

Research article

Open Access

Evaluating concentration estimation errors in ELISA microarray experiments

Don Simone Daly*¹, Amanda M White¹, Susan M Varnum²,
Kevin K Anderson¹ and Richard C Zangar²

Address: ¹Statistical and Mathematical Sciences, Pacific Northwest National Laboratory, PO Box 999, Richland, WA, USA and ²Biological Sciences, Pacific Northwest National Laboratory, PO Box 999, Richland, WA, USA

Email: Don Simone Daly* - DS.Daly@PNL.gov; Amanda M White - Amanda.White@pnl.gov; Susan M Varnum - susan.varnum@pnl.gov; Kevin K Anderson - Kevin.Anderson@pnl.gov; Richard C Zangar - Richard.Zangar@pnl.gov

* Corresponding author

Published: 26 January 2005

Received: 24 June 2004

BMC Bioinformatics 2005, 6:17 doi:10.1186/1471-2105-6-17

Accepted: 26 January 2005

This article is available from: <http://www.biomedcentral.com/1471-2105/6/17>

© 2005 Daly et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Enzyme-linked immunosorbent assay (ELISA) is a standard immunoassay to estimate a protein's concentration in a sample. Deploying ELISA in a microarray format permits simultaneous estimation of the concentrations of numerous proteins in a small sample. These estimates, however, are uncertain due to processing error and biological variability. Evaluating estimation error is critical to interpreting biological significance and improving the ELISA microarray process. Estimation error evaluation must be automated to realize a reliable high-throughput ELISA microarray system.

In this paper, we present a statistical method based on propagation of error to evaluate concentration estimation errors in the ELISA microarray process. Although propagation of error is central to this method and the focus of this paper, it is most effective only when comparable data are available. Therefore, we briefly discuss the roles of experimental design, data screening, normalization, and statistical diagnostics when evaluating ELISA microarray concentration estimation errors.

Results: We use an ELISA microarray investigation of breast cancer biomarkers to illustrate the evaluation of concentration estimation errors. The illustration begins with a description of the design and resulting data, followed by a brief discussion of data screening and normalization. In our illustration, we fit a standard curve to the screened and normalized data, review the modeling diagnostics, and apply propagation of error.

We summarize the results with a simple, three-panel diagnostic visualization featuring a scatterplot of the standard data with logistic standard curve and 95% confidence intervals, an annotated histogram of sample measurements, and a plot of the 95% concentration coefficient of variation, or relative error, as a function of concentration.

Conclusions: This statistical method should be of value in the rapid evaluation and quality control of high-throughput ELISA microarray analyses. Applying propagation of error to a variety of ELISA microarray concentration estimation models is straightforward. Displaying the results in the three-panel layout succinctly summarizes both the standard and sample data while providing an informative critique of applicability of the fitted model, the uncertainty in concentration estimates, and the quality of both the experiment and the ELISA microarray process.

Background

Proteomic approaches are resulting in the identification of large numbers of proteins that can potentially be used as disease markers or drug targets. Unfortunately, proteomic approaches currently lack the throughput or quality metrics necessary to evaluate hundreds or thousands of samples that may be required to determine clinical usefulness of a biomarker [1]. Traditionally, candidate biomarkers have been commonly evaluated using a 96-well enzyme-linked immunosorbent assay (ELISA). However, this approach is not suited for analyzing more than a few proteins when sample volumes are limited, as is commonly the case for early tumor samples. For this reason, we and others are developing ELISA microarray systems to evaluate 20 to 50 proteins using only a few microliters of sample in an efficient and quantitative manner [2,3].

Processing a ELISA microarray experiment produces large volumes of data of wide variety and high complexity. Similar to traditional 96-well ELISA data, ELISA microarray data often are perturbed by processing error [4-7]. Processing errors are introduced by unintended variation in sample preparation, slide or pin arrangement, printing, imaging, and estimation of spot summary statistics. The specific role of concentration error estimates and the general role of statistical diagnostics is to reveal process accuracy and precision. This evaluation then enables an insightful interpretation of the biological significance, an informative critique of the current experiment, and insights to improve the accuracy and precision of future experiments. In a high-throughput ELISA microarray system, there is a need to not only quickly and accurately generate the standard curves and estimate concentrations from the sample data, but also to quickly evaluate the quality of those estimates. The resulting information can be used in both the development stage for optimizing assay conditions and in the production phase for ensuring that the overall analytic process is working well on a day-to-day basis.

Statistically evaluating ELISA microarray concentration estimation errors depends upon both the availability of the appropriate set of comparable measurements and the choice of data analysis methods. Sufficient replication within and across arrays is key to making precise estimates of both concentrations and errors [8]. Hence, evaluating concentration estimation errors in an ELISA microarray experiment begins with the design of the experiment. Evaluation of these estimation errors also depends on the recording of the pedigree, or history, of each result from probe preparation and array printing through sample preparation and spot intensity estimation [9-12]. Screening for anomalous results and normalizing within and across arrays may significantly reduce obscuring variation and improve homogeneity [13-15]. Although the mathe-

matical statistics and algorithms are quite sophisticated, software makes actual estimation and application of the standard curve and the concentration error function straightforward. This is true also for the presentation of modeling results for diagnostic purposes.

In this paper, we describe and illustrate a methodology for calculating the concentration estimation error of each assay in an ELISA microarray experiment based on a statistical analysis of the most likely sources of error. We expect the resulting data analysis algorithms to be a key component in a bioinformatics package for evaluating ELISA microarray data.

Methods

Making concentration estimates and estimating their errors in our ELISA microarray studies involve a sequence of steps beginning with the layout of the ELISA microarray and design of the experiment. Following execution of the analytical components of the experiment, the statistical analysis proceeds with data screening, normalization, and model identification. Estimation and evaluation of the standard curves and error estimation functions come next. Finally, the standard curves and error estimation functions are applied and then evaluated using a modeling diagnostic.

