Animal Nutrition 10 (2022) 294–304

Contents lists available at ScienceDirect

Animal Nutrition

journal homepage: http://www.keaipublishing.com/en/journals/aninu/

Original Research Article

Abductive statistical methods improve the results of calibration curve bioassays: An example of determining zinc bioavailability in broiler chickens☆

Gene M. Pesti [a, b, *], Lynne Billard [b], Shu-Biao Wu [a], Robert A. Swick [a], Thi Thanh Hoai Nguyen [a], Natalie Morgan [a]

[a] The University of New England, Armidale, NSW, 2351, Australia
[b] The University of Georgia, Athens, GA, 30602, USA

ARTICLE INFO

Article history:
Received 10 August 2021
Received in revised form
31 December 2021
Accepted 22 April 2022
Available online 28 April 2022

Keywords:
Calibration curve
Standard curve
Confidence interval
Bioavailability

ABSTRACT

In this paper, we discuss the theory behind calibration curve experiments and their application to a zinc (Zn) bioavailability study with broiler chickens. Seven replicates of 16 male commercial broiler chicks were fed starter diets for 14 days. Six diets had different levels of a potential Zn source and one was a positive control with standard industry levels of Zn for comparison. Four commonly used methods of calculating bioavailability means and confidence intervals (CI) from a calibration curve (standard curve) experiment to estimate the bioavailability of a new zinc source in broiler chickens were compared. The methods compared were the following: 1) the Counter-Intuitive Method uses a multiple-range test to compare unknown test and standard samples; 2) the Intuitive Method uses standard linear regression and inverts the equation to predict Zn bioavailability for each replicate of test samples; 3) the Abductive Method uses Graybill's Equation, based on theory and observation, to estimate CI's; and 4) the Sophistic Method uses reverse regression, and calculates Zn bioavailability values directly from the equation. The Counter-Intuitive Method only gives information about which standards the test samples are, or are not, significantly different from respectively (average available Zn not predicted). The Intuitive Method ignores error about the standard curve and theoretically cannot estimate the CI directly ($\overline{X}$±SEM = 107.5 ± 15.8 mg Zn/kg). The Sophistic Method underestimates and overestimates the test sample mean values above and below the mean of the standards, respectively ($\overline{X}$ = 96.6 mg Zn/kg). The Abductive Method has an advantage over the other methods: The mean prediction estimation is consistent with theory (107.5 ± 6.1 mg Zn/kg; $\overline{X}$±SEM). When test or "unknown" samples are near the mean of the standard samples, the CI is smaller than when near the extremes of the calibration curve. When calibration curve error is small (R^2 > approximately 0.95), there is little advantage to using the Abductive Method, but when calibration curve error is larger, as in many bioassays with growing animals, the Abductive Method improves the accuracy of the CI calculations. The Abductive Method was used to demonstrate the influence of the number of replicate samples on experimental power and cost.

© 2022 Chinese Association of Animal Science and Veterinary Medicine. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Progress in all of biology, and especially in many applied biological fields such as agriculture and nutrition, has been largely dependent on progress in the fields of analytical chemistry and experimental statistics. The application of analytical chemistry to biology mainly uses a seemingly simple technique called the standard curve, or calibration curve. The calibration curve technique starts with a series of samples of known composition (independent or predictor variable X). It compares test samples of

☆ Mention of trade names or commercial products, proprietary produce or vendor in this publication does not imply endorsement by the authors, nor criticism of similar products not mentioned.

* Corresponding author.
E-mail address: gpesti@uga.edu (G.M. Pesti).

Peer review under responsibility of Chinese Association of Animal Science and Veterinary Medicine.

Production and Hosting by Elsevier on behalf of KeAi

https://doi.org/10.1016/j.aninu.2022.04.008

2405-6545/© 2022 Chinese Association of Animal Science and Veterinary Medicine. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

unknown composition (dependent variable, Y) to the standard samples to estimate the composition of the test samples. The calibration curve problem involves finding the variation or confidence interval (CI) around the predicted composition of a test sample's mean.

Interestingly, and perhaps the source of some confusion, calibration (or standard) curve methodology is practically always modeled as a straight line, not a curve. The levels of the property being estimated in the unknown sample(s), are under the control of the researcher (Kutner et al., 2005). The levels are chosen so that the response being measured is linear, and linearity is checked. The researcher chooses the extreme levels of X and spacing of intermediate ones as well as the amount of replication at each level to assure linearity. The unknown samples can be diluted or concentrated to fall as close to the center of the range of known samples as possible (where the CI for the line is minimal). The estimated CI of the unknown samples is used as a measure of quality control. If the CI of any unknown sample exceeds some predetermined level, the sample will be re-analyzed.

