JUNE 2023
ISSUE 06

In this issue:
Illustration of best-fit curve optimization in ELISA software

News bulletin of AMBI
West Bengal Chapter

The Biochemistry Chronicles

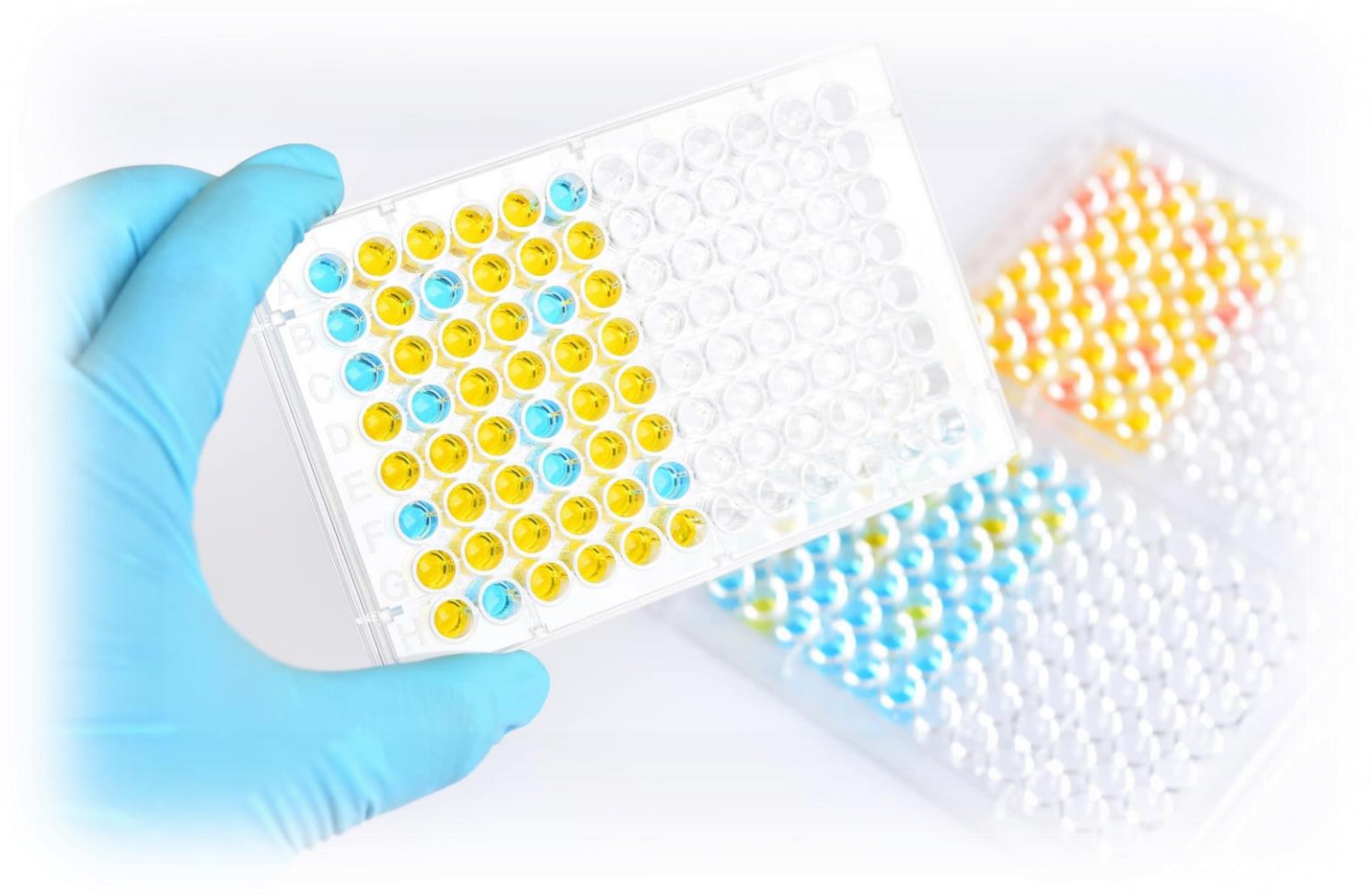


ILLUSTRATION OF BEST-FIT CURVE OPTIMIZATION IN ELISA

Traditional ELISAs are frequently used to detect and quantify specific analytes within a biological sample. These samples include serum, plasma, cell culture supernatants, and other biological matrices. In order to determine the concentration of an analyte within a sample, one must run a standard or calibration curve. The production of a standard curve requires the use of known concentrations of the analyte being assayed. Performing a quantitative immunoassay asks one, to plot an x-y plot, that shows the relationship between this standard (analyte of interest) with the readout of the assay, e.g., optical density (OD) for ELISA. The concentration of the analyte in the sample can then be calculated using the OD from the standard curve.

Before samples can be analysed, it is important to choose the best curve fit model to achieve the most accurate and reliable results.

- **Different Elisa standard curves:**
- **Linear Regression:**

Linear regression is the easiest regression model and is the basic regression analytical method for curve fitting.

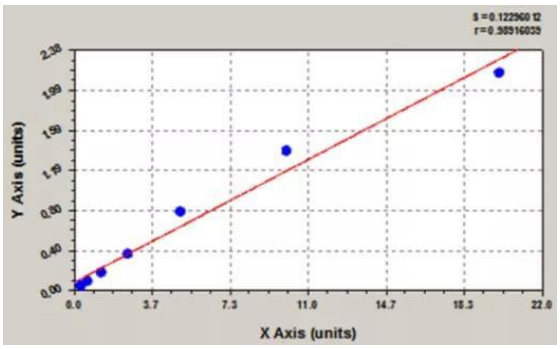


Figure 1. Fitting function equation: y=a+bx

The concentration is generally represented as x, the assay readout(OD) as y, with b referring to the slope and **a** referring to the y-intercept where x = 0.

The aim is to find values for the slope (b) and y-intercept (a) that minimize the absolute distance from the data point to the curve, also known as the “residual”.

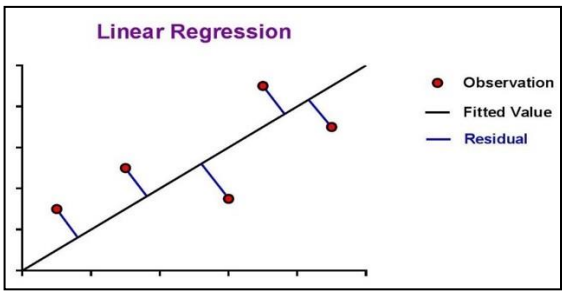


Figure 2.

