

The Advanced Precision Profile

An efficient tool for calibrating, characterizing
and optimizing assay methods.

A whitepaper discussing the mathematical and scientific aspects of the ProQuant™ software product

prepared by

Thomas M. Houts, Ph.D.

for

QIVX Inc. 357 C-1 Albion Way, Fort Collins, CO 80526

Abstract

Precision and accuracy are the two primary measures of the performance of a quantitative analytical test method or assay. Imprecision in concentration is rarely constant, and the precision profile, first described by Roger Ekins, is a convenient way to describe the precision of an assay throughout its working range. Other useful metrics, such as the upper and lower limits of quantification, are derived from the precision profile. While precision profiles have been used to characterize assay methods in routine use in clinical laboratories, they have rarely been applied in other settings, or during assay development. This is because published methods for calculating precision profiles generally require substantial amounts of data, or give unreliable estimates at the extremes of the assay range.

Developing a successful assay generally involves varying assay conditions (reagent concentrations, reaction times, measurement conditions, etc.); searching for the sets of conditions where the precision is adequate (optimal) over a suitable calibration range; and verifying that accuracy is maintained, even for "problematic" samples. Applying response-surface DOE techniques to assay development has been hampered by the difficulty in obtaining sufficient data for reliable estimates of precision, particularly since precision is a complex function of the concentration of the analyte in the sample. Some experimenters have tried optimizing on surrogate measures of performance, such as the separation in response between calibrators, with limited success. The precision profile would be the ideal tool for using DOE in optimizing assay conditions, provided that a useful profile could be obtained with reasonably small data sets.

A new approach is presented here, which efficiently utilizes data from calibrators, controls, (and unknowns, if present) to characterize a test method suitably for optimization. The precision profile is calculated, based on the pooled imprecision of the response replicates, and the slope of the dose-response curve throughout the calibration range. This calculation provides the relative standard deviation or %CV as a function of analyte concentration. Also provided are the upper and lower limits of quantification, based on the user-selected threshold for minimum acceptable precision. Accuracy can be directly estimated by including control samples with known target concentrations in the run. These direct measures of precision, accuracy, and range of quantification can be used for rapid assay characterization and optimization.

The power of the technique comes from pooling the variance in response units from all replicate groups in the run. This is based on the observation that, in many assay systems, the assumption of uniform variance in response units is valid. The software includes tests for the validity of the uniform variance assumption, and makes response transformation available should the assumption fail. Finally, the software provides confidence interval calculations for all reported values, thereby reducing the tendency for over-interpretation of the data.

Introduction

Before addressing the accuracy of an analytical method, it is first necessary to achieve adequate precision and understand its impact on estimates of bias. Keller and Passing describe the growing realization in the 1960s and 1970s that assay imprecision is rarely constant, but usually varies with the analyte concentration. This relationship between imprecision and analyte concentration is typically represented graphically as a precision profile, introduced

by Roger Ekins over twenty years ago. The precision profile plots analyte relative standard deviation or %CV against analyte concentration. Most commonly, it is generated in a two step process: estimating the relationship between response variation and response level (the Response Error Relationship or RER); and subsequently dividing the estimated standard deviation in response units by the slope of the dose response curve.

Sadler, et al observed problems with the RER/slope approach, particularly with the four-parameter logistic function typically used for calibration curves. They modified the approach by proposing a three-parameter equation to model the imprecision relationship directly in concentration units, and therefore obtain a continuous numerical estimate of imprecision throughout the calibration range. Subsequently, Sadler described the issues with estimating precision profiles from limited data sets, and discussed the importance of reporting precision profiles with confidence intervals.

The problem with previously described methods for estimating precision profiles is their data intensiveness. This is in part due to nature of the sampling uncertainty for estimating a standard deviation: the 95% confidence interval width for a standard deviation with 10 degrees of freedom is larger than the observed standard deviation itself. Fitting the response-error relationship, or directly estimating the precision profile according to the method of Sadler, requires large numbers of replicates at points throughout the calibration range. Sadler describes precision profiles based on 95 to 13,381 degrees of freedom, and still reports difficulty in getting stable estimates of the precision profile at higher concentrations. Given these limitations, precision profiles have been used primarily to characterize completed assays in routine use, rather than characterizing assays during method development.

At Syva Company, I saw many experts successfully optimize and develop immunoassays for precision and dynamic range by visually reviewing dose-response curves in duplicate or triplicate. They were intuitively evaluating the precision of the replicates and the slope of the dose-response curve at the low and high concentration ranges and making decisions about how to adjust reagent concentrations to improve precision over the calibration range of the assay. I began looking for ways to make this approach more systematic, and to apply response-surface DOE (design of experiments) techniques to general immunoassay development.