The calibration curve problem could probably be more appropriately called the calibration curve dilemma, because the literature offers seemingly no one exact or perfect answer as to how to calculate the confidence interval of the composition of the test samples. It seemingly involves a sometimes difficult choice. Eisenhart (1939) summarized the theory involved in choosing the right mathematical method to evaluate regression data. He emphasized the importance of understanding experimental design and its purpose in choosing the correct statistical method. For instance, when evaluating calibration curve studies with responses (Y_i's) dependent on different levels of input (X_p's), the relationship to be evaluated should have the response as the (dependent) variable, dependent on chosen levels of the independent variable ($Y_i = \beta_0 + \beta_1 X_i$), where β_n are constants. A different mathematical evaluation could be applied to dose–response trials where the responses become the independent variable used to predict the resulting appropriate dose ($X_i = \beta_0 + \beta_1 Y_i$). The resultant regression equations differ as shown in Fig. 1.

Eisenhart (1939) concluded in his paper: "Closer cooperation is possible between the practical man and the statistical theorist when the latter fully appreciates the problems of the former, and when the former, in turn, understands the methods advocated by the latter". We believe closer cooperation is not only possible but practically necessary to advance many areas of science.

Our goal is to bring clarity to the calibration curve situation by considering the assumptions behind the methods suggested along with comparisons. By so doing, the dilemma should disappear as we demonstrate what should be a straight-forward solution to the problem. The example used to illustrate the appropriate method for calculating means and confidence intervals from calibration curve bioavailability data is from a trial to evaluate zinc (Zn) bioavailability from a new chemical compound by broiler chickens. This paper may serve as a bridge to help "the practical man and the statistical theorist" (Eisenhart, 1939) understand each-other's perspectives.

1.1. Regression models

1.1.1. Standard model-predicting response Y_0 for a given X_0

Let us consider first the standard regression model. Suppose there is one independent variable X (also called regression or predictor variable) and suppose the response variable is Y; generalization to p independent variables $X_1, \ldots, X_p$ follows readily. We have observations Y_i, X_i ($i = 1, \ldots, n$). Then, the standard linear regression model is the following:

$$Y_i = \beta_0 + \beta_1 X_i + e_i, \tag{1}$$

where e_i are independent error terms with $E(e_i) = 0$ and variance, $Var(e_i) = \sigma^2$. For a given set of observations Y_i, X_i ($i = 1, \ldots, n$), the model parameters (β_0, β_1) can be estimated. The least squares estimators are

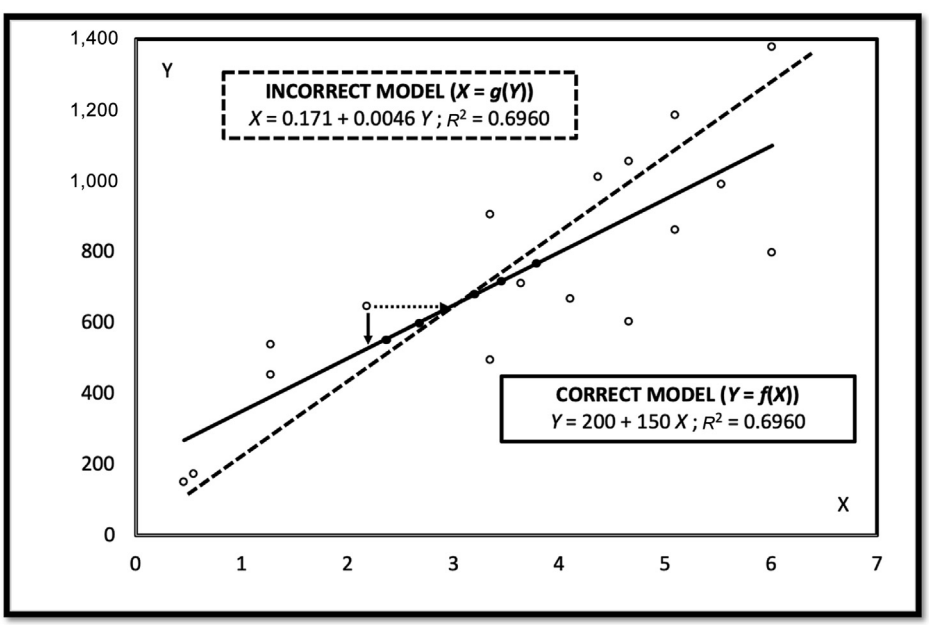

Fig. 1. Fitting of artificial X–Y data showing fits to linear increments. The solid line is the correct fit for calibration curve problems with $Y = f(X)$. The dashed line is the incorrect fit for calibration curve problems with $X = g(Y)$.

