LETTERS TO THE EDITOR

Re: Evaluation and Comparison of Two Fully Automated Radioassay Systems with Distinctly Different Modes of Analysis

Concerning the article by I. W. Chen et al. (*1*), which I read with much interest, I make the following remarks.

From Table 1, which shows the intra- and interassay precision at different concentrations, it appears that the interassay variability is often less than the intra-assay variability. This finding is rather puzzling. Indeed, as explained by D. Rodbard (*2*), the observed interassay variance, S^2, equals

$$S^2 = S_b^2 + S_w^2/r, \quad (1)$$

where S_b^2 would be the "intrinsic" component of the interassay variance if we had an infinite number of replicates in each assay; S_w^2 is the component of variation due to measurement error within an assay for a single tube (i.e., the intra-assay variance); and r is the number of replicates in each assay.

Since the intra-assay variance bears on the assay of a single tube, it is advisable to express the interassay variance in the same way. It suffices to solve for S_b^2 in Eq. 1, where S^2, S_w^2, and r are known, and to recompute S^2 for r = 1 using the same equation. Otherwise, when quantifying the interassay variability by the variance of the assay means, as I presume was done here, one must state the number of replicates in each run lest the resulting figure be meaningless. For example, let us suppose that the results bear on ten replicates per run. Without any change in the variability of the system, and computing the coefficient of variation (CV%) the way it is supposedly done, the use of only two duplicates in each run should result in a figure of about 6.23% CV for the interassay variability, and even of 8.32% CV if r = 1.

It is regrettable in a study comparing the performances of assay systems, where precision deserves most careful attention, that the method and computations used are not presented. If, by chance, the number of replicates used with the different assay systems (batch, sequential, and manual) was not the same, this would invalidate the comparison as far as interassay precision is concerned.

Using the method described by Rodbard (*2*), we measured the precision of the T_4 assay on the sequential system.* Commercial control sera† at three dose levels were assayed in triplicate in ten runs. The results are expressed as the coefficient of variation (%) to be expected for r = 1 (Table 1).

As can be seen, the intra-assay precision is in agreement with the results in the article, whereas the interassay precision is clearly much less.

TABLE 1. PRECISION EXPRESSED AS WITHIN-ASSAY (INTRA) AND BETWEEN-ASSAY (INTER) COEFFICIENTS OF VARIATION

Level μg/dl	CV % (r = 1) Intra	Inter
4.0	7.8	12.0
7.6	4.6	8.6
19.0	5.6	8.2

FOOTNOTES

* Aria II, Becton Dickinson Laboratory Systems.
† Hyland Diagnostics, Division of Travenol Laboratories, Inc.

P. DEVOS
University of Leuven
Medical School
Leuven, Belgium

REFERENCES

1. CHEN IW, MAXON HR, HEMINGER LA, et al: Evaluation and comparison of two fully automated radioassay systems with distinctly different modes of analysis. *J Nucl Med* 21:1162-1168, 1980
2. RODBARD D: Statistical quality control and routine data processing for radioimmunoassays and immunoradiometric assays. *Clin Chem* 20:1255-1270, 1974

Reply

Dr. Devos has commented on our paper dealing with the analysis of precision in automated radioimmunoassay systems (*1*). The percent coefficients of variation are given for duplicates within assays (CVw) and for singles between assays (CVb). Coefficients of variation were determined by analyzing each control sample in duplicate in each of several assays in all three systems studied. The variance between analyses was calculated according to standard analysis of variance without determining the components of variance of means (*2*). We prefer to estimate the total variance of means (variance of means alone plus variance of measurement error) when determining whether the variance between means is greater than that of analysis despite a limited number of replicates. Accordingly, the CVb reported in our paper represents the variance for means of duplicates, whereas CVw is for the single tube.

A CVw greater than CVb is unusual but may occur when CVb is small, as in the case of the T_4 assay in the sequential system, or when CVw is large because the number of replicates assayed in each run is small. The conditions of the assays in question lead to these results. We used duplicates in our studies instead of higher replication since the precision of the assay warrants this, and because we wanted to evaluate the precision of the automated radioassay systems under conditions as close to the routine laboratory conditions as possible. The CVw listed in Table 1 of our paper would apply to a single measurement of an unknown falling in the same general region on the dose-response curve. If the unknowns are analyzed in duplicate, as is the case in our laboratory, the values of CVw should be divided by $\sqrt{2}$. We, however, customarily express our results in terms of singles. In our laboratory different numbers of replicates may be used, and expression of CVw for singles provides a satisfactory estimate of the reproducibility of the assay for our purposes.

We have since accumulated more data on the sequential automated system, as shown in the accompanying table. Unfortunately, we no longer keep the other two systems in our laboratory, and thus

TABLE 1. WITHIN-ASSAY (% CVw) AND BETWEEN-ASSAY (% CVb) COEFFICIENTS OF VARIATION FOR DUPLICATE RUN (MEAN ± s.d.)