In one of the outstanding summary texts for experimental design, Box et al describe concepts and tools well established among DOE practitioners that can be brought to bear on this problem. The first concept is that it is frequently helpful to pool estimates of standard deviation *in response units* across multiple replicate groups. The second is to test the validity of the pooling by using a test for uniformity of variance, such as Bartlett's test. The third is to use variance stabilizing Box-Cox transformations if variance is not uniform. Finally, they emphasize the importance of understanding and reporting 95% confidence limits for estimated parameters. This paper describes the application of these concepts to the challenge of characterizing assay precision and accuracy suitably for assay optimization, using relatively small data sets.

Conceptual Approach

In many automated immunoassay systems, such as the EMIT® products, the range of response units spans less than an order of magnitude (frequently less than a factor of two) over the calibration range of the assay, and it was reasonable to assume uniform variance over such a small range. When tested, the assumption of uniform variance was usually found to be valid. This allows for a robust, single estimate of the pooled standard deviation in response units to be generated from all the replicated calibrators, controls, and unknown samples, if present. This pooled standard deviation, which has degrees of freedom equal to the total number of points minus the number of separate samples, efficiently utilizes all the precision information in the run.

A dose-response curve is fitted to the calibrators, using a modified 4- or 5-parameter logistic curve. The equation has the flexibility to accommodate dose-response curves ranging from straight lines through asymmetric sigmoids. The pooled standard deviation in response units is converted to standard deviation in concentration units at each concentration x by dividing by the slope of the dose-response curve at x . Dividing the result by x gives the relative standard deviation (%CV) at concentration x . Limits of quantification are determined as the concentration range where precision is better than a user-defined threshold (e.g., 20%CV).

The uncertainty in the precision profile is driven primarily by the uncertainty in the estimate of the pooled standard deviation in response units, particularly with small data sets. We estimate the 95% confidence interval for the pooled standard deviation in response units, and calculate the precision profiles at these upper and lower limits. These confidence limit precision profiles also provide upper and lower limits for the limits of quantification.

Confidence intervals for estimated concentrations can be made using the confidence interval range for the mean in response units for a given sample, using the t-distribution and the pooled standard deviation in response units. These limit values are converted to concentrations from the dose-response curve. Taking advantage of the larger number of degrees of freedom for the pooled standard deviation reduces the size of the t constant, and thereby provides narrower estimates than could be provided from the single replicate group. This improves the resolution of the analysis for detecting bias in control samples.

In some assay systems, the assumption of uniform variance in response units is not justified. These cases can be identified using Bartlett's test for uniformity of variance, or by examining a Box-Cox plot of $\log(\text{standard deviation})$ vs. $\log(\text{mean response})$. The slope of such a plot gives an indication of which response transformation will be effective at achieving uniform variance (see Numerical Methods Appendix for details). There are two advantages in using a response transformation that achieves uniform variance. First, the fitting of the dose-response curve is improved, because the residuals become independent of response level and concentration. The second is that it allows us to use a single pooled s with more degrees of freedom, thereby reducing the width of the confidence intervals.

A software program has been developed utilizing these techniques, which calculates dose response curves and reports the precision profile, analytical sensitivity, upper and lower limits of quantification, and quantified concentrations for calibrators, controls, and unknowns. Additionally, 95% confidence intervals are reported and displayed graphically for each of these values.

Results

Example 1: Immunoassay Precision Data

Data from an enzyme immunoassay were used to compare the estimated precision profiles using the ProQuant methodology with the within-run variance component analysis collected in the typical manner.

Over a five-day period, five calibration curves were run in duplicate, along with 4 control samples in 5 replicates each per run, one run per day. Control sample concentrations were quantified from the first day's calibration curve. The within-run component of variance for the quantified samples was calculated using ANOVA, and the results shown in Table 1.

Table 1: Within-Run Precision Component of Variance Estimated From Quantified Control Samples, with 95% confidence intervals (df = 20 throughout).

Control	Mean	Within-Run Variance	Within-Run CV	Lower limit	Upper Limit
A	4.65	0.2084	9.84%	7.53%	14.21%
B	9.27	0.3624	6.48%	4.96%	9.36%
C	16.05	0.3819	3.85%	2.94%	5.56%
D	24.24	1.3340	4.77%	3.65%	6.88%

The within-run precision profile was calculated using ProQuant and only the calibration curve data for each day. The results are summarized in Table 2.

Table 2: Within-Run Precision Component of Variance Estimated From Duplicate Calibration Curve Using ProQuant (df = 6 throughout).