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i X_i - n\overline{Y}\overline{X}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})(X_i - \overline{X})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}, \quad \widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\overline{X}, \tag{2}$$

where $\overline{Y} = \sum_i \frac{Y_i}{n}$ and $\overline{X} = \sum_i \frac{X_i}{n}$ are the sample means. Then, for each X_i, the estimated predicted response is

$$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i \quad (i = 1, \ldots, n). \tag{3}$$

From the squared residuals (SSE, $\sum_i R_i^2$ where $R_i = Y_i - \widehat{Y}_i$), an estimate of the underlying variance σ^2 can be found, i.e.,

$$SSE = \sum_{i=1}^{n} R_i^2 = \sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2, \quad \widehat{\sigma}^2 = SSE/n - 2. \tag{4}$$

If it is assumed that the error e_i $(i = 1, \ldots, n)$ are independent and normally distributed, and hence the Y_i $(i = 1, \ldots, n)$ are also normally distributed, then the least square estimators of Eq. (2) are maximum likelihood estimators. See any of the many elementary text on regression models, e.g., Kleinbaum et al. (1988), Draper and Smith (1981), Montgomery et al. (2001). Henceforth, we assume the data (Y_i's) are normally distributed; this assumption allows the calculation of confidence intervals.

The estimated regression parameters and the estimated responses, i.e., $\widehat{Y}_i$, $\widehat{\beta}_0$ and $\widehat{\beta}_1$, are random variables, because they are functions of the responses Y_i. Recall here that the independent variables X_i are fixed values, i.e., they are not random variables, and so do not have a distribution (unless we are dealing with structural models which is quite different from the current situation). These estimators are unbiased, in that, $E(\widehat{\beta}_0) = \beta_0$, $E(\widehat{\beta}_1) = \beta_1$ and $E(\widehat{Y}_i|X_i) = \beta_0 + \beta_1 X_i$ for given X_i. Note that these $\widehat{Y}_i$ values fall along the regression line, for each respective X_i. The variances are the following, respectively,

$$Var(\widehat{\beta}_0) = \frac{\sigma^2}{n}\left[\frac{\sum_{i=1}^{n} X_i^2}{\sum_{i=1}^{n}(X_i - \widehat{X})^2}\right], \quad Var(\widehat{\beta}_1) = \sigma^2\left[\sum_{i=1}^{n}(X_i - \overline{X})^2\right]^{-1}. \tag{5}$$

Hence, the variance of the predicted response of an individual observation $\widehat{Y}|X = X_0$, at a given X_0 is the following:

$$Var(\widehat{Y}_0) = \sigma^2\left[1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \widehat{X})^2}\right]^{1/2}. \tag{6}$$

When the errors are normally distributed, $\widehat{\beta}_0$ and $\widehat{\beta}_1$ and hence $\widehat{Y}_0$ are normally distributed. Thus, the confidence intervals (CI) for these predicted responses can be calculated from

$$\mathrm{CI} = \widehat{Y}_0 \pm t_\nu \widehat{\sigma}\left[1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \widehat{X})^2}\right]^{1/2}, \tag{7}$$

where $\nu = n - 2$ is the degree of freedom for the t_ν distribution and $\widehat{\sigma}^2$ is the estimated error variance given in Eq. (4). If we want to predict the average of k (say) responses at a given X_0, then Eq. (7) is replaced by

$$\mathrm{CI} = \widehat{Y}_0 \pm t_\nu \widehat{\sigma}\left[\frac{1}{k} + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \widehat{X})^2}\right]^{1/2}. \tag{8}$$

Notice that as a given X_0 moves away from the $\overline{X}$, the variances increase and hence the CIs curves diverge.

