

J. Clin. Chem. Clin. Biochem.
Vol. 24, 1986, pp. 1023–1028

© 1986 Walter de Gruyter & Co.
Berlin · New York

Evaluation of Radioimmunoassays: Comparison of Dose Interpolation Calculations by Four Parameter Logistic and Spline functions¹⁾

By M. Kraupp, R. März, E. Legenstein, B. Knerer and T. Szekeres

Institut für Medizinische Chemie, Universität Wien, Vienna, Austria

(Received July 31, 1985/July 15, 1986)

Dedicated to Prof. Erich Kaiser on his 60th birthday

Summary: Two computerized methods for dose interpolation calculation were compared. Generated data sets with a known coefficient of variation as well as laboratory RIA data were analysed. The four parameter logistic method, which is based on an approximation of the mass action law, performed better than the Spline method, a procedure which makes no a priori assumptions about the data. Correct weighting of the data was important for obtaining satisfactory fits. The determination of the response error relationship proved to be the most satisfactory approach in obtaining suitable weighting factors.

Auswertung von Radioimmunoassays. Vergleich der Interpolation von Standardkurven mittels der vier Parameter Logistik und der Spline Funktion

Zusammenfassung: Zwei Methoden der computerisierten Dosisinterpolation wurden verglichen. Sowohl computergenerierte Daten mit einem vorgegebenen Variationskoeffizienten als auch Labordaten wurden ausgewertet. Die vier Parameter Logistik Methode, ein Modell, das auf einer Annäherung des Massenwirkungsgesetzes beruht, ergab bessere Resultate als die Spline Funktion, eine Methode ohne Annahme der Datenstruktur. Das richtige Gewichten der Daten war wichtig, um den Kurvenverlauf zu bestimmen. Berechnung des Verhaltens der experimentellen Varianz im Bereich der Standardkurve (response error relation) war die beste Methode einer zuverlässigen Berechnung der Gewichtungsfaktoren.

Introduction

A linear signal to dose relationship provides an ideal basis for quantitative analysis, since the results obtained with standard concentrations are easy to interpolate. However, for many assays routinely used in the clinical laboratory, the signal to dose relationship is much more complex. Until quite recently a combination of data transformation and graphic methods had to be employed to convert the assay results into the desired parameter, e. g. concentration.

These methods can be tedious, error prone, inherently inaccurate, and subject to interpreter bias. But worse, after data transformation the unavoidable experimental error of standards, unknowns, and controls alike will distort the result. Although this bias is systematic and in principle predictable, even the best methods can only minimize, but not eliminate it. The continuing popularity of the *Lineweaver-Burk* transformation proves that in fact the problem is often simply ignored.

Ready access to computing facilities has considerably broadened the options available to most laboratories. Many computerized methods have been developed

¹⁾ Supported by grant 31/83 of the Anton Dreher-Gedächtnisschenkung für Medizinische Forschung.

for calculating dose interpolations. However, this convenience carries a high price-tag, since the investigator often relegates the critical evaluation of the assay data to the computer.

Many computerized methods for dose interpolation have been developed for radioimmunoassays and related procedures. These methods can be roughly separated into two categories:

1. General purpose fitting methods. For these calculations no a priori assumptions concerning the shape of the curve to be fitted is made. Examples include the fitting of polynomials (e. g. cubic) or Spline functions (1, 2).

2. Model-based methods. Here multi-parameter expressions derived from the mass action law are used to generate the shape of the curve. The actual fit is calculated by a regression method which optimizes each parameter individually. This method can be simplified and generalized by using instead an approximation of the mass action law, e. g. the four parameter logistic (3, 4).

Methods from both categories are in fact in use for automatic data reduction. This paper examines both simulated and laboratory data of radioimmunoassays. The performance of one example of each of the two types of data fitting algorithms are evaluated. Finally the relationship between quality of assay data and satisfactory performance of each of these computerized data interpolation methods is analysed.

Materials and Methods

Radioimmunoassays were performed with commercial kits. The Insulin-MAIA kit of Sero Diagnostic (Chavannes-de-Bogis, Switzerland) was used to measure serum insulin levels. Neopterin concentrations in serum and urine were assayed using the appropriate kits from Henning (West Berlin, FRG). Replicate serum neopterin samples of known concentrations were used in testing the performance of the fitting algorithms.