Concentration	Day	%CV	Lower Limit	Upper Limit
4.65	1	8.35%	5.38%	18.37%
4.65	2	13.27%	8.55%	29.21%
4.65	3	6.17%	3.97%	13.57%
4.65	4	10.15%	6.54%	22.33%
4.65	5	7.07%	4.56%	15.57%
9.27	1	5.49%	3.54%	12.09%
9.27	2	8.41%	5.42%	18.52%
9.27	3	3.63%	2.34%	8.00%
9.27	4	6.30%	4.06%	13.87%
9.27	5	4.32%	2.78%	9.50%
16.05	1	4.74%	3.06%	10.43%
16.05	2	7.11%	4.58%	15.64%
16.05	3	3.08%	1.99%	6.78%
16.05	4	5.13%	3.30%	11.28%
16.05	5	3.55%	2.29%	7.82%
24.24	1	4.91%	3.17%	10.81%
24.24	2	7.35%	4.74%	16.17%
24.24	3	3.43%	2.21%	7.56%
24.24	4	5.09%	3.28%	11.19%
24.24	5	3.66%	2.36%	8.06%

The 95% confidence interval range for the precision estimated from the duplicate standard curve data includes the observed value for the within-run precision component calculated from the quantified samples (Table 1) in all but one case (day 2 at concentration 16.05).

The ProQuant methodology provides estimates of the within-run imprecision over the entire range of the dose response curve. This is represented graphically in Figure 1 for day 1 of the immunoassay data.

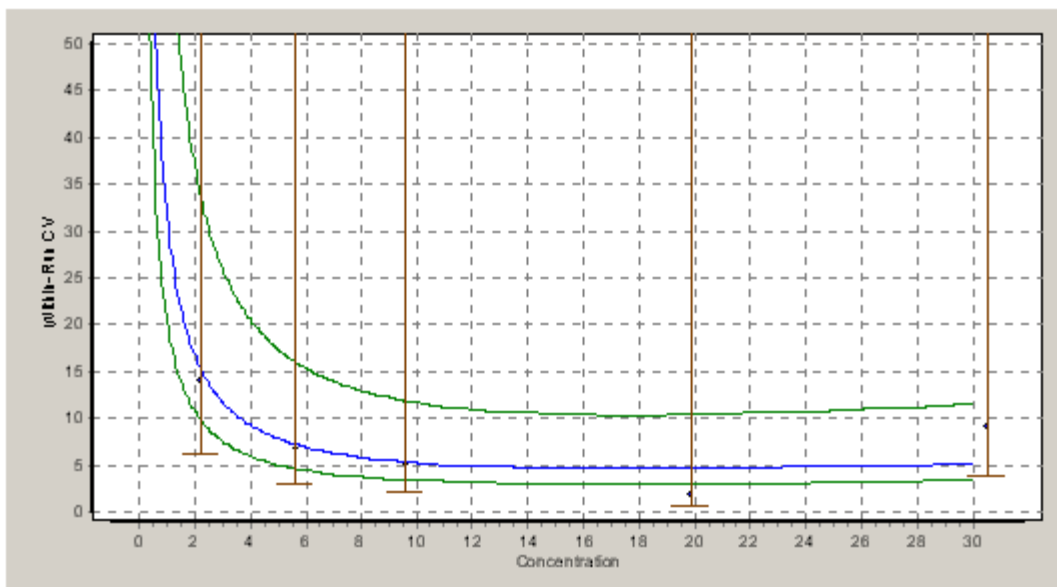


Figure 1 The blue curve represents the precision profile, and the green curves surrounding it represent the 95% confidence limits for the precision profile. The data points represent the observed mean and %CV of the quantified values for each replicated calibrator, with error bars showing the 95% confidence interval in the calculated %CV (the error bars are so wide, because each standard deviation is based on $n=2$, so has only 1 degree of freedom).

Example 2: Simulated Linear Data with Response Transformation

To further challenge the approach, data were simulated with a linear calibration relationship ($y = 20x+10$) with non-uniform response variance coming from two sources: a constant variance source with a standard deviation of 3, and a constant CV source with $CV=5\%$. The response standard deviation σ at any response y is therefore given by

$$\sigma_y = \sqrt{3^2 + (0.05y)^2}$$

Random data were generated for a calibration curve at concentrations 0, 2, 4, 6, 8, and 10, to give 10 replicates at each concentration with mean response and standard deviation of response from the two equations above.