1.1.2. Calibration problem-what independent X_0 gave known responses Y_0?

Many practical researchers will recognize that their standard curve assay is a regression problem and apply a simple linear model. The technique is called "Inverse prediction" by Kutner et al. (2005). It can also be called the Intuitive (instinctive, untaught) method. Animal nutritionists, analytical chemists and many other scientists start by estimating β_0 and β_1 using ordinary least squares methods. They understand that the responses are dependent on the standards (predictor or independent variable). The fitted equation is as follows:

$$Y = \widehat{\beta}_0 + \widehat{\beta}_1 X. \tag{9}$$

Then, to find that unknown X_0 value for their Y_0 value, Eq. (9) is turned around, inverted, to obtain an estimated dependent value $\widehat{X}_0$ as follows:

$$\widehat{X}_0 = (Y_0 - \widehat{\beta}_0)/\widehat{\beta}_1. \tag{10}$$

This approach was first suggested by Krutchkoff (1967). Although Williams (1969) showed why this was not valid, its use has unfortunately continued to recent times (e.g. Parker et al., 2010). For each of the m replicate's responses, many practical researchers then estimate a separate X_0. Then they average the m estimates of X_0 to get the correct $\overline{X}$. The $\widehat{\sigma}^2$ is then calculated from the m $\widehat{X}_0$ values, not the one in Eq. (4). The last step is to calculate the CI as they always have with m observations to obtain

$$\mathrm{CI} = \widehat{X_0} \pm t_\nu \widehat{\sigma} \Big/ m^{1/2} \tag{11}$$

as is illustrated by Fig. 2. More critically, this result Eq. (11) assumes that $Var(\widehat{X}_0) = \widehat{\sigma}^2/m$, which is clearly incorrect, as explained in Section 1.3. This approach estimates a separate X_0 for each of m dependent responses Y. However, in contrast, the problem at hand is that these m Y values occur for the same unknown X_0 value (Fig. 3).

1.2. Graybill's calibration method

Researchers may recognize that there is variation in both the calibration curve and between replicate samples, and find an abductive (based on theory and observation) calibration method. One such method was described by Graybill (1976): Suppose now that in addition to the observations Y_i, X_i $(i = 1, \ldots, n)$, of Section 1.1, we have m responses Y_i $(i = n + 1, \ldots, n + m)$, for an unknown independent value X_0. The model of Eq. (1) becomes

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + e_i, \quad (i = 1, \ldots, n). \\ Y_i &= \beta_0 + \beta_1 X_0 + e_i, \quad (i = n + 1, \ldots, n + m). \end{aligned} \tag{12}$$

Then, the first n observations are used to find the estimators ($\widehat{\beta}_0$, $\widehat{\beta}_1$) as given in Eq. (2). The estimated value for the unknown X_0, provided that $\widehat{\beta}_1 \neq 0$, is

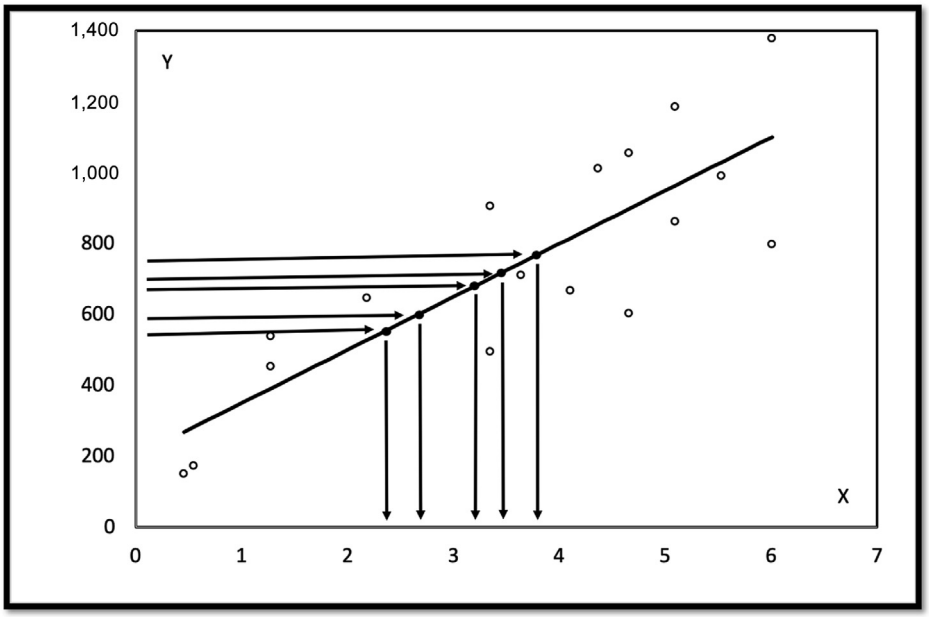

Fig. 2. Fitting of artificial X–Y data showing fits to linear increments. The solid line is the correct fit for calibration curve problems with $Y = f(X)$. The arrows show how a distinct X is found from responses from each replicate Y (unknown test sample replicates). Confidence intervals are calculated from the individual predicted values of X.