The ideal assumption is that the best-fit linear curve will be a line that passes as close as possible to all data points from the standard curve. The question that arises from this is, "How is this assessed?" This is where the concept of a 'residual' is introduced. Since the best fit line will be the one that passes closest to all data points, it should seem natural that we could simply sum the residuals of all data points and the line with the lowest sum would be the best. However, there is an underlying problem here that needs to be addressed. Take an over-simplified example where we are looking at residuals from just 2 data points, A & B. Now, imagine we fit 2 linear curves to the data. The first gives residuals of A = 1 and B = 9, and the second gives A = 5 and B = 5. If we sum the residuals, both curves give the same answer of 10. This is problematic since mathematically they are

"equivalent", but clearly the second curve fits the data better as it passes closer to both data points. More simply put, 5 for each is a better fit than 1 and 9.

The solution to this issue is to square the residual values first, and then add them together. By transforming the data like this, curves with poorer fits and larger residuals will be scored higher and become less desirable. To revisit the example from above, $5^2 + 5^2 = 50$, and $1^2 + 9^2 = 82$. Rather than being mathematically equivalent, now the better fit curve has the lower sum of squared residuals.

- **Non Linear Curves:-**

⇒ **Quadratic Polynomial Regression Equation:**

Quadratic polynomial is shown in the form of parabola which opens upward or downward. In many ELISA tests, either ascending or descending order of quadratic polynomial is fitted. Due to features of the curve, the same concentration value may differently appear in the curve: no relevant OD value, one OD value, or two OD values. Thus, when using quadratic polynomial for fitting, the range of selected value should entirely fall in the ascending or descending curve. Otherwise, the good correlation coefficient is probably inconsistent with the actual value.