Layout of the ELISA microarray and design of the experiment

To estimate errors in concentration estimates, it is necessary to carefully lay out the microarray and design the experiment. Our layout features several distally separate replicates of each assay spot on each microarray to evaluate local processing effects. Our design addresses selection and application of treatments – in particular, replicate treatments – to a collection of arrays. This replication facilitates adjustments for the sources of variability that lead to ambiguous concentration estimates [16,17]. In array experiments featuring relatively small numbers of assays, usually 50 or fewer analytes, thoughtful design is critical to normalization, calibration, and estimation of concentrations due to the significant lack of technical replicates found in arrays with thousands of assays. With regard to error estimation, the major consideration in the design of the experiments is replication of treatments across arrays to capture the effects of process error.

To illustrate our technique for evaluating estimation errors in an ELISA microarray experiment, we used a subset of data from an ELISA microarray investigation of breast cancer biomarkers. The ELISA microarray experiments were performed as previously described [2,3]. Briefly, capture antibodies were covalently attached to an aminosilanated glass slide surface (Sigma, St. Louis, Missouri, USA) using a Microgrid 2 robot from Genomic

Solutions (Ann Arbor, Michigan, USA) equipped with ChipMaker2 split pins from TeleChem (Sunnyvale, California, USA). As demonstrated previously, these spots are typically uniform in shape with a reasonable homogeneous distribution of protein across the spot [1-3]. That is, "donut" formation is not normally observed. These spatially confined antibodies bind a specific antigen from a sample overlaying the array. A second, biotinylated antibody that recognizes the same antigen as the first antibody but at a different epitope is then used for detection. Detection of the second antibody is based upon streptavidin (which binds biotin) and an enzymatic signal enhancement method known as tyramide signal amplification (TSA). The resultant fluorescence was detected at 10-micron scan resolution using a ScanArray 3000 from General Scanning (Billerica, Massachusetts, USA). The experiment used 94 arrays printed in pairs on 47 slides. Each array contained 4 (2×2) replicate subarrays of 25 (5×5) spots. A subarray contained 21 unique assays, 1 positive control and 3 negative control spots. A set of 7 known standard concentrations and a buffer blank was assembled by performing a three-fold dilution series of a single mixture of all the standards. Each standard concentration was applied to duplicate slides. The remaining 39 slides were treated with serum samples from women with or without breast cancer. These sera were encoded to prevent knowledge of the study group during sample processing. The treated microarrays were imaged with a ScanArray microarray scanner (PerkinElmer, Boston, Massachusetts, USA). The spot fluorescence estimates were calculated with custom array-image-analysis software that was developed in-house.

Data screening, normalization and model identification

Data screening, an exploratory data analysis, serves several purposes – identifying outliers, anomalous values, and experimental design shortcomings; identifying data transforms to improve curve-fitting and application; identifying measurement trends and other processing effects; and suggesting an appropriate functional form for the standard curve [6,18-21]. This exploratory analysis combines simple summary statistics and graphical displays. For instance, graphs of control spot intensities versus processing variables such as array print order or pin number may reveal variability due to processing. These processing trends can be made more apparent with locally weighted regression, or loess, a statistical technique to fit a smooth curve through the scatterplot [22,23]. These graphs can be used as the basis for modifying the process or for data normalization.

Because our protein arrays feature fewer spots per array than do typical gene expression microarrays, a different approach to normalization, suitable for low spot-frequency arrays, is required. This normalization is critical,

given that array-to-array processing error is common and that standard curves are estimated from reference spot intensities calculated from one set of arrays and then applied to sample spot intensities estimated from a separate set of arrays.

A scatterplot of intensity estimates of standard spots versus concentration is particularly useful. First, outliers and anomalies may be readily apparent. Second, the spacing between concentration values may be assessed. If standard concentrations follow from a dilution series, then the separation between concentrations decreases significantly with the decrease in concentration. This results in spot intensities measured at higher concentrations having much more leverage on the fit of the model than may be desirable. It should also be apparent whether the variability in spot intensity is increasing with mean spot intensity. Both increasing spacing in the concentrations and heteroskedasticity in the measured intensities affect the model fit and follow-on statistical inferences [24]. These may be minimized with \log_e transformations of both concentrations and spot intensities.

A scatterplot of raw or transformed standard spot intensity versus concentration also provides an indication of the appropriate model for the data. In particular, data following a sigmoid curve favor the logistic curves while data apparently lacking the horizontal asymptotes of a sigmoid curve favor a linear or power law model. Although several models may be fit and one selected based on a goodness-of-fit statistic (see next section), the scatterplot is a useful visual check on this selection.

Several plots provide useful information about the quality of the fitted model. Of special importance are the scatterplot of residuals versus concentration and the scatterplot of residuals versus estimated intensity. In both cases, the variability of the residuals should be centered about zero and constant across concentration or intensity. Model bias is indicated by a systematic drift of residuals to one side of the zero line. Heteroskedasticity is indicated by a systematic change in the variation of the residuals. Both may indicate that a better model is necessary before proceeding to estimation of sample concentrations and estimation of concentration errors.

Standard curves and estimation errors

An ELISA standard curve expresses protein concentration as a function of spot intensity. One standard curve is required for each assay. In an ELISA microarray experiment, the standard data are collected by fixing a set of concentrations and measuring spot intensities via imagery of the treated arrays. A standard curve is estimated by fitting an appropriate function to the set of (concentration,

intensity) measurement pairs [25]. This equation is then inverted to obtain the standard curve.

Common parametric choices for standard curve models are multiparameter logistic functions and power law functions. For an ELISA microarray, a strictly monotone model is consistent with the belief that a monotone change in concentration should result in a monotone change in spot intensity.

We estimate standard curves with both logistic and power law parametric models. The four-parameter logistic model [26], expressing intensity I as a function of concentration C and parameters P_1, P_2, P_3 and P_4 , is defined as

$$I(C|\mathbf{P}) = P_1 + \frac{P_2 - P_1}{1 + \exp\left[\frac{(P_3 - C)}{P_4}\right]} + \varepsilon \quad (1)$$

The two-parameter power law model [27] expressing intensity I as a function of concentration C and parameters P_1 and P_2 , in \log_e terms, is

$$\log_e(I) = P_1 + P_2 \log_e(C) + \varepsilon$$

We assume the errors, denoted by the term ε , are independent and normally distributed with mean 0 and variance σ^2 . With either of these parametric models, concentration estimation errors may be estimated using propagation of error, also known as the delta method.

