The Impact of Laboratory Error on the Normal Range: A Bayesian Model


Interpretation of clinical laboratory results, aside from clinical considerations, is based on the probability of the result being within a given normal range. This probability is influenced by the degree of error inherent in the analytical method. It would be advantageous to assign a more definite probability to the result of the measurement by combining the error distribution of the result around the true value and the distribution of the healthy population that serves as a reference. Bayesian statistics permits the revision of this prior information into a single probability.

Additional Keyphrases: statistics • error distribution • probability statements • precision • normal values • quality control

Automation of the clinical laboratory has made possible health screening on a large scale. The medical profession has accepted health screening as a valuable tool for refining the diagnostic capabilities, but only limited advances have been made in improving the laboratory statement (1–3). This statement is usually expressed as a numerical value without reference to the error of measurement. On the basis of earlier work by Cotlove and coworkers (4–7) and by Gwendynock (8), we have analyzed the interaction of the analytical error on the normal range and we will show that the error of measurement can be incorporated into the laboratory statement by means of Bayesian statistics. No attempt will be made in this presentation to analyze the usefulness of the normal range for diagnostic purposes; difficulties in establishing normal ranges were recently discussed by Reed et al. (9, 10). The use of Bayesian statistics as an aid in clinical diagnosis has been proposed elsewhere (11).

If a laboratory result falls near the center of the normal range, no problem is usually encountered in assuming “normality” for the measured biological constituent. If it falls near the upper or lower limit of the normal range (e.g., mean ±2 SD), a decision has to be made whether the datum indicates a pathological process, particularly if the patient’s history and other tests are inconclusive and that laboratory result is to be used as the basis for further studies of the patient. Compounding this problem is the fact that the true concentration of a constituent cannot be determined with certainty owing to the analytical error associated with any chemical measurement.

It would be advantageous if one could assign to a given laboratory result a definite probability that the true value is within the normal limits, a probability that would take into account both the distribution of values of that chemical constituent in healthy individuals and the distribution of the laboratory result around the true value for that sample caused by the analytical error. Bayesian statistics offers a method by which a more definite probability may be calculated.

Division of Clinical Chemistry, Department of Pathology (R. D. K., V. D. A., and H.-D. G.) and the Department of Statistics (T. A. W.), Ohio State University, Columbus, Ohio 43210.

1 Postdoctoral Fellow in Clinical Chemistry, 1972–1974. Present address: Department of Pathology, Deaconess Hospital, Spokane, Wash. 99210.


3 Address reprint requests to H.-D. Gruemer, M.D.

A preliminary report was presented by V. D. Anand at the 25th National Meeting of the AACC, New York City, July 1973. Received June 17, 1974; accepted Dec. 6, 1974.
Method of Approach

The standard deviation \( (S_b) \) of a normal healthy reference population as determined in the laboratory may be considered a two-component system: the elements of the person-to-person biological (group) variation \( (S_b) \), and the variation \( (S_m) \) associated with the long-term error in measurement. These quantities are related by the expression:

\[
S_p = \sqrt{S_b^2 + S_m^2} \tag{1}
\]

Because every member of this population has been tested once only, intrapersonal variance is not included.

If the magnitude of the long-term analytical error is determined at the time the normal range is established, it is possible to calculate \( S_b \). Because \( S_b \) may be assumed to be constant for a given physiological constituent that is in a steady-state condition and to be independent of the magnitude of the analytical error, it would be more valuable than \( S_b \) for use in interlaboratory comparisons of normal ranges. Calculation of \( S_b \) when the normal range is established allows easy re-evaluation of \( S_b \) when the magnitude of the laboratory error changes.

The Bayesian Model

Bayes’ theorem is essentially an algebraic relationship by which prior information is revised in view of additional data as it becomes available. The prior information is represented by the distribution of concentrations of a biochemical constituent in the sera of a presumably healthy reference population and the sample information by the error distribution of analytical results around the observed value. The prior information and the sample information are combined to obtain a distribution (12, 13) around the most likely true value—in Bayesian language, the “posterior distribution” (Figure 1). The posterior distribution may be used for statistical manipulations that are impossible to carry out on two or more component distributions.

