Simplified Approach to Confidence Limits in Radioimmunoassay

Robert C. Baxter

A simple method of calculating confidence limits for radioimmunoassay data is presented. The method involves the use of the within-assay variation in dose estimate of three routine quality-control specimens, measured in repeated assays, to estimate the confidence limits for results on unknown samples. Results for control specimens are combined by calculating the unique quadratic curve fitting a graph of within-assay standard deviation vs mean value for each control. This method requires no special data accumulation or advanced computing equipment. For cortisol, lutropin, and thyroxine radioimmunoassays, confidence limits calculated in this way have been compared with those calculated from the variance of the response variable "B/B0" in repeated standard curves. Both methods agree well with actual limits observed when plasma pools containing a wide range of hormone concentrations are assayed repeatedly.

Additional Keyphrases: statistics • cortisol • lutropin • thyroxine • quality control

Rodbard (1) and Rodbard and Frazier (2) have described the relationship between the response variable in radioimmunoassay, Y, and its variance, by the general equation

\[ \text{var}(Y) = a_0 + a_1Y + a_2Y^2 \]  

Note: \( Y = \frac{B/B_0}{(B - \text{NSB})/(B_0 - \text{NSB})} \), where \( B \) = counts bound at a given antigen dose, \( B_0 \) = counts bound at zero dose, and NSB = counts bound nonspecifically.

When dose-response data are linearized by using the logit-log method (2-4), the variance of log Y can be approximated by \( \text{var} \) (log Y) = \( \text{var} \) (Y)/\( Y^2(1-Y)^2 \). I have used Rodbard's model (generally with \( a_2 = 0 \)) to calculate confidence limits for routine radioimmunoassay results. Values for \( a_0 \), \( a_1 \), and \( a_2 \) have been obtained empirically for each assay by plotting var (Y), the within-assay variance (obtained from replicates at each dose concentration in a standard curve, then pooled for several standard curves), against the value of Y for each standard.

I observed that confidence limits calculated by analysis of standard-curve performance gave an excellent prediction of the actual performance of plasma samples assayed repeatedly as routine quality-control specimens. Thus it seemed likely that the confidence limits of assay results on routine specimens could be estimated from the performance of quality-control samples in multiple assays. This approach was first described by Rodbard et al. in 1968 (5). Several methods of combining results from quality-control samples have been suggested (1, 6, 7). In this report, I demonstrate that quality-control data can conveniently be combined by fitting a quadratic curve to a graph of within-assay standard deviation vs mean value for each of three samples assayed repeatedly. I compare confidence limits calculated in this way with those derived from standard curve performance for cortisol, thyroxine, and lutropin (luteinizing hormone) radioimmunoassays.

Materials and Methods

Hormone Measurements

Plasma cortisol and lutropin were measured by using Cortisol Premix RIA and LH RIA kits, respectively (Diagnostic Products Corp., Los Angeles, CA 90064). Plasma thyroxine was measured by using Gammacoat T4 RIA Kita (Clinical Assays, Cambridge, MA 02139). The number of standards supplied with these kits is five, six, and six, respectively. Other hormones were measured by using conventional double-antibody radioimmunoassays. Hormone dose estimates in unknown samples were based on logit-log transformed standard curves.

Calculation of Confidence Limits

Method 1: Standard curve method. Values of \( a_0 \), \( a_1 \), and (in some cases) \( a_2 \) (equation 1) were determined by plotting the within-assay variance of Y vs the mean value of Y for each standard, data being pooled from 10 standard curves with use of normal analysis of variance methods (8). Standards were run in triplicate, yielding 20 degrees of freedom for the determination of each point. This is the "indirect approach" described by Rodbard (7). For unknowns, the variance of the logarithm of hormone dose estimate was then calculated as described by Chang et al. (4). Ninety-five percent confidence limits of dose estimates were calculated as exp (z ± 2.0 \( s_e \)), where \( z = \ln \) (dose estimate) and \( s_e \) = the standard deviation of z. Limits calculated in this way are not symmetrical about the mean value.

Method 2: Laboratory control method. For each hormone, three quality-control plasma specimens, containing low, medium, and high concentrations of hormone in the assay range, were run in duplicate in 20 routine assays. The relationship between the mean dose estimate in each specimen (x) and its within-assay standard deviation based on 20 degrees of freedom (\( s_x \)) was approximated by the unique quadratic curve fitting a graph of \( s_x \) vs x for the three specimens, a simple form of "precision profile" (9). Although this procedure can be performed mathematically, without actually plotting data points, such curves are valuable in assessing assay performance, and so it may be useful routinely to plot these data, either manually or by computer. For unknowns, dose estimates (x) were obtained from log-log standard curves and values of \( s_x \) were calculated from the fitted quadratic equation. Ninety-five percent confidence limits were calculated as \( x \pm 2.0 \times s_x \).

Method 3: Experimental confidence limits. Ten plasma pools with hormone values spanning the expected clinical range were assayed in triplicate in 10 successive assays, yielding values for within-assay standard deviation (s) based on 20 degrees of freedom. Ninety-five percent confidence limits are mean values ±2.0 s.

Results

Figure 1 shows the relationship between var (Y) and Y for 10 replicate cortisol, lutropin, and thyroxine standard curves. All data sets could be approximated by linear relationships (i.e., \( a_2 = 0 \)), with correlation coefficients of 0.870, 0.757, and 0.911, respectively. Typical values of \( a_0 \) and \( a_1 \) for these and some other hormones routinely measured in this laboratory are shown in Table 1. Values of \( a_0 \) are relatively small, confirming the approximately proportionality between var (Y)
Fig. 1. Relationship between \( \text{var}(Y) \) and \( Y \) (the response variable)\(^\dagger\) \(\text{B}_1\text{B}_0^{-1}\) for (a) cortisol, (b) lutropin, and (c) thyroxine radioimmunoassays

Each line represents pooled data from standards run in 10 standard curves, as described in the Methods section.