To choose between competing candidate models, a number of measures exist for evaluating model fit when replicate observations of each assay are available. These include partitioning the mean squared error, or MSE, into components representing pure error and lack of fit [28], and penalized measures such as Akaike (AIC) and Bayesian (BIC) information criteria [29]. We also examine the PRESS statistic, a direct measure of the predictive capability of each candidate model [30].

To calculate the PRESS statistic for each candidate model, suppose we exclude each point (x_j, y_j) in turn and fit the model to the remaining points. We predict the value $\hat{y}_{j,-j}$ at the excluded point x_j and calculate the PRESS residual defined by $e_{j,-j} = y_j - \hat{y}_{j,-j}$. Then, the PRESS statistic is the sum of the squared PRESS residuals

$$PRESS = \sum_j (e_{j,-j})^2$$

The candidate model with the lowest PRESS score as the best predictive model to estimate concentrations.

The basic approach to estimating concentration errors with the propagation of error method has three steps [31]. First, fit intensity as a function of concentration and estimate the covariance among model parameter estimates. Next, solve the fitted function for concentration as a function of intensity. Finally, propagate error estimates from the fitted model through the inverted model and combine with the error estimate of the sampled spot intensity to estimate the concentration estimation error.

Let $C(I|\hat{\mathbf{P}})$, with $\hat{\mathbf{P}} = (\hat{P}_1, \hat{P}_2, \dots, \hat{P}_N)$, denote the inverted N parameter model expressing concentration C as a function of intensity I and the parameter estimates $\hat{\mathbf{P}}$. Suppose $\hat{\Sigma}_P$ is the $N \times N$ parameter covariance matrix estimated by fitting I as a function of C , say $I(C|\mathbf{P})$. Now, let C_S be the estimated concentration from the sample intensity estimate I_S , say $C_S = C(I_S|\mathbf{P})$ and $\hat{\sigma}_S$ be the corresponding estimated standard error of I_S . Then, the propagation of error estimate for the concentration estimate C_S is the square root of the product of $\hat{\Sigma}$, the sample covariance matrix augmented with $\hat{\sigma}_S$, and the Jacobian matrix J evaluated at I_S and the parameter estimates $\hat{P}_1, \hat{P}_2, \dots, \hat{P}_N$. In this application, the Jacobian is the matrix of partial derivatives of $C(I|\mathbf{P})$ with respect to the intensity I and the parameters \mathbf{P} . Hence, the concentration estimation error of $C(I|\mathbf{P})$ is the square root of the concentration estimation variance $V(C(I|\mathbf{P}))$

$$V(C(I|\mathbf{P})) = J(C(I|\mathbf{P}))^T \Sigma J(C(I|\mathbf{P}))$$

where the Jacobian is

$$J(C(I|\mathbf{P}))^T = [\partial C/\partial I, \partial C/\partial P_1, \dots, \partial C/\partial P_N]$$

and the augmented covariance matrix is

$$\Sigma = \begin{bmatrix} \sigma_S^2 & 0 \\ 0 & \Sigma_P \end{bmatrix}$$

Hence, the formula for estimated standard error of C_S is

$$SE[C_S] = \sqrt{V(C_S, I_S, \mathbf{P})}$$

For a given intensity estimate I_S and standard error $\hat{\sigma}_S$, the estimated concentration and approximate 95% confidence interval $(C_{95\%L}, C_{95\%U})$ are

$$C_S = C(I_S)$$

$$C_{95\%L} = C_S - 2SE[C_S] \quad (2)$$

and

$$C_{95\%U} = C_S + 2SE[C_S] \quad (3)$$

For example, consider the four parameter logistic model, Eqn. 1. The concentration estimation equation is obtained by solving this equation for C in terms of I and the four parameters

$$C(I | \mathbf{P}) = P_3 - P_4 \log_e \left(\frac{P_2 - I}{I - P_1} \right) \quad (4)$$

The Jacobian matrix is obtained by taking the partial derivatives of the inverted four-parameter logistic function of C (Eqn. 4) with respect to I and the parameters P_1, P_2, P_3 and P_4

$$\frac{\partial C}{\partial I} = \frac{P_4 (P_2 - P_1)}{(I - P_1)(P_2 - I)}$$

$$\frac{\partial C}{\partial P_1} = \frac{-P_4}{I - P_1}$$

$$\frac{\partial C}{\partial P_2} = \frac{-P_4}{P_2 - I}$$

$$\frac{\partial C}{\partial P_3} = 1$$

$$\frac{\partial C}{\partial P_4} = -\log_e \left(\frac{P_2 - I}{I - P_1} \right)$$

Diagnostic visualizations

A three-panel display combining a histogram of normalized sample spot intensities for a given antigen, its corresponding standard curve, and the graph of the concentration coefficient of variation, or relative error, versus concentration provides pertinent information about the conduct of the current experiment as well as information to improve future experiments. The standard curve panel presents a scatterplot of normalized standard spot intensities versus standard concentrations. The scatterplot is overlain with the estimated standard curve expressing concentration as a function of spot intensity. This panel also includes approximate 95% confidence intervals. These intervals summarize the uncertainty in concentration estimates due to both the uncertainty in estimating the standard curve and the uncertainty in the sample spot intensity estimate. Finally, a highlighted region helps distinguish concentration estimates with acceptable errors from concentration estimates with possibly less than acceptable errors.