In this discussion, it is assumed that the prior distribution and the methodological distribution—and hence the posterior distribution—are gaussian. If so, the posterior mean \( (M_p) \) is an average of the prior mean \( (M_b) \) of the biological range \( (M_b \pm 2 S_b) \) and of the value of the observation \( (X) \), weighted by their respective precisions (precision is the reciprocal of the variance, \( 1/S^2 \)). This relationship can be stated mathematically as:

\[
M_p = \frac{M_b S_m^2 + X S_b^2}{S_m^2 + S_b^2} \tag{2}
\]

The precision of the posterior distribution is equal to the sum of the prior and methodological precisions. Thus, the standard deviation \( (S_p) \) of the posterior distribution is given by the expression:

\[
S_p = \frac{1}{\sqrt{S_b^2 + S_m^2}}
\]

or,

\[
S_p = \sqrt{\frac{S_m^2 S_b^2}{S_m^2 + S_b^2}} \tag{3}
\]

Serum calcium has been chosen as an example to illustrate the use of these procedures. This constituent is measured in our laboratory by an automated atomic absorption spectrophotometric method (14). The normal range was established by analysis of a large number of sera from blood donors. The distribution of calcium values for these samples was found to be essentially gaussian with a mean \( (M_b) \) of 9.5 mg/100 ml and a standard deviation \( (S_b) \) of 0.45 mg/100 ml. The conventional normal range \( (M_b \pm 2 S_b) \) is thus 8.6 to 10.4 mg/100 ml. At the time the normal range was established over some months, the long term analytical error \( (S_m) \) was 0.25 mg/100 ml, giving the biological variation \( (S_b) \) a value of 0.374 mg/100 ml \( (S_b = \sqrt{(0.45)^2 - (0.25)^2}) \). Quality-control data show that the \( S_m \) for calcium is concentration-independent.

Bayesian statistics may be used to determine the probability that an experimentally determined serum calcium value of 10.8 mg/100 ml (for example) is within the normal range.

To do this the mean \( (M_p) \) and standard deviation \( (S_p) \) of the posterior distribution must first be determined, by using equations 2 and 3:

\[
M_p = \frac{(9.5)(0.25)^2 + (10.8)(0.374)^2}{(0.374)^2 + (0.25)^2} = 10.399 \text{ mg/100 ml}
\]

and

\[
S_p = \sqrt{\frac{(0.374)^2(0.25)^2}{(0.374)^2 + (0.25)^2}} = 0.208 \text{ mg/100 ml}
\]

The relationship between these three distributions (prior, methodological, and posterior) is shown in Figure 1. Because the precision of the posterior distribution is equal to the sum of the precisions of the
prior and methodological distributions, the standard deviation for posterior distribution curve is smaller than those for the other two. The mean of the posterior distribution curve is the average of the prior mean and the observed result, weighted by their precisions. Therefore, the mean of the posterior distribution in our example is nearer to the observed result than to the mean of the prior distribution, because the methodological distribution is more precise than the biological (prior) distribution.

Because the posterior distribution contains information from both the biological distribution and the methodological distribution, the posterior distribution can be used in place of the other two distributions. Therefore, the clinical significance of a serum calcium level of 10.8 mg/100 ml is influenced by both the prior distribution and the methodological distribution. The decision as to the degree of normality of such a concentration can be made solely with knowledge obtained from the posterior distribution curve.

In the example given, the problem was to calculate the probability that a serum calcium concentration of 10.8 mg/100 ml is normal. The problem would be difficult if the computation were based on two independent distributions, but becomes trivial when only one distribution, the posterior, is involved. Thus, the problem reduces to the determination of the fraction of the area under the posterior distribution curve that lies to the left of the upper limit \(L_u\) of the biological range (10.248 mg/100 ml), where:

\[
L_u = M_b + 2S_b
\]

(4)

Inspection of Figure 1 shows that this fraction is approximately \(\frac{1}{4}\).

A numerical probability may be obtained by calculating the difference \(c\) between the mean of the posterior distribution and the upper limit of the biological range in units of standard deviation by using the equation:

\[
c = \frac{L_u - M_b}{S_b}
\]

(5)

and then looking up the answer in a standard probability table (15). For instance, the value of \(c\) in this example is:

\[
c = \frac{10.248 - 10.399}{0.208} = -0.726
\]

This value corresponds to a probability of 23.4%. Therefore, there is only about a 23% probability that a serum calcium level of 10.8 mg/100 ml is within the biological range \((M_b \pm 2S_b)\). A similar treatment applies in determining the probability that a laboratory value is above the lower limit of the biological range.