Two procedures were used to fit standard curves and calculate dose interpolations. The first of these employed Spline functions to describe the standard curve mathematically. This method can be regarded as a computerized version of the "French curve" or mechanical "flexicurve".

The standard curve is subdivided into local regions between two successive standards, which are individually described by a polynomial. The polynomials are then spliced together by "knots" thus giving one knot for each standard. At these knots the derivatives (slopes) of two joining polynomials have to be identical to ensure a smooth and continuous curve (1, 2). The automatic program used in this study tries to find a curve with either one or no turning points and no maxima or minima, by varying data weighting and smoothing factor.

In the second procedure the weighted four parameter logistic model of *Rodbard* (3, 4) was used to describe the standard curve:

$$B' = ((a-d)/(1 + (X/c)^b)) + d.$$

The terminology is the one used by *Rodbard* (3, 4), where:

- a) is counts/min bound to the antibody in the absence of unlabeled substrate (B_0)
- b) is an allosteric factor determining the shape of the curve
- c) is the midpoint of the curve (ED_{50})
- d) is the non-specific binding (Nsb)
- B is the measured response (counts/min)
- X is the dose.

The model is derived from the widely used logit log transformation of RIA data. This transformation of response data requires accurate estimates of zero and infinite dose response, in order to obtain a linear calibration curve. Since this requirement is unrealistic, the four parameter logistic model is a better approach. Here, zero dose response and non-specific binding are included as parameters and are adjusted by the fitting algorithm.

Fitting to the standards was carried out with the weighted non-linear regression algorithm of *Marquart* (5), which minimizes the weighted sum of squared deviations of the measured from the calculated response. In addition, estimates of the standard error of the curve parameters are computed.

Three different methods were used to assign weights to each point of the standard curve:

- 1) each point was weighted equally,
- 2) the weights were calculated using the relationship:

$$W_i = 2/(B - B_i)^2$$

W_i is the weight assigned to an individual data point; B and B_i (the "response") are duplicate determinations of the counts/min of labeled substrate bound to the antibody in the presence of unlabeled substrate (the "dose")

- 3) the weights were estimated by analysing the behaviour of the standard deviation of the measurements over the whole response range. The response error relationship was evaluated as described in the literature (6, 7). Assay results for standards as well as unknowns from several experiments were pooled. The difference between duplicate estimates divided by $\sqrt{2}$ was plotted against the mean of B/B_0 . To these data of the response error relationship a quadratic polynomial:

$$S_y = a_0 + a_1 \cdot Y + a_2 \cdot Y \cdot Y \quad (Y = B/B_0)$$

or, if appropriate, a straight line through the origin:

$$S_y = a_1 \cdot Y \quad (Y = B/B_0)$$

was fitted. The weights were then calculated as:

$$W_i = 1 / S_y \cdot S_y$$

In another approach these data were grouped (binned). The median in each group was determined and plotted against the mean B/B_0 in this group. Again a quadratic polynomial was fitted to these data with a weighting equal to the number of data points in each group.

In the special case of Monte Carlo simulations method 3 could be implemented in a simplified form, since the response error relationship is linear. Here the first and the third term of the polynomial (a_0, a_2) were set to zero. The formula for calculating the weights thus became:

$$W_i = 1/(CV \cdot B)^2$$

where CV = coefficient of variation.

Monte Carlo experiments were carried out as follows: A standard curve was calculated with the four parameter logistic formula using an arbitrary set of parameters ($a = 19870, b = 1, c = 10, d = 0$). A set of 8 duplicate "standards" and another

set of 99 "unknowns" covering the whole response range of the standard curve were generated and the corresponding true doses calculated. The "standards" were randomized to give them a normal distribution around the standard curve with the variance proportional to the measured response (see preceding paragraph). The coefficient of variation used in randomizing the data was set between 1 and 10%.

The randomized points served as raw data for calculating the shape of the standard curves with the two different fitting algorithms. The concentration of an "unknown" was computed from the true dose by interpolating these standard curves. The interpolated values were then compared with the true ones to evaluate the quality of the fitting methods.

Radioactivity was determined using either a Beckman LS 7500 beta or a LKB 1275 Minigamma gamma counter. Programs were written and implemented on a Hewlett Packard 9826 computer in extended BASIC 1.0. Program listings are available upon request.