Fig. 2. Relationship between within-assay standard deviation (SD) and mean dose estimate for three quality-control specimens in (a) cortisol, (b) lutropin, and (c) thyroxine radioimmunoassays. Shaded areas represent 95\% "confidence envelopes" for estimates of SD, calculated from the chi-square distribution with 20 degrees of freedom. For mean dose estimates, 95\% confidence limits are contained within data points and Y previously reported (10). It is important to note, however, that the negative values of \( \text{B}_0 \) seen in some assays demonstrate the limitations of using a linear relationship between var \( (Y) \) and \( Y \), and could be avoided by the inclusion of the term \( \text{B}_2 \).

For a variety of hormone assays, 95\% confidence limits calculated by Method 1 agreed well with the actual within-assay variation that occurred when routine quality-control specimens were assayed repeatedly. Therefore I tested whether the variation of control specimens (Method 2) could, as previously suggested (5), be of value in predicting the variation of unknowns, using three assays run on a daily basis, for a steroid hormone, a thyroid hormone, and a peptide hormone.

Figure 2a shows, for cortisol, the relationship between mean values of three quality-control specimens (\( x \)) and their within-assay standard deviation (\( s_x \)). The quadratic equation is: \( s_x = a + bx + cx^2 \), where \( a = 4.912, b = 1.675 \times 10^{-2}, \) and \( c = 1.568 \times 10^{-6} \) (Table 2). Thus the estimated within-assay standard deviation of a cortisol result of 500 nmol/L is 4.91 + 8.38 + 3.92 = 17.2 nmol/L. The 95\% confidence limits of this estimate are 13.2-24.8. Similar curves for lutropin and thyroxine are shown in Figures 2b and 2c. Values of \( a, b, \) and \( c \) for the three assays are given in Table 2.\(^\dagger\)

For each assay, 10 plasma pools containing hormone concentrations spanning the expected range for unknown samples were run in triplicate in 10 assays. In Table 3, 95\% confidence limits for the cortisol assay determined by using these pools (Method 3) are compared with those calculated by using Methods 1 and 2. Figure 3 shows, for each assay, the mean dose estimate for each pool plotted vs its within-assay CV. These experimentally obtained precision profiles (Method 3) are compared on the same graph with theoretical profiles obtained either by Method 1 or Method 2. The three methods give results that agree closely over most of the analytical range, although some discrepancies are obvious at extremely low values.

Discussion

The statistical techniques based on the variance of the response variable \( Y \), originated by Rodbard and coworkers, allow the accurate monitoring and prediction of radioimmunoassay performance and are of immense value to practicing assayists. They are, however, most easily run with advanced computing facilities to handle large accumulations of data, although a program has recently become available for the hand-held T-59 calculator and PC-100A printer (Texas Instruments, Houston, TX 77001).
In the alternative method described in this report (Method 2), the only data accumulated are the assay-to-assay values of three quality-control specimens, values normally accumulated as a matter of routine assay control. Calculations of means and within-assay standard deviations, and the quadratic curve joining them, require only simple calculating facilities.

The use of variation in control samples to assess the error in estimates of unknowns is not in itself new, having first been suggested over a decade ago (5). However, the need has remained for a simple method of combining information derived from control samples covering the expected concentration range of analyte in unknown samples. Rodbard (11) used a linear relationship between $s^2$ and $x$; implying a constant CV over the expected concentration range, but suggested that a curvilinear relationship might be more appropriate. I have only rarely observed constancy of CV, and in this report I have demonstrated empirically that a quadratic relationship between three control samples allows excellent prediction of the errors of unknown estimates. The cases of constant CV are particular cases of my more general description. It is important to note that the parameters $a$, $b$, and $c$, defining the quadratic curves, are meaningful only for a particular assay in an individual laboratory, and must be calculated separately for each assay in each laboratory.

As with any method, the use of Method 2 to estimate confidence limits of unknowns is subject to experimental error. This can be minimized by increasing the degrees of freedom on which calculations of the SD of each control specimen are based. For example, it may be calculated by reference to chi-square tables that the 95% "confidence envelopes" shown in Figure 2, based on 20 degrees of freedom (i.e., 20 estimates in duplicate) would be widened by 62% if only 10 duplicate measurements of each control sample had been made (10 df) but narrowed by 50% if 35 measurements in triplicate were made (70 df).

Method 2 may also be unreliable for extremely low or high samples, if the three quality-control specimens are not chosen adequately to span the analytical range. This is seen particularly in Table 3 for the cortisol specimen with mean value 26.7 nmol/L (normal range: 150–600 nmol/L), where in fact neither Methods 1 nor 2 accurately predicted the confidence limits.

Despite these limitations, it is clear from Figure 3 that in general both Methods 1 and 2 predict the variation of unknown specimens accurately. Thus, in a sense, these methods are validated by each other and could both profitably be used by practicing assayists.

In conclusion, the use of a quadratic relationship to fit a graph of within-assay standard deviation vs mean dose estimate for three quality-control specimens provides a simple and accurate method for estimating confidence limits for unknown samples in radioimmunoassay.

The expert assistance of R. Ristuccia, D. Chan, and R. Grigg in performing radioimmunoassays is gratefully acknowledged. I thank Dr. D. Rodbard for his valuable comments on the manuscript.

References