⇒ **Cubic Polynomial Regression Equation:**

Cubic polynomial is presented in inverted "s-shape". If the test result just falls in the ascending or descending curve, the effect is good. However, if the value range is wide, due to the fluctuation of the curvature, the simulation of cubic polynomial **may not be good**.

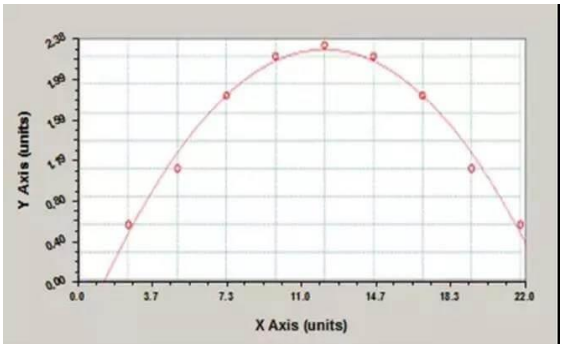


Figure 3. Fitting function equation: $y = a + b\,x + c^2$

Like quadratic equations, both the correlation coefficient of the curve and the distribution of calculated points should be concerned. Then, the ideal result can be obtained. During calculating values with the software, selectively choose relative concentration or OD value and list the result corresponding with original OD.

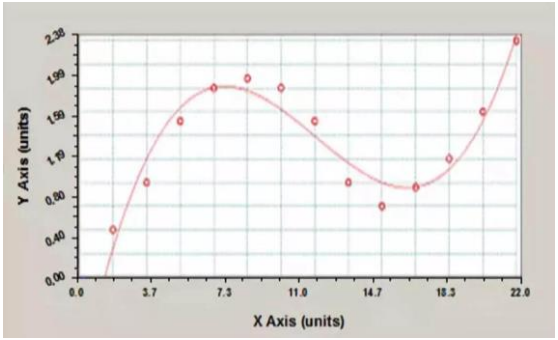


Figure 4. Fitting function equation: $y = a + b\,x + c^2 + d^3$

⇒ **Semilogarithmic Regression Equation:**

Semilogarithmic fitting refers to take logarithm of concentration value and make linear regression in appropriate OD value. The ideal result is a line under semi log

coordination, showing concentration logarithmically increases or decreases with the related variation of OD values. In other words, the change of concentration is stronger than the variation of OD values. Semilogarithmic fitting is used in ELISA test (Semi logarithm is also used for drawing in EXCEL).

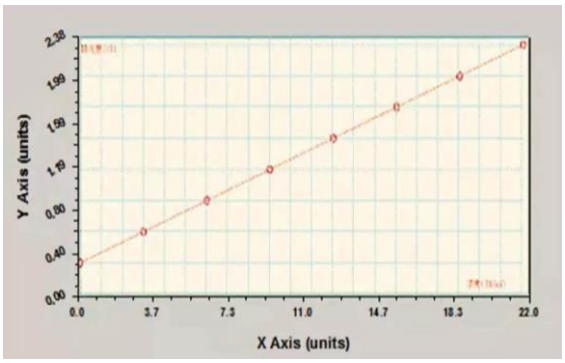


Figure 5. Fitting function equation: y = a lg(x) + b

⇒ Log-Log Regression Equation:

Log-Log fitting is similar with semilogarithmic fitting. OD values and relevant concentration values are made by the linear regression and log-log curve is drawn.