The segment of the standard curve corresponding to acceptable concentration errors may be determined using the 95% confidence intervals. The lower and upper end-

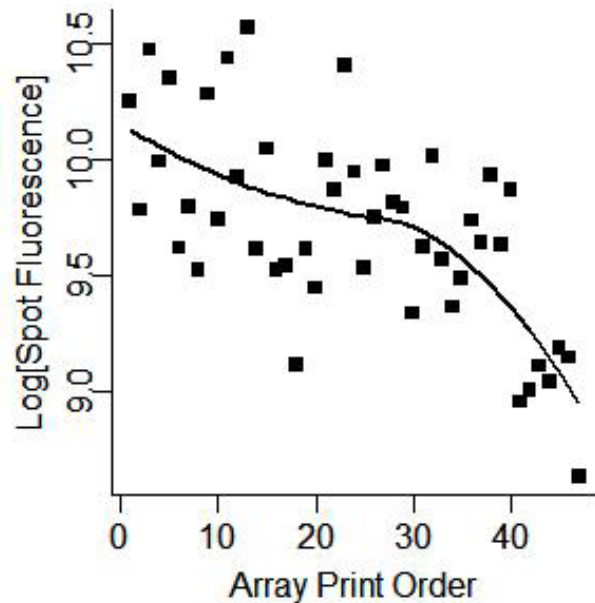


Figure 1
Control spot fluorescence and slide print order We expect control spot intensity to be constant across slide print order. The trend, estimated with loess regression, indicates a processing effect due to printing order. In this case, the decrease in control spot intensity, coupled with knowledge of the printing process, suggests that the quantity of printed material is decreasing over a printing run.

points of this segment, (I_L, C_L) and (I_U, C_U) , are the two points such that the confidence intervals begin to increase significantly in length. This segment generally corresponds to the linear segment of a standard curve. We identify the intensity I_L of the lower pair as the smallest intensity such that $95\% \text{ UB}(I_L)$ is less than $95\% \text{ UB}(I)$ for intensity values I less than I_L . Similarly, we identify I_U as the largest intensity such that $95\% \text{ LB}(I_U)$ is greater than $95\% \text{ LB}(I)$ for intensity values I greater than I_U . We define C_L and C_U to be $C_L = C(I_L)$ and $C_U = C(I_U)$, respectively. We believe that this is a conservative approach to identifying intensities that generate concentration estimates with acceptable errors.

An informative visualization of acceptable concentration estimates may be generated using the points (I_L, C_L) and (I_U, C_U) . Consider the union of the two rectangular regions defined by the two sets of vertices $[(I_L, 0), (I_L, C_L), (I_U, C_U), (I_U, 0)]$, and $[(0, C_L), (0, C_U), (I_U, C_U), (I_U, C_{LU})]$. This union defines an L-shaped region covering the standard curve segment and bound at its extremes by the inten-

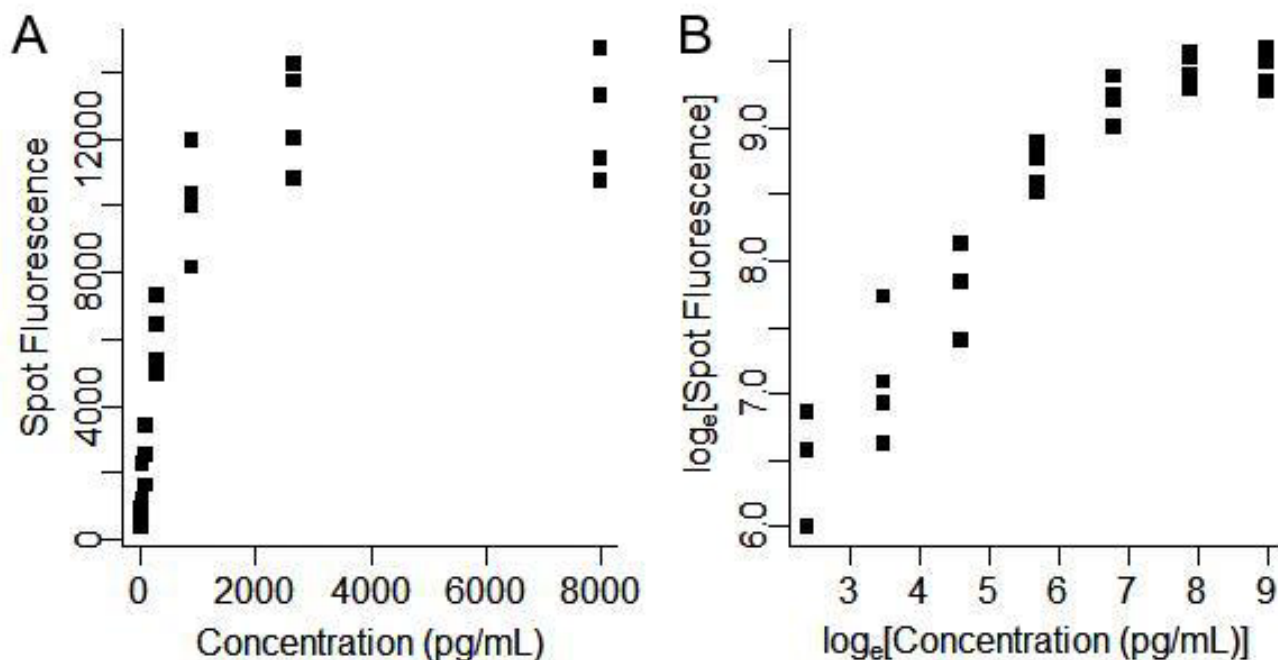


Figure 2
Spot fluorescence and standard concentration A graph of spot fluorescence as a function of standard concentration (A) shows that both the difference in concentration and the variability in spot fluorescence increase with concentration. Both negatively affect model fitting and subsequent statistical inferences due to the excessive leverage of the larger concentrations and the heteroskedasticity of the spot intensities. In this case, log_e transforms of both concentration and spot fluorescence reduce the excessive leverage and improve the homogeneity (B). The log transformation is consistent with the featured concentration dilution series and is a common transformation of chemical measurements that often are assumed log normally distributed.

sity and concentration segments. From this visualization, one can quickly grasp the dynamic range of acceptable intensities and the potential range of acceptable concentration estimates.