These calculations can be incorporated as part of the normal computer handling technique in larger clinical laboratories. The results of determinations that are known to follow a gaussian distribution may be expressed both as concentration and as the probability of being within the normal range. If it is assumed that \(S_b\) and \(M_b\) remain constant, the probability of a given analytical result being within the biological range depends on the value of \(S_m\) at the time the test was done. If \(S_m\) becomes larger, laboratory results that have 50% probability of being within the biological range will move farther away from the midpoint \((M_b)\) to a higher or lower concentration, thus broadening the normal range (Figure 2). Therefore, the probabilities should be calculated by using the currently appropriate value of \(S_m\), which should be determined and updated as frequently as possible by a concomitant quality-control program.

**Discussion**

The proposed model in its present form should not be applied to nongaussian distributions, unless they can be separated into gaussian subdistributions (16) and the subdistributions used, or the nongaussian distribution can be transformed to a gaussian curve. The standard deviation \((S_b)\) may have to be considered to be the composite of various biological components \((S_{b1}, S_{b2}, \ldots)\) as for example, age, sex, nutritional status, circadian rhythm, and seasonal changes. Such factors have to be isolated and evaluated for their impact on the normal or biological range and whether the standard conditions of the established normal range of a healthy reference population apply to the patient or to the condition under which the sample was obtained. However, this is true for all methods and is not limited to the use of Bayesian statistics.
Another prerequisite for the application of Bayesian statistics is that the measurement of the standard deviation ($S^m$) by the quality-control program must truly reflect the long-term laboratory error or at least be a good approximation of it.

In this model we have purposely ignored the existence of intra-individual variation. If intrapersonal variance, $S^p^2$, were considered, equation 1 would become:

$$S^p^2 = S^b^2 + S^i^2 + S^m^2$$

and the definition of normal limits as the 95% limit of a healthy reference population would be changed. Cotlove and coworkers (4–7) estimated the magnitude of intra-individual variation of some biological constituents. Unfortunately, intra-individual variation of calcium increases considerably in early hyperparathyroidism over that seen in normal individuals, even if the serum calcium concentrations are still within the normal population range. It is just the abnormal fluctuation of results that the physician wants to observe in repeated studies. Incorporation of an average $S^i$ from a healthy population could then easily be misleading, complicate the model considerably, and move the model presented here unwittingly away from the real-world situation.

For the purpose of presenting this model, we expressed the current state of information on a patient’s true value in terms of a distribution of concentrations (i.e., the posterior) or as the probability of being within ±2 biological standard deviations of the normal population mean. An experimental result of 10.8 mg/100 ml for calcium in the above example would have been +3.5 $S_b$, while after combination with the prior distribution the posterior mean was only +2.4 $S_b$ from the population mean $M_b$.

The advantage of the proposed model lies in the recognition of the uncertainties in an experimentally determined value because of methodological errors. The combination of the information from the laboratory measurement with the prior distribution, as expressed by the normal distribution, exhibits a trend toward the mean. In theory, a correction toward the mean is justified only if the concentration of the measured substance is not influenced by disease. In practice, the biological range (mean ±2 $S_b$) may be rightfully used as a prior in applying Bayesian statistics, because the overwhelming majority of specimens examined in the clinical laboratory are normal, particularly if they are part of a screening battery of tests. Nevertheless, most clinical laboratories will have a small but significant number of patients for whom an abnormal result has been anticipated clinically and can be confirmed in the laboratory. In these patients, the prior assumption is an abnormal population distribution. Unfortunately, distribution curves of abnormal results for any constituent in a sick population are ill-defined at best and cannot be used to calculate a posterior distribution. This limitation may be negligible at pathognomonic values if they are far outside of the range of the normal reference population. At borderline values, usually up to 3–5 standard deviations below or above the mean, the trend toward the mean (normality) may generally be considered to be a correction in the right direction, provided no other hard evidence makes the assumption of a pathological process likely. Such hard evidence, besides clinical facts, may be a small but significant simultaneous increase in blood urea nitrogen, creatinine, and uric acid in the case of renal disease, or a low phosphate concentration in combination with an elevated calcium value in the case of hyperparathyroidism. In the past, one of the problems encountered in developing computer programs for pattern recognition was that the error in measurement, which at times may be quite significant, was not considered. Here the application of Bayesian statistics may be helpful in improving the probability statements of pattern recognition.

We are indebted to Dr. C. Lam for his help in setting up the Bayesian model for the application presented here, and to Drs. F. Crump, G. F. Grannis, and J. A. Lott for their helpful suggestions in the formulation of this manuscript.

References