When the data are analyzed using ProQuant, the results indicate non-uniformity of variance (Bartlett's p-value 0.005). The slope of the Box-Cox plot is 0.38, suggesting a square root ($y^{0.5}$) transformation. After transformation, the data passed the uniformity of variance test (Bartlett's p-value 0.232). The resulting precision profile is shown in figure 2.

Precision Profile, Simulated Linear Data

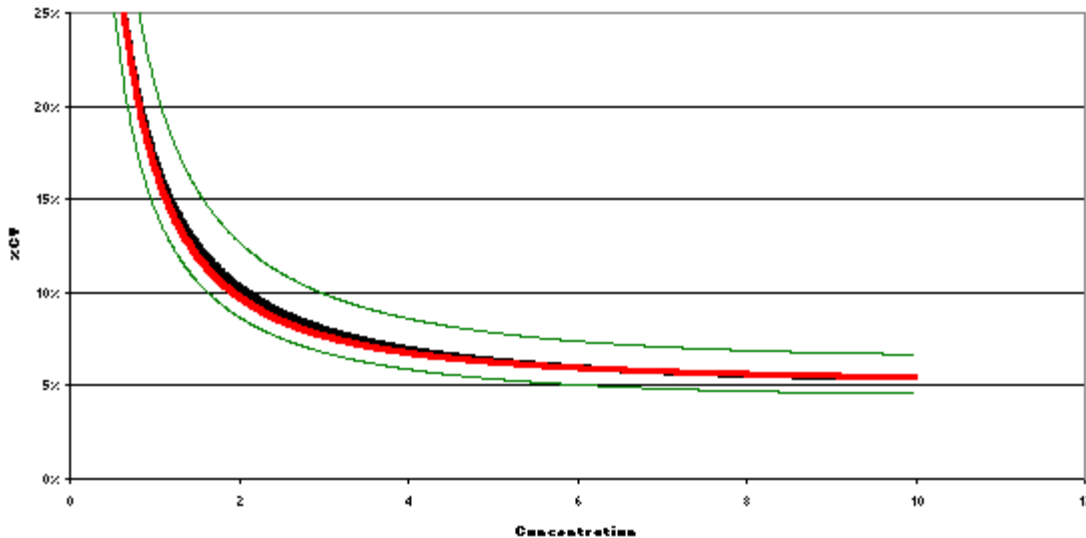


Figure 2. Precision profile for simulated linear data. Details of simulation are in text. The red line represents the true precision profile. The blue line represents the precision profile estimated using the ProQuant™ methodology from the calibration curve replicates after square root transformation. The green lines represent the upper and lower 95% confidence limits for the estimated precision profile.

The results for the other key metrics are shown in Table 3:

Metric	True Value	Estimated Value	Lower Limit	Upper Limit
Sensitivity / LOD	0.45	0.44	0.37	0.55
LLOQ	0.82	0.84	0.69	1.08

Discussion

This new tool provides estimation of the primary measures of quantitative assay performance: precision, accuracy, and upper and lower limits of quantification, using as little data as 5 calibrators in duplicate. It also provides 95% confidence limits for all measures, which allows understanding the limitations of the estimates for small data sets, and demonstrates the narrower confidence intervals with additional data.

This tool, which efficiently generates quantitative measures of assay performance, is ideal for use with DOE in assay optimization. We recommend optimizing to simultaneously achieve smallest %CV at both the lowest positive and the highest calibrator, with verification that accuracy for controls is not systematically affected.

Although this paper has focused on within run imprecision, total imprecision can be estimated by running (or analyzing) singleton calibration data across multiple days. Alternatively, variance components in response units can

be used, in combination with the slope of the dose-response curve, to generate variance components in concentration units at any point within the calibration range.

The response transformation feature allows almost any data sets to be used effectively. The required transformation provides information regarding the response-error relationship. Fitting of a dose-response curve after transformation should also improve calibration performance. The software tests for uniformity of variance, allowing the user to verify that the transformation has been successful in achieving uniform variance. If no transformation achieves uniform variance, the Box-Cox plot can be evaluated to determine if extreme outliers are present in one or more of the replicate groups.

This approach is limited to monotonic dose-response curves (e.g., no high-dose hook effect).

These techniques are available in a software package, which may be downloaded for a free trial period at <http://www.qivx.com/ProQuant>

Appendix: Numerical Methods

The source data are response measurements for calibrators and (optionally) controls or unknowns. To estimate a 4-parameter dose-response curve, we recommend at least 5 calibrator levels. Up to 20 different calibrators or controls can be included, each replicated up to 60 times.

A pooled estimate of the variance in response units is calculated from all replicate groups, according to the standard formula:

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{N - k}}$$

Where there are N total data points among k groups, with n_i replicates in the i th group. Variance uniformity among the groups is evaluated using Bartlett's test.