In regard to this first panel, two notable aspects of this propagation of error methodology are noteworthy. First, the error bands are computed pointwise and provide reasonable error estimates for a small number of concentrations. As the number of concentration estimates grows, the impact of the multiple testing problem grows [32]. This is a problem in any biomedical testing that features numerous simultaneous tests and has spawned considerable debate and research. The second aspect of note is the divergence of the error bands from the estimated standard curve as the standard curve approaches a horizontal asymptote. We see this apparent deficiency in the method as a plus. This divergence is a clear indicator that concentration estimates in the segment of a standard curve approaching a horizontal asymptote are highly suspect.

The second panel in this display shows the concentration coefficient of variation – that is, $\%CCV = 100 * SE(C|I)/$

$C(I)$, or relative error of a concentration estimate – as a function of concentration. This provides an alternative view of the error in concentration estimation over the concentration range covered by the concentration estimation equation. A standard curve modeled with a four-parameter logistic function generally will have a bathtub shape due to the increasing uncertainty in concentration estimates at the two ends of the concentration range where the curve approaches horizontal asymptotes.

The third panel in this display features an annotated histogram of sample spot intensity estimates on the intensity axis opposite the scatterplot. In this representation, it is easy to see the extent of overlap between the distribution of sample intensity estimates and the range of intensities that result in concentration estimates with acceptable errors.

Results and discussion

To evaluate concentration estimation errors in the example analysis, we attempted to quantify or understand those errors that we can and minimize those errors that we cannot. We began with data screening. The most signifi-

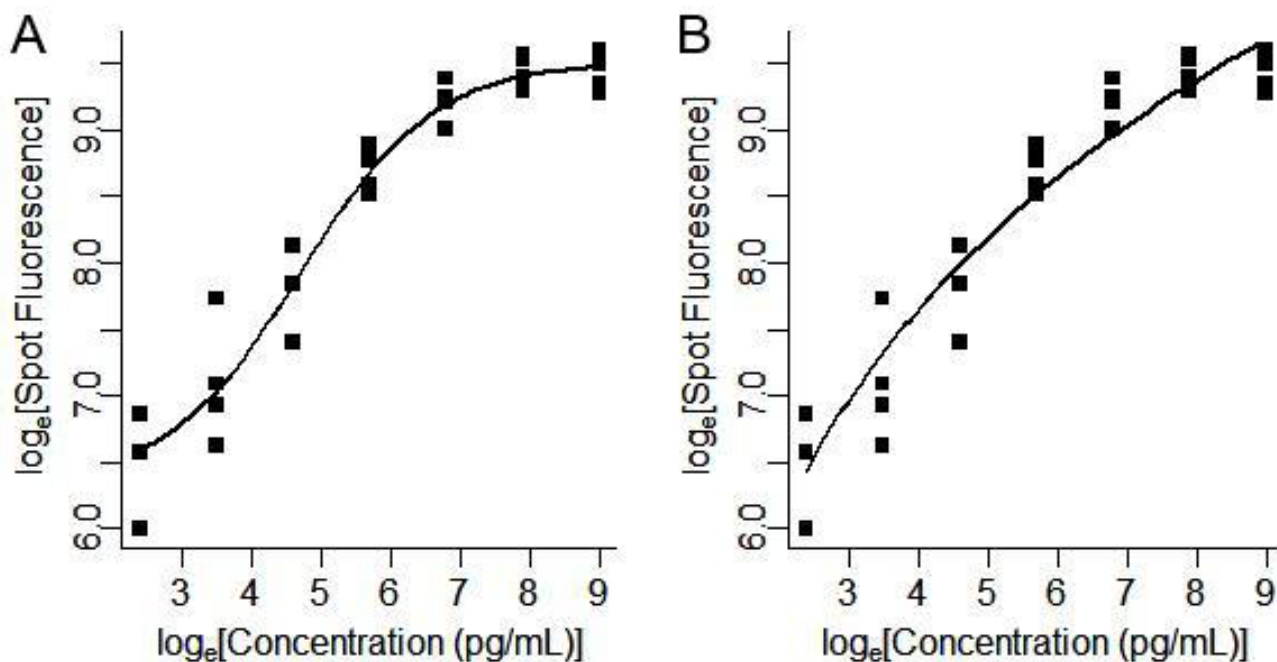


Figure 3
Candidate standard curves A review of the graphs of a four-parameter logistic curve (A) and a power law curve (B) suggests that the former shows higher fidelity to the data and is a better choice for the standard curve in this case.

cant anomaly uncovered during this exploratory analysis of the cancer biomarker data was a decreasing trend in control spot intensity as a function of array print order (Figure 1). The trend was quantified using loess, a flexible, nonparametric method to fit a smooth curve through a scatterplot to uncover trends in data [22,23]. This trend suggests that 1) normalizing across arrays would improve precision; 2) in future experiments, assigning study groups to arrays should address array print order; and 3) array printing should be monitored and, if possible, modified to reduce this source of obscuring variation. In this case, we normalized for slide-level processing errors by subtracting from each spot's \log_e (fluorescent intensity) the difference between the mean of its slide's control spot \log_e (intensities) and the corresponding loess estimate.

Our evaluation addressed the selection and fitting of an acceptable concentration estimation model. To that end, we examined two plots. The first displays the fluorescent intensities of the standards as a function of concentration (Figure 2A). Two characteristics of the data that significantly affect selecting and fitting the model and then interpreting the results in a statistically meaningful way are apparent.

The first is heteroskedasticity, or the increasing variation in fluorescent intensity with increasing concentration. Meaningful statistical inferences about concentration estimation errors depend upon correct modeling assumptions. To apply propagation of error when estimating and then interpreting the approximate 95% confidence intervals, we rely on normal distribution theory and require that the random variability in spot intensities be homogeneous across concentrations [28]. In this case, a \log_e transformation of the intensity estimates stabilizes the variability across concentrations (Figure 2B).

The second characteristic is the undue leverage of data at high concentrations due to the increasing separation between standard concentrations with increasing concentration. Although both are expected (the first due to the randomness generally observed when counting photons, and the second due to the use of a concentration dilution series in the design), each must be addressed to achieve the best fit of the standard curve and resulting concentration estimation inferences. A \log_e transformation of the concentrations standardizes the separation in concentrations (Figure 2B).