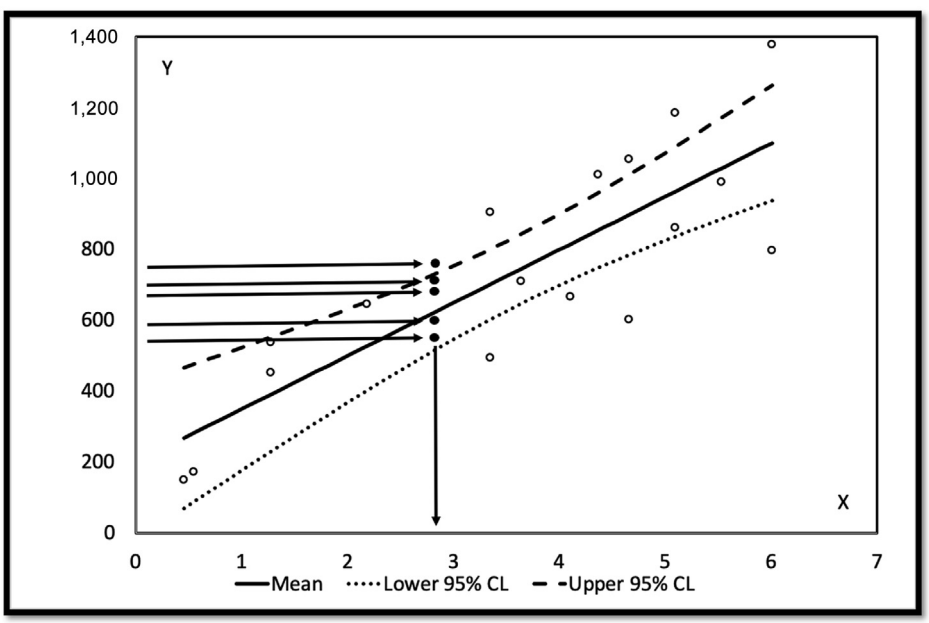

Fig. 3. Fitting of artificial X–Y data showing fits to linear increments. The solid line is the correct fit for calibration curve problems with $Y = f(X)$. The dashed and dotted lines represent the confidence interval for the line. The arrows show how one X value is found from responses from each replicate Y (unknown test sample). Confidence intervals are calculated from Graybill's Abductive Method (Eq. (14)).

$$\widehat{X}_0 = \frac{\overline{Y}_0 - \widehat{\beta}_0}{\widehat{\beta}_1} = \overline{X} + \frac{\overline{Y}_0 - \overline{Y}}{\widehat{\beta}_1}, \overline{Y}_0 = \frac{1}{m}\sum_{i=n+1}^{n+m} Y_i, \tag{13}$$

which is the maximum likelihood estimator and is often referred to as the classic estimator. The error variance σ^2 is now estimated by $\widehat{\sigma}_c^2$ (say) using all $n + m$ observations, i.e.,

$$\widehat{\sigma}_c^2 = \frac{1}{n+m-3}\left[\sum_{i=n}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_1)^2 + \sum_{i=n+1}^{n+m}(Y_i - \widehat{Y}_0)^2\right]. \tag{14}$$

Notice this differs from the estimated error variance $\widehat{\sigma}^2$ of Eq. (4) unless $m = 1$ when the second term on the right-hand-side of Eq. (14) is zero.

To calculate the CI around this estimated $\widehat{X}_0$ requires knowledge of its distribution and (estimated) variance. However, from Eq. (13), we see that $\widehat{X}_0$ is a function of the inverse of the normally distributed $\widehat{\beta}_1$ which inverse follows a Cauchy distribution whose mean and variance do not exist. One solution to this problem is to use the abductive method sometimes called

Graybill Calibration Method, e.g., Graybill (1961, 1976), also called the Inverse Method. Other approaches have been proposed in the literature, but unfortunately these can not be sustained theoretically, as we shall see.

1.3. Other approaches

Another approach offered in the literature (Krutchkoff, 1967; Parker et al., 2010) is the Sophistic (plausible but fallacious) Method, or Reverse Regression approach, based on a reverse regression model, as in

$$X_i = \beta_0^{(r)} + \beta_1^{(r)} Y_i + e_i^{(r)}, \tag{15}$$

where now the superscripts (r) have been used to distinguish between the standard and reverse regression models. The estimators of $\beta_0^{(r)}$ and $\beta_1^{(r)}$ are

$$\widehat{\beta}_1^{(r)} = \frac{\sum_{i=1}^n Y_i X_i - n\overline{Y}\overline{X}}{\sum_{i=1}^n Y_i^2 - n\overline{Y}^2} = \frac{\sum_{i=1}^n (Y_i - \overline{Y})(X_i - \overline{X})}{\sum_{i=1}^n (Y_i - \overline{Y})^2}, \quad \widehat{\beta}_0^{(r)} = \overline{X} - \widehat{\beta}_1^{(r)} \overline{Y}. \tag{16}$$