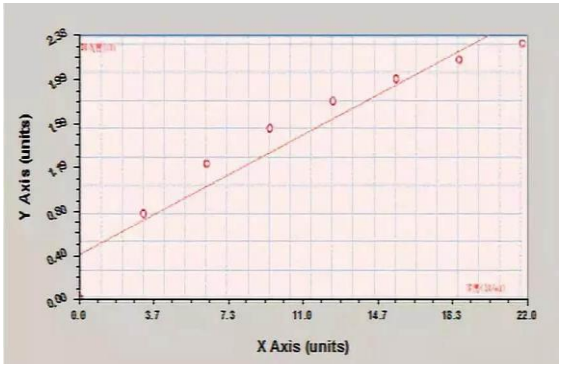


Figure 6. Fitting function equation: lg(y) = a lg(x) + b

⇒ 4-Parameter Logistic (4PL):

Immunoassay standard curves typically produce an S-shaped sigmoidal curve, which requires a different kind of mathematical modelling called logistic regression, that allows for curve fitting beyond the linear range of the curve. This new range is referred to as the logistic range, and is most simply described by a 4PL curve. This type of modelling still uses the underlying concept of summing the square of the residuals, but instead of minimizing residuals for a straight line, we're now doing so with an S-shaped curve that is defined by the following parameters.

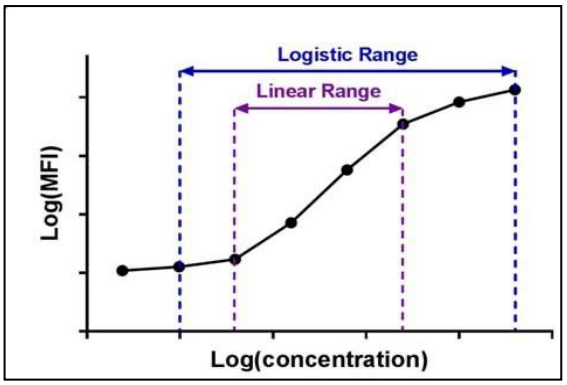
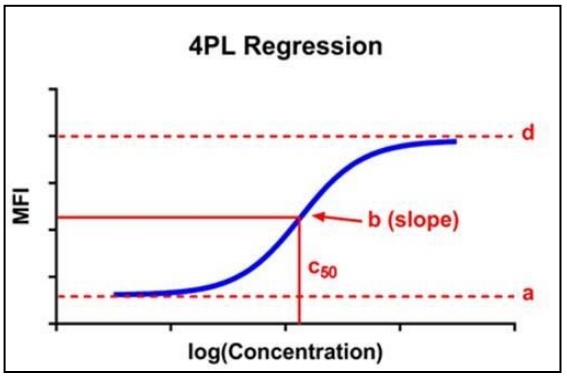


Figure 7



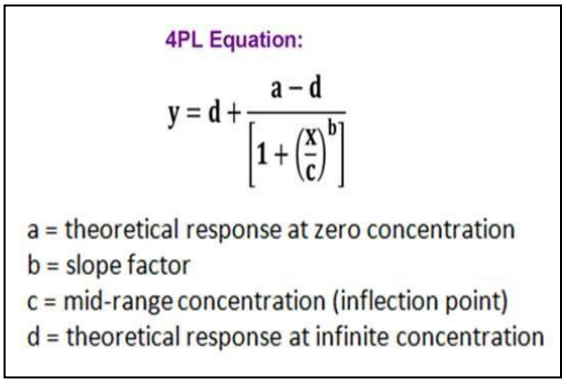


Figure 8

his type of analysis uses an equation that has a maximum and minimum incurporat - ed into it, and 4 parameters, hence the name. If your data produces a symmetrical, S-shaped curve, a 4PL fit should be sufficient to analyse your data.

⇒ 5-Parameter Logistic (5PL):