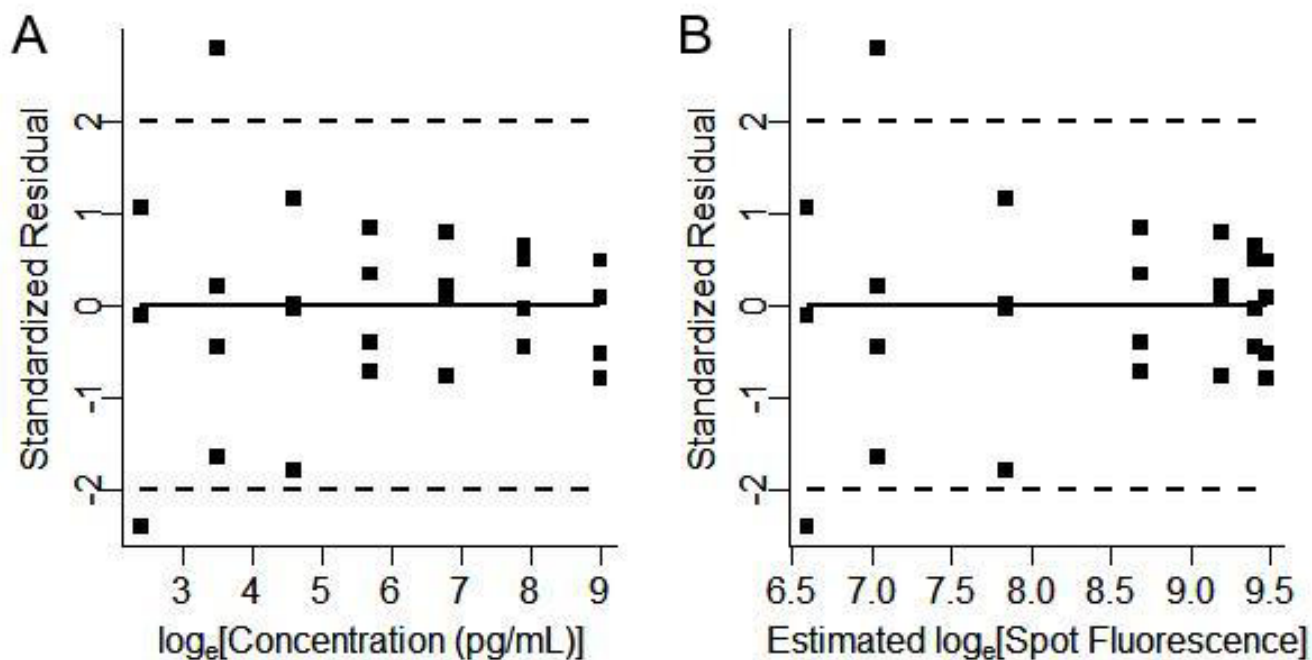


Figure 4
Modeling diagnostics The points in graphs of the standardized residuals versus the \log_e standard concentrations (A) and the standardized residuals versus the estimated \log_e spot intensities (B) are reasonably well behaved, showing no strong systematic trends or deviations from the zero line. These indicate that the four-parameter logistic model is acceptable, that subsequent statistical inferences are reasonable, but that a better model may exist in another model family.

With the heteroskedasticity and undue leverage addressed, we estimated a standard curve by selecting one of two models: a four-parameter logistic model (Eqn. 1) and a power curve model (Eqn. ??). We chose the logistic model as the model that fits the data best visually and in terms of the PRESS statistic (Figure 3). The logistic curve more closely follows the data points, while the power curve is too high in the lower concentrations and too low in the higher concentrations. We confirmed our choice with a review of the modeling diagnostics. In this case, we examined graphs of the standardized residuals as a function of concentration and the estimated intensities (Figure 4). In both graphs, the residuals show no significant systematic trends or deviations from the zero line and vary uniformly. Further, the preponderance of standardized residuals falls between -2 and 2, indicating that a statistical interpretation of the 95% confidence intervals is warranted.

Figure 5 presents the three-panel diagnostic visualization for the HER-2 data. HER-2 belongs to the family of epidermal growth factor receptors and has been used as a serum biomarker for the detection of breast cancer. This figure illustrates how data from a large study measuring HER-2

levels in the serum of women with and without breast cancer can be visualized using this statistical approach. A standard curve of HER-2 was generated, and the concentration of HER-2 in 39 samples was determined. To estimate manually the concentration for a sample HER-2 spot intensity, say I , locate I on the vertical axis, then scan across horizontally to the standard curve and 95% confidence intervals (Figure 5A). Scan down from these points to find the appropriate estimated concentration and lower and upper 95% concentration confidence bounds. In this manner, the estimated concentration and confidence interval can be determined.

In the standard curve panel (Figure 5A), we see that near the asymptotes of the standard curve, the uncertainty grows much more quickly than the curve, causing the concentration confidence bounds to diverge. Although this divergence is due to the approximation (Eqns 2 and 3), it is true that near the asymptotes, the uncertainty of the estimated concentration increases greatly. For this reason we have defined our optimal region of this curve to be the range of spot fluorescence values such that both the upper and lower bounds are monotonically increasing in intensity. The boundaries of this range are indicated by the