Fig. 1 compares the dashed regression equation obtained when fitting the model of Eq. (1), and the solid regression equation is fitting the model of Eq. (15). For the standard regression model, the least squared estimates of the parameters are found by minimizing the sum of the squared Y residuals $R^2 = (\widehat{Y}_i - Y)^2$ corresponding to the solid blue vertical distance between the Y_i and $\widehat{Y}_i$ in Fig. 1; whereas, for the reverse regression model, the estimators are found by minimizing the sum of the squared "reverse" X residuals $\sum_i (R^{(r)})^2 = \sum_i (X_i - \widehat{X}_i)^2$ corresponding to the dotted horizontal distances between the X_i and $\widehat{X}_i$ in Fig. 1.

If we continue with this model approach, we would obtain an estimate of the reverse error variance as

$$SSE^{(r)} = \sum_{i=1}^n \left(R_i^{(r)}\right)^2 = \sum_{i=1}^n (X_i - \widehat{X}_i)^2 = \sum_{i=1}^n \left(X_i - \widehat{\beta}_0^{(r)} - \widehat{\beta}_1^{(r)} X_i\right)^2, \quad \widehat{\sigma}^2 \widehat{\sigma}_{(2)}^2 = SSE^{(r)} \Big/ (n-2). \tag{17}$$

There are difficulties with this approach. One is that in the derivation of the parameter estimators of Eq. (16), there is an apparent but spurious implicit assumption that the errors $e_i^{(r)}$ are independent of the Y_i values, which violates the original assumption (that Y is the dependent variable). This implies that minimizing the sum of the squared "reverse" residuals $[R^{(r)}]^2$ is meaningless mathematically, because the X_i values have been pre-determined. Another problem is the obvious fact that the reverse regression equation is different from the real regression equation, as is evident from Fig. 1. The model assumptions apply to the model of Eq. (1) but not to the model of Eq. (15). Further, if Eq. (15) is used to estimate an unknown X_0 value for given Y_0 value in a bioassay, we would obtain

$$\widehat{X}_0^{(r)} = \beta_0^{(r)} + \beta_1^{(r)} Y_0. \tag{18}$$

Thence, if we ignore for the moment the mathematical difficulties, the CI for the estimated $\widehat{X}_0^{(r)}$ would follow (analogous to Eq. (7) but with X's and Y's reversed) as

$$\mathrm{CI} = \widehat{X}_0^{(r)} \pm t_\nu \widehat{\sigma}_{(r)} \left[1 + \frac{1}{n} + \frac{(Y_0 - \overline{Y})^2}{\sum_{i=1}^n (Y_i - \widehat{Y})^2}\right]^{1/2}. \tag{19}$$

Eq. (19) cannot be appropriate for calibration curve bioassays because $\widehat{X}_0^{(r)}$ is not an appropriate estimate of the mean, X_0, quite apart from the mathematical difficulties discussed above.

Krutchkoff (1967) and Osborne (1991) refer to this $\widehat{X}^{(r)}$ as an inverse estimator of X_0. Crucially, $\widehat{X}_0^{(r)} \neq \widehat{X}_0$ for the same Y_0 value. Furthermore, the CIs around these estimates are also different, because the estimate of the underlying error variance is based on the X residuals ($R^{(r)}$) in the former and on the Y residuals (R_i) in the latter. What is wanted is the expectation, $\widehat{X}_0$ of Eq. (13).

Advocates of this method for calibration curves point to the fact that it is relatively easy to implement, and that the estimator $\widehat{X}_0^{(r)}$ has a finite variance (compared to the infinite variance of $\widehat{X}_0$); even there, this smaller (mean square error, based on the X residuals) variance only holds when estimating X_0 from a single Y_0 value with this unknown X_0 being in a very narrow region around $\overline{X}$ and for large samples, but not otherwise. Also, these advocates completely overlook the fact that model assumptions are violated which has the consequence that answers are invalid. The fact that the assumptions are invalidated is hard to overlook; notice from Fig. 1: $Y = f(X)$ and $X = g(Y)$ give different lines despite having identical coefficients of variation. Furthermore, there is a large body of literature harshly critical of using $X = g(Y)$, primarily because it has no theoretical justification. In other words, using $X = g(Y)$ cannot be sustained particularly because of the violations of the basic model assumptions (e.g. Williams, 1969; Berkson, 1969; Montgomery et al., 2001).

