

CONCLUSION

The above discussions are intended to illustrate that uncertainty is unavoidable and should be expected in the practice of medicine (and in life as well), but useful guidelines are available. In the words of Sir William Osler:

We must collect facts in order to establish general principles. But in the practice of medicine, where our strength should be, lies our greatest weakness. Our study is man, as the subject of accidents or diseases. Were he always, inside and outside, cast in the same mold, instead of differing from his fellow man as much in constitution and in his reaction to stimulus as in feature, we should ere this have reached some settled principles in our art (2).

Still, he advised that probability guides us when knowledge fails.

In conclusion, the subtle theories of subjective and mathematical probability lie at the very roots of human knowledge and form the basis of much of scientific knowledge as we attempt to predict outcomes from scientific research and day-to-day experiences. Most agree that Pascal and Fermat stated and solved a genuine problem, that of bringing the superficial lawlessness of pure chance under the domination of law, order, and regularity (5). And, to our physicians in training, don't be surprised by diagnostic uncertainty, especially considering today's complex patient populations. Remain courageous and heed Dr. Osler's advice regarding life's challenges in general and difficult diagnostic and therapeutic choices in particular: “If the fight is for principle and justice, even when failure seems certain, where many have failed before, cling to your ideal, set the horn to your lips, sound the challenge, and calmly await the conflict” (2).

1. Spitz W, Spitz D. *Spitz and Fisher's Medicolegal Investigation of Death: Guidelines for the Application of Pathology to Crime Investigation*, 4th ed. Springfield, IL: Charles C. Thomas, 2006.
2. Osler W. *Aequanimitas, with Other Addresses*, 3rd ed. Philadelphia: Blakiston, 1932.
3. Bryan CS. *Osler: Inspirations from a Great Physician*. New York: Oxford University Press, 1997.
4. Cowles M, Davis C. On the origins of the .05 level of statistical significance. *American Psychologist* 1982;37(5):553–558.
5. Bell ET. *Men of Mathematics*. New York: Touchstone, Simon & Schuster, 1937.