Results and Discussion

Both of the programmed fitting routines were checked by applying them to published RIA data (3). For unknowns within or close to the concentration range of the standard curve, both the four parameter logistic and the Spline interpolation gave results almost identical to the literature values (tab. 1, lines 1 and 2). These data are, however, based on an excellent standard curve with very little error, so that any method of interpolation would be expected to work. There are also too few unknowns to allow any detailed comparison of the fitting methods.

Tab. 1. Interpolation of published RIA data.

	Dose (pg)		
	Literature (Finney, 1976 (3))	Four parameter Logistic	Spline
1	57.4	58	60
2	217	216	218
3	1750	2073	1920
4	14100	6900	3796

The standard curve is based on quadruplicate determinations for six concentrations ranging from 6.25 to 200 pg (3).

Two of the unknowns (tab. 1, lines 3 and 4) lie far outside the concentration range of the standard curve; thus extrapolation rather than interpolation became necessary. Both the Spline and the four parameter logistic algorithm did not perform well in this region. However, since all these routines are only designed to be used for interpolation this does not argue against their utility.

To investigate the performance of both types of data fitting procedures, test data were generated using the Monte Carlo approach. Figures 1 and 2 show the result of such a simulation. The scattered "measurements" of a simulated standard curve were used as the raw data for both fitting procedures. In figure 1 the "true curve" (the same one was used in all cases before the data were scattered), as well as both fitted standard curves are shown for a coefficient of varia-

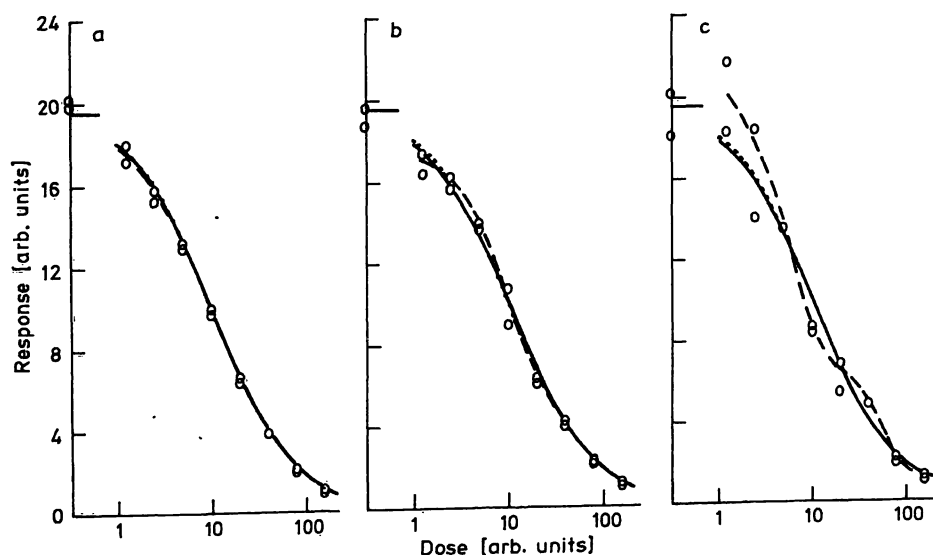


Fig. 1. Monte Carlo simulation: comparison of standard curves obtained by fitting Spline functions and the four parameter logistic model to three sets of data randomized with increasing coefficients of variation (a: 2%, b: 5%, c: 10%).

Y-Axis: response in arbitrary units
X-Axis: dose in arbitrary units on a log scale
(.....) ideal standard curve
(——) four parameter logistic fit
(---) Spline fit

tion of 2, 5, and 10% (panels a–c). Since the error envelope is known for these data, weights were calculated using method 3, arbitrarily setting the coefficient of variation to 4%.

In all three panels the line representing the four parameter logistic method seems virtually superimposed on the theoretical curve, indicating a perfect fit. The Spline curve shows excellent agreement at low scatter, but the fits get poorer as the scatter increases.

Figure 2 shows a more revealing representation of the same data. The sample concentrations calculated from the fits of the corresponding standard curves are expressed as a fraction of the actual value and plotted against the actual concentrations. The weighted four parameter logistic fit was indeed almost perfect at low degrees of scatter, while the Spline method already resulted in some deviation in the low dose range.