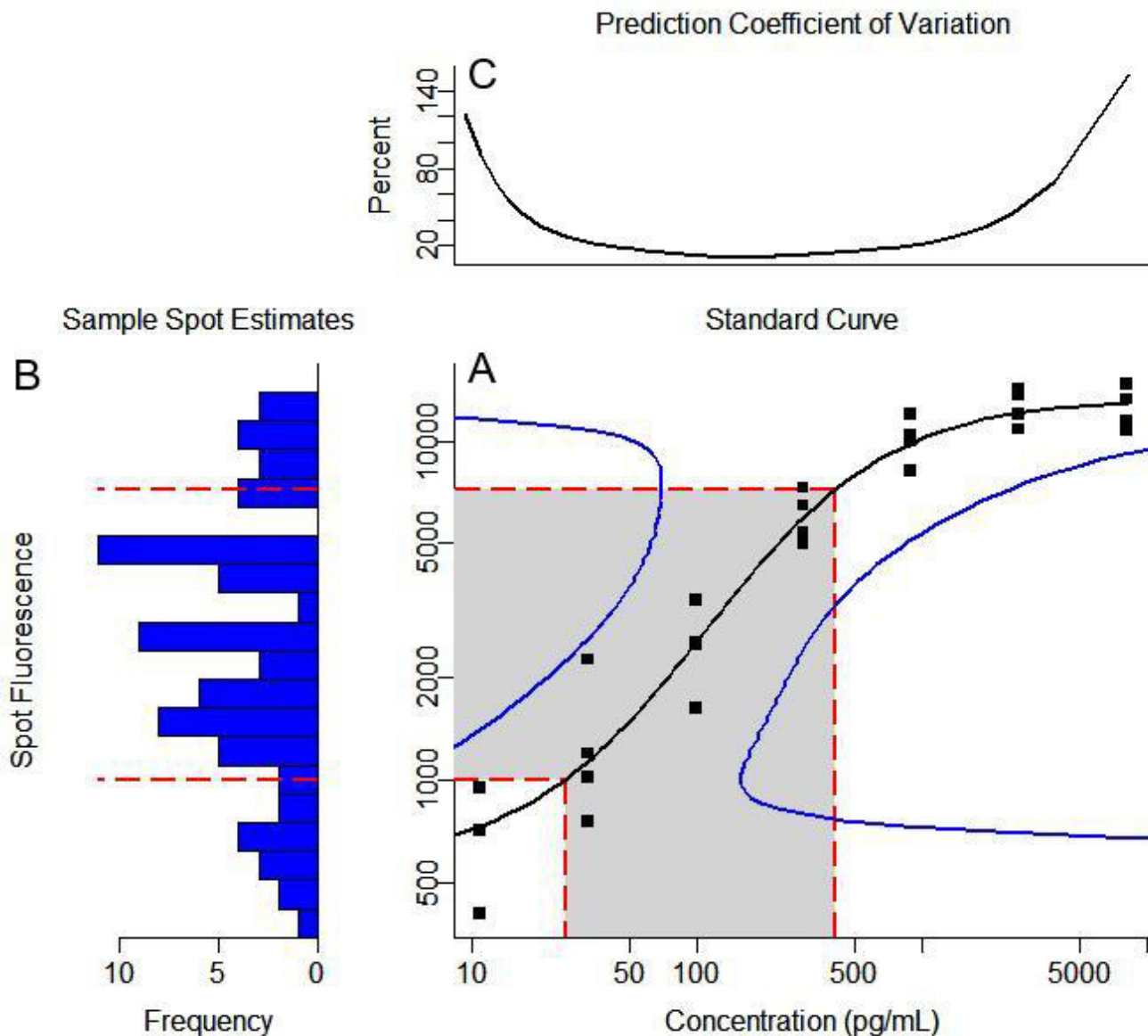


Figure 5

Three-panel diagnostic summary The standard curve panel (A) of the three-panel diagnostic summary features a scatter-plot of the data, the estimated standard curve (black line), the 95% confidence intervals (blue lines), and the region of acceptable errors (grey). For this example, the acceptable segment of the standard curve (i.e., the segment with concentration estimation errors acceptable to us) covers spot intensities from about 1000 to 7500 intensity units and concentrations from about 25 to 500 pg/ml. The histogram of sample intensities (B) suggests that about one-fifth of the sample spot intensities is below the detection limit of about 1000 units, about three-fifths will produce estimated concentrations in the acceptable range, and one-fifth exceeds the acceptable range. With regard to remeasuring the standards in this experiment or imaging in future experiments, the three-panel graph suggests that it may be worthwhile to attempt to extend the acceptable range. It also suggests that the normalization between the standards measurements and sample measurements should be revisited. The graph of the concentration coefficient of variation as function of concentration (C) offers an alternative summary of the estimation errors over the range of concentrations.

shaded area and the dashed red lines, which also show the concentration values corresponding to the acceptable fluorescence range. Our optimal concentration range spans approximately two orders of magnitude.

The histogram (Figure 5B) shows the fluorescence values of the sample spots for which this curve may best be used to estimate concentration. The plot shows that many of the sample values lie outside our optimal region. The researcher must then decide if too many of these values lie outside this range and, if so, what can be done to fix this problem. Nevertheless, we were able to compare the HER-2 concentrations and found a 3.5-fold increase in HER-2 protein levels in women with stage III/stage IV breast cancer (7 samples) compared to women without breast cancer (12 samples).

The concentration coefficient of variation, or the ratio of the concentration estimation error to the corresponding estimated concentration, offers an alternative expression to a confidence interval as a means to evaluate concentration estimation error. A graph of this estimation error as a function of concentration offers a comprehensive summary of the variation in the concentration coefficient of variation over the concentration range (Figure 5C).

Presenting the results in this type of plot allows us to immediately look for several potential problems. First, does the fitted curve seem reasonable, given the data points to which we are fitting? We also can determine whether most of the unknown sample data fall within the acceptable range of the curve. The usable concentration range is made clear and, if it is too limited in range, it is immediately apparent. If problems are identified, several fixes are available, including changing the settings on the imager or using a different concentration range to create the standard curves.

Conclusions

Evaluation of errors in estimating concentrations is important to establishing confidence in protein concentration estimates. Propagation of error provides a straightforward approach to estimating concentration estimation errors in ELISA microarray experiments. When presented in a simple multi-panel visualization, the propagated errors provide valuable information about individual concentration estimates, the applicability of the estimated standard curve, quality of the experiment, and the conduct of the ELISA microarray processing. The visualization provides a rapid assessment of the quality of the data, particularly in regard to the goodness of fit of the estimated standard curve and its capability to estimate concentrations over the observed range of intensities of biological samples.