To help unravel the confusing terminology, we observe that the so-called classic estimator $\widehat{X}_0$ of Eq. (13) is calculated using the calibration regression method (sometimes called the inverse method) of model Eq. (1); whereas the inverse estimator $\widehat{X}_0^{(r)}$ of Eq. (18) is calculated from the reverse regression method using model Eq. (15).

1.4. Calibration problem—what is the CI for X_0 given known responses Y_0?

Let us return to the question of calculating a CI for the classic estimate $\widehat{X}_0$ developed in Section 1.2. (the abductive method—based on Graybill's Method). In particular, it is difficult to obtain $Var(\widehat{X}_0)$. However, Graybill (1976) has shown that the CI is given by

$$\overline{X} + d_1 \le X_0 \le \widehat{X} + d_2, \tag{20}$$

where d_i $(i = 1, 2)$ are the roots of the quadratic equation,

$$d^2 \left[\widehat{\beta}_1^2 - \frac{t_\nu^2 \widehat{\sigma}^2}{\sum_{i=1}^n (X_i - X)^2}\right] - 2d\widehat{\beta}_1 (\overline{Y}_0 - \overline{Y}) + \left[(\overline{Y}_0 - \overline{Y})^2 - t_\nu^2 \widehat{\sigma}^2 \left(\frac{1}{m} + \frac{1}{n}\right)\right] = 0. \tag{21}$$

Montgomery et al. (2001) have the same CIs except they assume just one observation at Y_0 (i.e., $m = 1$), so that $\widehat{Y}_0 = Y_0$ and the degrees of freedom in the t-distribution are $\nu = n - 2$; whereas, Graybill (1976) has a more general m observations from this X_0 with $m \ge 1$ and uses σ_c^2 of Eq. (14) and so $\nu = n + m - 3$.

By using the so-called delta method (Miller, 1991; Parker et al., 2010; Casella and Berger, 2002) and abductive reasoning, an approximation is given by

$$Var(\widehat{X}_0) \approx \frac{\sigma^2}{\beta_1^2}\left[\frac{1}{m}+\frac{1}{n}+\frac{(Y_0-Y)^2}{\beta_1^2\sum_{i=1}^{n}(X_i-\overline{X})^2}\right]. \quad (22)$$

Hence, the CI is approximately

$$\widehat{X}_0 \pm t_v \frac{\sigma^2}{\widehat{\beta}_1}\left[\frac{1}{m}+\frac{1}{n}+\frac{(Y_0-\overline{Y})^2}{\beta_1^2\sum_{i=1}^{n}(X_i-\overline{X})^2}\right]^{1/2}. \quad (23)$$

This is called the Abductive Method.

1.5. Specific objectives

The analyses reported in this paper compare 4 statistical methods of evaluating bioavailability experiments and estimate the CI of a new source of zinc for growing broiler chickens: 1) the Counter-Intuitive Method compares the various points on the standard curve and test samples with a multiple range test (there is no estimate of the mean or CI of the test sample, only what standards the test sample is not different from); 2) the Intuitive Method uses Eq. (3) to determine the mean value of each test sample replicate, and the CI is calculated from those values; and 3) Graybill's Abductive Method uses Eq. (3) to estimate the mean bioavailability value, and Eq. (21) to calculate the CI; 4) the mean values from the Sophistic Reverse Regression Method in Eq. (18) were calculated to show the magnitude of differences between mean estimates under different circumstances.

2. Materials and methods

2.1. Animal ethics

All the experimental procedures applied in this study were reviewed and approved by the University of New England Animal Ethics Committee.

2.2. Birds and diets

A total of 784 Ross 308 male chicks at 1 d old were from a commercial hatchery (Darwalla Poultry Distributors Pty Ltd., Redland, Queensland, Australia). Chicks were weighed and randomly assigned to 7 dietary treatments, each replicated 7 times in floor pens, with 16 chicks per replicate. Basal wheat-soybean meal diets were formulated to meet or exceed the requirements for starter (0 to 14 d) (Table 1; Aviagen, 2019). The 7 dietary treatments consisted of the following: 1) a positive control diet (PC) with 50 mg/kg Zn as ZnO and 50 mg/kg Zn as $ZnSO_4$, 2) a negative control basal diet (NC) without any added Zn, and 3) to 7) 5 diets supplemented in basal diet with 20, 40, 60, 80 or 100 mg/kg of supplemental Zn as zinc hydroxychloride (Selko IntelliBond Zn, Trouw Nutrition, Netherlands).