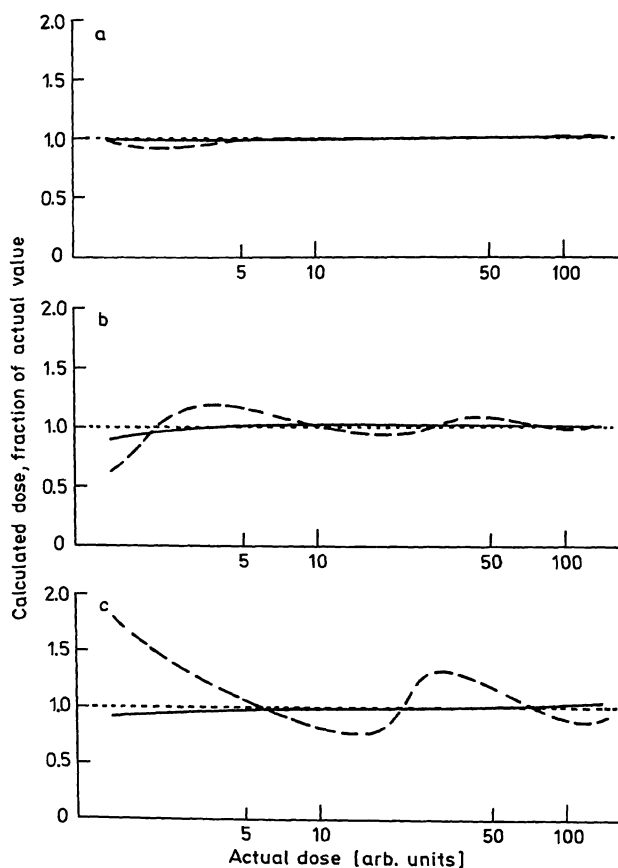


Fig. 2. Monte Carlo simulation: comparison of dose interpolations calculated from the fitted Spline functions and the four parameter logistic model of fig. 1.
Y-Axis: interpolated doses as fraction of the actual value
X-Axis: actual dose in arbitrary units on a log scale
(-----) ideal interpolation (100%)
(——) four parameter logistic interpolation
(---) Spline interpolation

The results obtained by both methods deviate more and more from the true values with increasing coefficients of variation, but the four parameter logistic method gave consistently better results. Spline interpolations displayed a strong tendency to oscillate around the true values, whereas the four parameter logistic method gave results differing somewhat systematically from the actual dose.

The behaviour of the Spline function in any one interval of the curve has some impact on how the two neighbouring portions will be fitted. Each segment will transfer its properties to some extent to the adjoining regions. In this way a poorly fitted segment may cause the whole curve to oscillate.

Weighting of the data used for the non-linear least squares regression will effect the resulting fits. In figures 3 and 4 several methods for assigning weights were compared. The data presented in figures 1 c and 2 c were evaluated by the four parameter logistic method using the three weighting procedures described in "Methods". The interpolation obtained using the response error relationship (method 3) to calculate the weights gave the best result, while assigning each point an equal weight was clearly the worst option.

The previous analysis established that the estimation of the response error relationship should indeed be used as a valid basis for further fitting operations. This relationship is a priori unknown for actual RIA assay data. The response error relationship can be estimated, however, by analysing replicate data points as described in "Methods". This procedure was applied to laboratory data from insulin and neopterin RIA determinations. These two assays employ differing experimental protocols, so the question arose as to how different their respective response error relationship would be.

Figure 5 a shows the result of such an analysis performed on the pooled data from six insulin RIA determinations. Each point represents the difference of a duplicate determination of the measured response divided by $\sqrt{2}$ plotted against the mean of the duplicate determination as a fraction of the zero-dose (B_0) response.

For panel b the same data were binned into groups of 6 and the median of each bin determined. This median was plotted against B/B_0 . Both panels a and b show that the standard deviation changes over the entire response range, a behaviour known as heteroscedasticity (4, 6). A straight line through the origin was found to describe the data adequately. The slope of the line (a_2) was estimated by weighted regression. For panel a the points were weighted

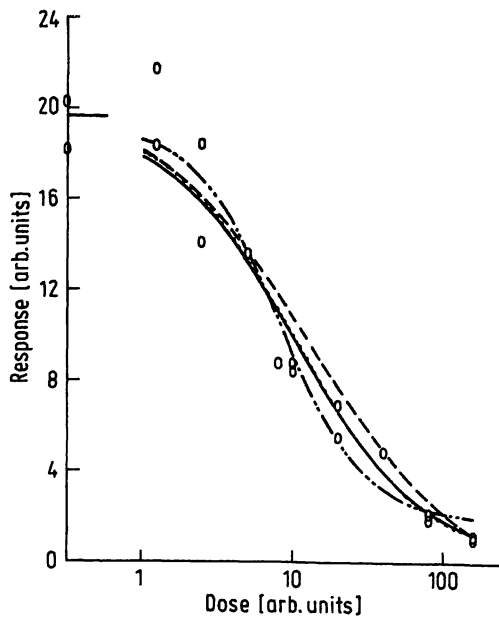


Fig. 3. Monte Carlo Simulation: comparison of 3 methods for calculating weighting factors. The four parameter logistic model was fitted to the data of fig. 1 c. Axes as in fig. 1.