Control sera	T_4		T_3		Digoxin	
	% CVw	% CVb	% CVw	%CVb	% CVw	%CVb
Low	4.4 ± 1.2	5.3 ± 1.6	5.5 ± 0.6	8.3 ± 1.4	3.2 ± 1.0	4.3 ± 1.7
Medium	3.3 ± 0.8	3.5 ± 1.0	3.8 ± 1.1	6.0 ± 1.1	2.8 ± 0.4	3.9 ± 0.7
High	3.5 ± 0.7	4.6 ± 1.5	3.3 ± 0.7	5.4 ± 1.5	3.8 ± 0.8	5.4 ± 1.8

no additional data can be given for making comparisons. The data shown represent mean values obtained from the last 10-mo period, using the same three control sera as in our previous study. The CVw and CVb were determined every month for digoxin and T_4 assays (on the average, about 25 assays per month) and every 2 mo for T_3 assays (about 24 assays per 2 mo), and they represent the variation one should expect if this instrument is used to run assays in duplicate.

ACKNOWLEDGMENT

We thank Dr. Devos for giving us this opportunity to clarify our study protocol.

I-WEN CHEN
HARRY R. MAXON
University of Cincinnati College of Medicine
Cincinnati, Ohio

REFERENCES

1. CHEN IW, MAXON HR, HEMINGER LA, et al: Evaluation and comparison of two fully automated radioassay systems with distinctly different modes of analysis. *J Nucl Med* 21:1162–1168, 1980
2. SNEDECOR GW: *Statistical Methods.* 4th ed. Ames, IA, Iowa State College Press, 1946, p 228

On the Improvement of Analyses of Xenon-133 Lung Washin and Washout Curves

The recent article by van der Mark et al. (*1*) claims to describe an improved method for analyzing Xe-133 lung washin and washout curves. This method involves a "pragmatic" approach whereby a human lung histogram is described as a sum of a single exponential function and an approximation of a quadratic Taylor's series expansion for residual exponentials (*2*). Justification of the use of Taylor's series is based upon the difficulty of fitting more than one exponential function to physiological curves obtained from a region of interest. It is our opinion that the improvement of fitting seen with their exponential plus polynomial model is more directly attributable to the increased number of parameters available in that case. The authors, in fact, use seven adjustable constants in fitting their pragmatic picture to both simulated and patient data. These constants include the time the fit began (T_0), the time of equilibrium concentration (T), and the required three Taylor's series multipliers: a_0, a_1, and a_2. In addition, equilibrium count rate (N_∞) and the exponential rate constant (k) were also available for manipulation. These seven values were determined simultaneously by a minimization of the reduced χ^2 statistic.

By way of contrast, three of the four alternative forms of analysis described in the article depend upon only a single parameter. Moreover, this parameter is not determined by a least-squares algorithm but by a simplistic analysis of the pulmonary histograms. For example, in the $t_{1/2}$ method, the authors measure the time to reach $\frac{1}{2} N_\infty$. This result is inverted and multiplied by ln2 to define a "rate constant." Understandably, the result, in the case of the simulation-curve comparison, is a relatively poor fit compared with that of the pragmatic seven-parameter model.

A second example of the type of simplistic alternative analysis cited by the authors can be seen in their consideration of the moments method (*3*). They state that a first temporal moment of clearance times is given by

$$M_1 = \frac{\int tN(t)dt}{\int N(t)dt},$$

where N(t) is the histogram. Since no limits of integration are included, the exact meaning of the equality is uncertain. This calculation appears, however, to be based upon the assumption that N(t) is the probabilistic distribution of Xe-133 lung clearance times. Because of the inhalation of xenon over an extended period of several minutes, N(t) actually represents the net result, in a given region of interest, of concurrent inflow and outflow. It is most certainly not the distribution of clearance times—unless a very sharp bolus injection of xenon was delivered to the region in question. Mathematically, N(t) is better represented as the convolution of the distribution of transit times with the inhaled curve of xenon presented to the lung volume. Finally in this example, the authors assume that the inverse of M_1 is a rate constant comparable to the (k) rate constant of their seven-parameter model. This last step is, again, an assumption that need not be the case (*4*). As a consequence the moments method appears to be the least desirable of any alternative model investigated by van der Mark et al. (*1*).

Similar ad hoc manipulations occur for the height-over-area method and when a single exponential curve is fitted, by least squares, to the washin segment of the histogram. This last alternative is merely a subset of their pragmatic approach and thus expected to be a relatively worse approximation to the fitting of any regional curve.

Minimally, the authors should use somewhat more realistic models having a similar number of parameters as their proposed functional representation. Second, the fitting should be done, in all cases, with the same statistical optimization technique. One could then, at the termination of the algorithm, compare the goodness of fit and draw more valid conclusions. While their model is clearly superior to simplistic analyses, it is not necessarily an improvement in any statistical sense of the word. For example, we are not given reduced χ^2 values for any of the four alternative representations of the histograms.

A question can also be raised as to the ultimate interpretation of their model. While most observers would agree that multiple exponentials are difficult to determine by any algorithm, at least the comprehension of the resultant rate constants is somewhat more straightforward than that of a set of Taylor's series coefficients. As the authors remark, their series expansion is, effectively, a sum over all other exponentials not explicitly considered in their analysis. Thus in their case the description of the patient's histogram is being forced *a priori* into a single-exponential format. The lung volume under observation, however, may contain more than one population of alveoli, so that more than one exponential function would be required in a first-order compartment model. This type of behavior is, in fact, not unexpected in victims of