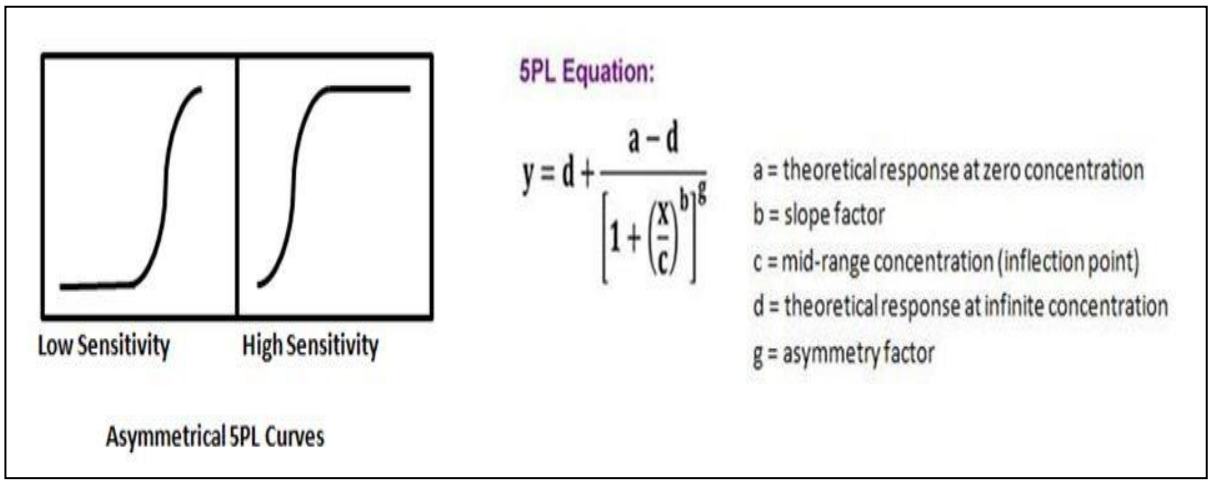
At times when running an ELISA, or more complex multiplexing assays such as LEGEND plex™, you may not get a pretty, symmetrical curve. What do you do then? There is an additional parameter that can be added to the 4PL equation, thus allowing one to do a 5PL curve fit. This fifth parameter takes into account an asymmetry factor, g, and provides a better fit when the curve does not have symmetry.

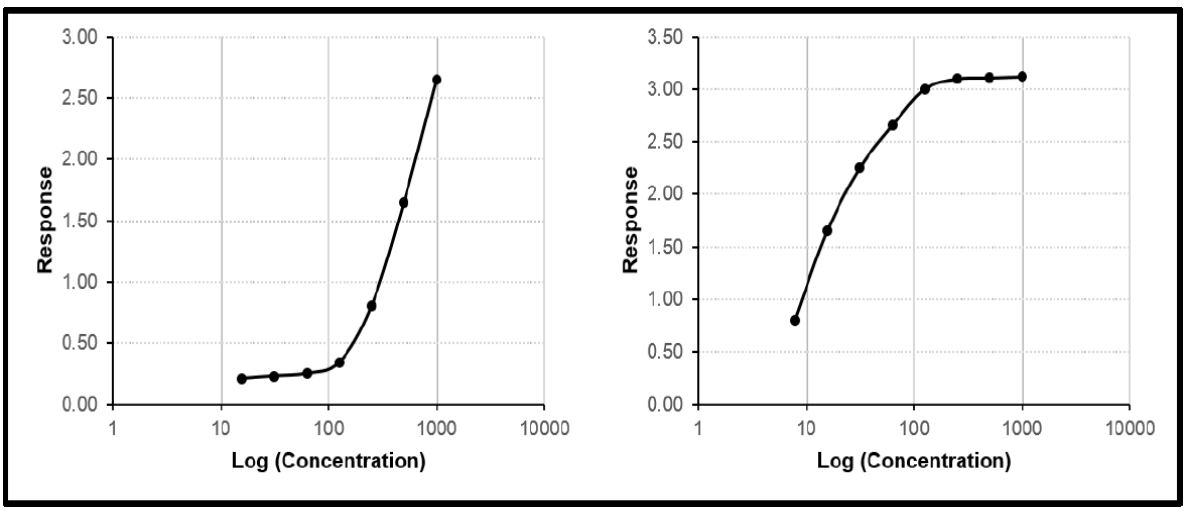
For asymmetric calibration curves (Figure 2), a 5-PL regression analysis may give a better fit, because the regression equation takes into account the asymmetry with an additional parameter g:

$$\text{5-PL Regression: } y = d + \frac{a-d}{\left[1+\left(\frac{x}{c}\right)^{b}\right]^{g}}$$

Figure 9: Above

Figure 10: Below: Examples of asymmetric immunoassay curves **which may require 5-PL regression** analysis





- **How Well Does my Model Fit the Data?**

  The answer to this question is essential for generating high quality data. For linear regression analysis the regression coefficient R is most commonly used to describe the goodness of fit. For non-linear regression models, the evaluation is slightly more complex and requires the investigation of residual variances over the calibration range. There are several practical ways to determine the goodness of fit without the need of sophisticated statistical software, two of them are presented here.