- (.....) ideal standard curve
- (——) $W_i = V \cdot B$
- (- - -) $W_i = 1$
- (- · - ·) $W_i = 2/(B - B_1)^2$

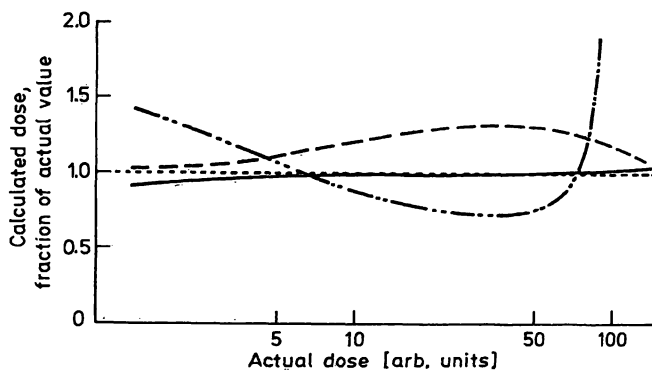


Fig. 4. Monte Carlo Simulation: effect of the 3 different weighting procedures on the accuracy of interpolations. The fitted standard curves of figure 3 were used to calculate the unknowns. Axes as in fig. 2. Line symbols and weighting as in fig. 3.

equally, whereas for b weights were equal to the number of points in each bin. The slopes obtained by both methods are in good agreement; a: 0.099, b: 0.080. These estimated coefficients were used to calculate weights according to method 3.

Figure 5c shows the same type of analysis as performed in 5a but this time for neopterin data. A slope of 0.108 was obtained. Binning of the data gave a similar result (data not shown).

Performance of the interpolation methods was also tested by analysing serum neopterin samples of

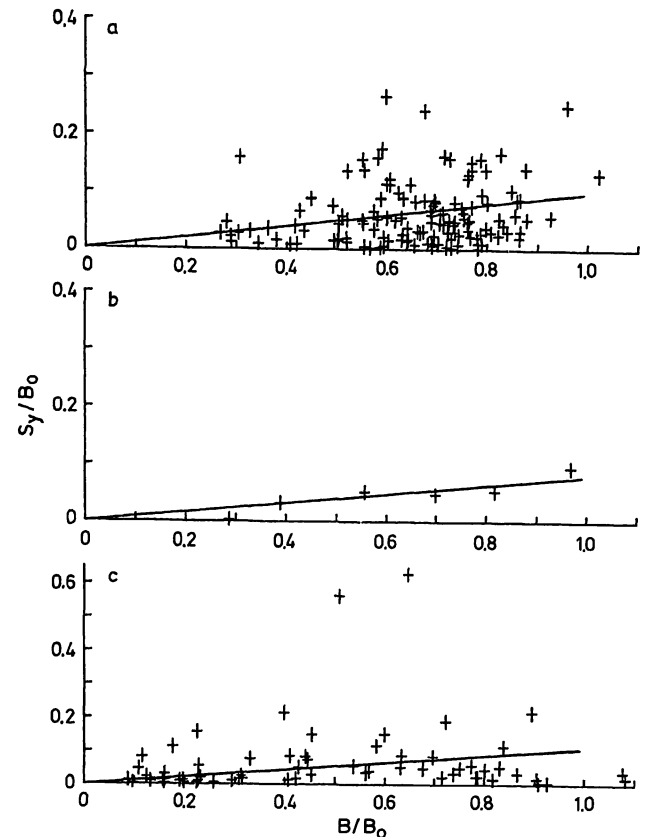


Fig. 5. Analysis of the response error relationship for insulin and neopterin RIAs. All data are expressed as fractions of zero dose response (B_0). A straight line through the origin was fitted to all data.