A dose-response curve is generated by non-linear regression to the calibrator data. We have selected a calibration curve equation that adapts well to both immunoassay data and linear data types:

$$y = C_0 + \frac{C_1}{1 + e^{C_2 \ln\left(\frac{x}{C_3} + C_4\right)}}$$

where y is response at concentration x , and C_0 , C_1 , C_2 , C_3 , and C_4 are curve fit parameters. We typically assign parameter C_4 the value of 0.5 and fit the remaining 4 parameters. When $C_4 = 0$, this is equivalent to a 4-parameter logistic dose response curve.

The slope of the dose response curve at any concentration x can be calculated by differentiation:

$$slope = \frac{dy}{dx} = -\frac{C_1 e^{(C_2 \ln\left(\frac{x}{C_3} + C_4\right))} \left(\frac{C_2}{x + C_3 C_4}\right)}{(1 + e^{(C_2 \ln\left(\frac{x}{C_3} + C_4\right))})^2}$$

The standard deviation in concentration units at any concentration x is estimated by dividing s_{pooled} by the slope of the dose response curve at that concentration. Dividing the result by concentration x gives the imprecision of the method, as %CV or relative standard deviation in concentration units. Confidence intervals can be calculated for the estimated imprecision based on the uncertainty in the estimate of s_{pooled} based on $N-k$ degrees of freedom. In relatively small data sets, the contribution from uncertainty in the slope of the dose response curve was found to be negligible.

Accuracy is estimated by comparing the known concentration for calibrators and target concentration for controls

with those estimated from the calibration curve based on the mean response (\bar{y}), using the equation:

$$x = C_3 e^{(\ln(-1+C_1(y-C_0))/C_2-C_4)}$$

Upper and lower confidence intervals for the estimated concentration are calculated for each replicate group by first calculating the upper and lower confidence intervals for the mean response for the group according to the expression

$$\bar{y} \pm t s_{pooled} / \sqrt{n}$$

where n is the number of replicates in the group, and the value of t is based on the desired confidence level and the number of degrees of freedom in the estimate of s_{pooled} ($=N-k$). These boundary values for the response are subsequently converted into concentration units.

Central to this approach is that a single pooled estimate of s is valid. This is tested using Bartlett's test for uniformity of variance⁷. If the variance is found to be non-uniform we use the response-transformation approach taught by DOE practitioners. We examine the slope of a plot of $\log(s)$ vs. $\log(\bar{y})$; we calculate the Box-Cox parameter lambda as 1-slope. The Box-Cox transformation applied is

$$y' = y^\lambda \text{ for } \lambda \text{ not equal to zero;}$$

$$y' = \log(y) \text{ for } \lambda \text{ equal zero.}$$

If response transformation is indicated, a new dose-response curve is calculated, relating the transformed responses to the concentration, and precision estimates are calculated using the pooled s in transformed response units.

Lower limit of detection is calculated as the concentration corresponding to the predicted response at zero dose plus 3 times pooled.

Upper and lower limits of quantification are calculated as the concentrations where the precision profile crosses the user-specified precision threshold, typically 20% CV.

References

1. Keller, H. and Passing, H.: Performance profiles: new tools for characterization and comparison of clinical chemical results. *J Clin. Chem. Clin. Biochem.* 27:613-629 (1989).
2. Ekins, R.P. (1976) In: *Hormone assays and their clinical application* (Loraine, J.A. & Bell, T.T, eds.) Churchill Livingstone, Edinburgh.
3. Ekins, R.P. (1983) The precision profile: its use in assay design, assessment and quality control. In: *Immunoassays for clinical chemistry*, 2nd edition (Hunter, W.M. & Corrie, J.E.T., eds.) Churchill Livingstone, Edinburgh.
4. Sadler, W.A., Smith, M.H., and Legge, H.M.: A method for Direct Estimation of Imprecision Profiles, with Reference to Immunoassay Data. *Clin. Chem.* 34:1058-1061 (1988).
5. Sadler, W.A., and Smith, M.H.: Use and Abuse of Imprecision Profiles: Some Pitfalls Illustrated by Computing and Plotting Confidence Intervals. *Clin. Chem.* 36:1346-1350 (1990).
6. Box, G.E.P, Hunter, W.G., and Hunter, J.S. *Statistics for Experimenters.* John Wiley and Sons, New York. 1978.
7. Dixon, Wilfrid J. and Massey, Frank J. Jr.: *Introduction to Statistical Analysis.* 4th Edition. McGraw-Hill, San Francisco. 1983.