- **Residual Sum of Squares (RSS) Method:**

  This method calculates the distance of the computed response (y-value) based on the chosen regression model from the measured response value at each concentration x. The Sum of Squares (SS) is calculated according to the following equation:

$$\text{RSS} = \sum(y_{observed} - y_{calculated})^2$$

  A lower RSS value indicates a better fit.

- **Recovery of Calibration Standards:**

  This method investigates the accuracy of the observed concentration calculated by the curve-fitting model for each calibrator concentration (expected value). The recovery for each concentration x is calculated according to this formula:

$$\text{Recovery} = \frac{c_{observed}}{c_{expected}} \text{ x } 100\%$$

  The closer the recovery is to 100%, the better is the applied regression model. For accurate quantification, recovery values should be within 80 – 120%.

- **Weighting in Curve-Fitting Models:**

  Unweighted 4-PL or 5-PL regression models assume equal response variance across all protein standards ('homoscedasticity'). However, immunoassays usually show unequal variances ('heteroscedasticity') across the calibration range. The variability in response and thus the measurement error usually increases with higher response values, and at lower protein concentrations small changes in response have a larger effect on accurate determination of protein concentration. Thus, weighting algorithms are often used to offset these effects which eventually leads to the optimization of the curve-fitting model. One way of adjusting the weight is to use the reciprocal of the variance. By doing so, standards with high variance will have less weight on the calibration function, while standards with low variance will have more weight.

  Weighting in curve-fitting models requires statistical software tools and we refer to the respective instruction manuals for detailed information.

  In our Institution **(Medical College Kolkata)**, we execute this procedure in MAGELLAN software which helps us to interpret the best fit curve.

References:

1) Findley J.W.A and Dillart R.F. Appropriate Calibration Curve Fitting in Ligand Binding Assays. The AAPS Journal. 2007; 9(2) Article 29.
2) Nix B. and Wild D. 2005, Calibration Curve-fitting. In: Wild D. ed. The Immunoassay Handbook. Oxford, UK. Elsevier Ltd, pp. 233-245.
3) Cox, K.L.et al. 2001. Immunoassay Methods. In: Sittampalam, G.S. et al. eds. Assay Guidance Manual. Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004-.
4) Updating the ELISA standard curve fitting process to reduce uncertainty in estimated microcystin concentrations, PMID: 30023312.
5) Comparisons of standard curve-fitting methods to quantitate Neisseria meningitidis group A polysaccharide antibody levels by enzyme-linked immunosorbent assay, PMID: 1909345.

- **Contributors:**
  - Dr. Soumika Biswas, Assistant Professor with & Dr. Viveka R Srivastava & Dr. Sourish Ghosh, SR
  - Dr. Suparna Datta, Assistant Professor with Dr. Jayeeta Pandey, Dr. Aparna Majumder & Dr. Anirban Sinha,PGT
  - Dr. Sangita Biswas, Assistant Professor with Dr. Partha P Naskar, Dr. Mainak Dutta& Dr. Pritilata Saha,PGT
  - Dr. Neha Karar, Assistant Professor with Dr. Senjuti Ghosh& Dr. Debolina Chatterjee,PGT
  - Dr. Papia Sen, Demonstrator with Dr.Sujoy Naskar & Dr. Pritha Banerjee, PGT
  - Guided by Prof. (Dr.) Subhramay Chatterjee
- **Illustrator:**

  Dr Manali Sinha Ray, Associate Professor, Department of Biochemistry, CNMCH, Kolkata

  Dr Abhra Ghosh, Associate Professor, Department of Biochemistry, DMCH, Ludhiana