- Panel a: Standards and unknowns of 6 insulin RIA sets
- Panel b: Data of panel a binned into 6 groups
- Panel c: Standards of 8 neopterin RIA sets
- Y-Axis: Difference of duplicate determinations divided by $\sqrt{2}$
- X-Axis: Mean of the corresponding duplicate determination.

known concentrations (tab. 2). The mean found for each sample group was in good agreement with the actual value except at low concentrations.

A pairing design test which compared the actual and the found concentration gave $t = 0.87$ for the four parameter logistic and $t = 1.43$ for the Spline fits (255 degrees of freedom). By this criterion the four parameter logistic method proved superior since the null hypothesis is in fact true. The rather large values of t are the consequence of the large coefficient of variation associated with the calculated concentrations (tab. 2). The measurements themselves had low coefficients of variation but data transformation magnified the experimental errors. This problem is inherent in the method and can not be eliminated by use of the computer.

Together these results establish that the fits obtained with the Spline function are more problematic than those which make use of the four parameter logistic

Tab. 2. Analysis of serum neopterin samples with known concentrations. Results computed with the different interpolation programs as well as coefficients of variation of the measured counts/min and the calculated concentrations are given.

Concentration (nmol/l)	Concentration found mean (nmol/l)		Coefficient of variation (%)		
	Spline	Four parameter logistic	Counts/min measured	Concentration calculated	
				Spline	Four parameter logistic
2.5	2.9	2.9	4.4	77	57
5	3.2	3.4	4.5	73	50
10	10.5	9.7	5.2	23	32
20	19.5	19.8	3.2	15	14
40	40.4	39.6	3.2	10	11
80	81.6	81.2	5.8	19	26
5.3	7.1	6.4	4.4	31	31
19.3	21.1	19.9	4.7	16	18

Four standard curves with duplicate determinations for eight concentrations ranging from 1.25 to 160 nmol/l were used for data evaluation. Samples with nominal concentrations of 5.3 and 19.3 nmol/l were control sera supplied by the manufacturer of the RIA kit. All assays were performed on the same day. The number of determinations was 56 for control sera samples, 24 for all others.

method, an appropriate mathematical model to describe the dose response relationship. Nevertheless the merits of the Spline approach will be decisive in many cases: no a priori assumptions about the dose response relationship must be made and less computational effort is needed. However, for RIA assays where a mathematical model exists, the four parameter logistic approach is clearly appropriate.

Knowledge of the behaviour of the response-associated error improves the calculation of the weighting factors considerably. This in turn allows increased accuracy in the calculation of the interpolated results. Any uncertainty in the response error relationship, which always exists for laboratory data, lowers the accuracy of the interpolations.

The relative weighting of the data points is the most important factor while the absolute values are not decisive. For the four parameter logistic interpolations shown in figure 2, the correct shape of the response error relationship was used. The absolute

value of the coefficient of variation was arbitrarily set to 4%, while the true values varied from 2 to 10%. The accuracy of the interpolations was not improved significantly by using the true values (data not shown).

However, there are limitations in the obtainable accuracy of interpolations. The Monte Carlo simulations allow a prediction of the accuracy that can be expected for data interpolation, using experimental data with a certain coefficient of variation. The tolerable error will of course depend on the parameter being determined and the appropriate accuracy requirement. However, the investigator should in all cases be able to estimate the accuracy of the assay employed, and statistical analysis as presented here is a valuable tool for achieving this goal.

Acknowledgement

Henning Berlin generously supplied one serum neopterin kit free of charge.

References

1. Reinsch, C. H. (1967) *Numerische Mathematik* 10, 177–183.
2. Reinsch, C. H. (1971) *Numerische Mathematik* 16, 451–454.
3. Finney, D. J. (1976) *Biometrics* 32, 721–740.
4. Rodbard, D. & Frazier, G. R. (1975) *Meth. Enzymol.* 378, 3–22.
5. Marquart, D. W. (1963) *J. Soc. Ind. Appl. Math.* 11, 431–441.
6. Rodbard, D., Lenox, R. H., Wray, H. L. & Ramseth, D. (1976) *Clin. Chem.* 22, 350–358.
7. International Atomic Energy Agency (1981) *Programs for Data Processing in Radioimmunoassay using the HP-41 C Programmable Calculator*, Vienna, Austria.

Dr. Martin Kraupp
Institut für Medizinische Chemie
Universität Wien
Währingerstraße 10
A-1090 Vienna, Austria