Authors' contributions

DSD framed evaluation of ELISA microarray concentration estimation errors as a statistical problem. KKA suggested the use of propagation of error and outlined the initial derivation. DSD and AMW derived the appropriate propagation of error equations, developed the algorithms, and designed the diagnostic visualizations. AMW encoded the algorithms, analyzed the microarray imagery, and produced the statistical results. SMV designed and printed the arrays and then carried out the ELISA microarray experiments. RCZ conceived the study, designed it with SMV, and then coordinated all our efforts. All authors submitted comments on drafts, then read and approved the final manuscript.

Acknowledgements

This research was supported by the Biomolecular Systems Initiative at the Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830.

References

1. Zangar RC, Varnum SM, Covington CY, Smith RD: **A rational approach for discovering and validating cancer markers in very small samples using mass spectrometry and ELISA microarrays.** *Disease Markers* 2004 in press.
2. Varnum SM, Woodbury RL, Zangar RC: **A protein microarray ELISA for screening biological fluids.** *Methods Mol Biol* 2004, **264**:161-172.
3. Woodbury RL, Varnum SM, Zangar RC: **Elevated HGF levels in sera from breast cancer patients detected using a protein microarray ELISA.** *Journal of Proteome Research* 2002, **1**:233-237.
4. Racine-Poon A, Weihs C, Smith AFM: **Estimation of relative potency with sequential dilution errors in radioimmunoassay.** *Biometrics* 1991, **47**:1235-1246.
5. Giltinan DM, Davidian M: **Assays for recombinant proteins: a problem in non-linear calibration.** *Statistics in Medicine* 1994, **13**:1165-1179.
6. Jones G, Wortberg M, Kreissig SB, Hammock BD, Rocke DM: **Sources of experimental variation in calibration curves for enzyme-linked immunosorbent assay.** *Analytica Chimica Acta* 1995, **313**:197-207.
7. Fossceco SL, Curtis NA: **Exploring enzyme-linked immunosorbent assay (ELISA) data with the SAS Analyst Application.** In *Proceedings of the Twenty-Fourth SAS Users Group International Conference* Edited by: Lajiness M. SAS Institute, Inc. (Cary, NC); 1999:1635-1643.
8. Lee Mt, Kuo Fc, Whitmore GA, Sklar J: **Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations.** *PNAS* 2000, **97**(18):9834-9839.
9. Stoekert CJ Jr, Causton HC, Ball CA: **Microarray databases: standards and ontologies.** *Nature Genetics* 2002, **32**:469-473.
10. Ball CA, Sherlock G, Parkinson H, Rocca-Sera P, Brooksbank C, Causton HC, Cavalieri D, Gaasterland T, Hingamp P, Holstege F, Ringwald M, Spellman P, Stoekert CJ Jr, Stewart JE, Taylor R, Brazma A, Quackenbush J: **Standards for microarray data.** *Science* 2002, **298**(5593):539. Also published simultaneously in *Bioinformatics and The Lancet*.
11. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoekert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M: **Minimum information about a microarray experiment (MIAME): toward standards for microarray data.** *Nature Genetics* 2001, **29**:365-371.
12. Brazma A: **On the importance of standardisation in life sciences.** *Bioinformatics* 2001, **17**(2):113-114.

13. Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19(2)**:185-193.
14. Edwards D: **Non-linear normalization and background correction in one-channel cDNA microarray studies.** *Bioinformatics* 2003, **19(7)**:825-833.
15. Yang YH, Buckley MJ, Dudoit S, Speed TP: **Comparison of methods for image analysis on cDNA microarray data.** *Journal of computational and graphical statistics* 2002, **11**:108-136.
16. Rocke DM, G J: **Optimal design for ELISA and other forms of immunoassay.** *Technometrics* 1997, **39**:162-170.
17. Sittampalm GS, Smith WC, Miyakawa TW, Smith DR, McMorris C: **Application of experimental design Techniques to optimize a competitive ELISA.** *Journal of Immunological Methods* 1996, **190**:151-161.
18. Chatterjee S, Hadi AS: *Sensitivity analysis in linear regression* New York, New York: John Wiley and Sons; 1988.
19. Cook RD, Weisberg S: *Residuals and Influence in Regression* New York, New York: Chapman and Hall; 1986.
20. Milliken GA, Johnson DE: *Analysis of Messy Data* New York, New York: Van Nostrand Reinhold; 1984.
21. Robard D: **Statistical quality control and routine data processing for radioimmunoassays and immunoradiometric assays.** *Clinical Chemistry* 1974, **20(10)**:1255-1270.
22. Cleveland WS: **Robust locally weighted regression and smoothing scatterplots.** *Journal of the american statistical association* 1979, **74(368)**:829-836.
23. Venables WN, Ripley BD: *Modern Applied Statistics with S-Plus* New York, New York: Springer-Verlag; 1996.
24. Larholt KM, Sampson AR: **Effects of heteroscedasticity upon certain analyses when regression lines are not parallel.** *Biometrics* 1995, **51**:731-737.
25. Bunch DS, Rocke DM, Harrison RO: **Statistical design of ELISA protocols.** *J Immunol Methods* 1990, **132**:245-247.
26. Pinheiro JC, Bates DM: *Mixed-effects models in S and S-plus* New York, New York: Springer-Verlag; 2000.
27. Yeargers EK, Shonkwiler RV, Herod JV: *An introduction to the mathematics of biology: with computer algebra models* Boston, Massachusetts: Birkhauser; 1996.
28. Neter J, Wasserman W, Kutner KH: *Applied linear regression models* Boston, Massachusetts: Richard D Irwin; 1989.
29. Box GEP, Jenkins GM, Reinsel GC: *Time Series Analysis: Forecasting and Control* New York, New York: Prentice Hall; 1994.
30. Myers RH: *Classical and modern regression with applications* Boston, Massachusetts: PWS-KENT Publishing Company; 1990.
31. Mood AM, Graybill FA, Boes DC: *Introduction to the theory of statistics* New York, New York: McGraw-Hill Book Company; 1974.
32. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *Journal of the Royal Statistical Society, Series B, Methodological* 1995, **57**:289-300.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