On d 14, 3 birds per replicate were each given (gavage) 1 mL of fluorescein isothiocyanate-dextran solution (FITC-d) (100 mg MW 4000, Sigma Aldrich Co., Castle Hill, NSW, Australia). At 2.5 h after inoculation, the birds were stunned and decapitated. Right tibias were collected. All soft tissues and cartilage were removed before drying in an oven for 24 h at 105 °C (Qualtex Universal Series 2000, Watson Victor Ltd., Perth, Australia). The dried tibias were then ashed in a Carbolite CWF 1200 chamber furnace (Carbolite, Sheffield, UK). The ashing started at 300 °C and increased to 600 °C in the first hour. Samples were in the oven for a total of 6 h. The mineral contents of the tibia ash and diet samples were determined by inductively coupled plasma-optical emission spectrometer (ICP-OES) (Agilent, Mulgrave, Victoria, Australia). Briefly, 0.1 g subsamples were weighed in Teflon tubes (Milestone, Sorisole, Italy) and then subjected to digestion in 1 mL distilled water and 4 mL concentrated HCl (70%) in an Ultrawave Microwave Digestion system (Milestone, Sorisole, Italy) for 45 min. The solution was cooled to room temperature and quantitatively transferred into a 30-mL volumetric flask. The solution was made to 25 mL total volume with distilled water and mixed well for analysis of trace mineral concentration by the ICP-OES instrument.

Table 1
Diet composition of starter from d 0 to 14 (%, as-fed basis).

Ingredients	Content	Calculated nutrients	Content
Wheat	56.10	ME, kcal/kg	3,000
Soybean meal (dehulled)	29.70	Crude protein	23.96
Canola meal	5.63	Crude fat	4.42
Rice bran	3.87	Crude fiber	3.18
Canola oil	2.00	d Arg	1.33
Limestone	1.17	d Lys	1.24
Dicalcium phosphate[1]	0.38	d Met	0.53
Sodium chloride	0.17	d M + C	0.90
Sodium bicarbonate	0.12	Calcium	0.85
Mineral premix[2]	0.10	Phosphorus avail.	0.43
Vitamin premix[3]	0.09	Sodium	0.17
Choline chloride (60%)	0.06	Chloride	0.20
L-Lysine	0.20	Choline, mg/kg	1,600
D,L-Methionine	0.21	Linoleic, 18%:2%	1.32
L-Threonine	0.05		
Xylanase	0.02		
Phytase	0.01		
Total	100		

[1] Dicalcium phosphate contained: phosphorus, 18%; calcium, 21%.
[2] The Zn-free trace mineral concentrate supplied per kilogram of diet: Cu (sulfate), 16 mg; Fe (sulfate), 40 mg; I (KI), 1.25 mg; Se (Na selenate), 0.3 mg; Mn (sulfate and oxide), 120 mg; cereal-based carrier, 128 mg; mineral oil, 3.75 mg.
[3] Vitamin concentrate supplied per kilogram of diet: retinol, 12,000 IU; cholecalciferol, 5,000 IU; tocopheryl acetate, 75 mg, menadione, 3 mg; thiamine, 3 mg; riboflavin, 8 mg; niacin, 55 mg; pantothenate, 13 mg; pyridoxine, 5 mg; folate, 2 mg; cyanocobalamin, 16 μg; biotin, 200 μg; cereal-based carrier, 149 mg; mineral oil, 2.5 mg.

2.3. Statistical analyses

One-way ANOVA and Duncan's New Multiple Range Test were performed using SAS (SAS 9.4, Carey, N.C. 2008) with the following statements:

```
Data a; input Tibia Zn Treatment;
Datalines;
(…data lines…)
Proc glm; Class Treatment; model Tibia Zn = Treatment; means
Treatment/Duncan; run;
```

The remainder of the calculations were made using MicroSoft Excel for Mac Version 16.45 (Microsoft Corporation, Redmond WA USA). An Excel workbook, the Calculation Curve Confidence Calculator (CCCC.xlsx) is available from the authors.

2.4. Experimental power calculations

The m and n terms of Eq. (21) were varied together to predict the expected standard error of the mean when using different numbers of replicates for 6 levels of the calibration curve and one unknown test sample. Total costs were based on experiments conducted in Australia in 2020 and included housing costs at $1.28/pen per day, chick costs at $3.33/chick, feed costs at $0.86/chick, bedding costa

at $5.60/pen, labor at $11.70/pen, transportation at $229/experiment and miscellaneous at $103/experiment (gloves, containers, laundry), tibia analyses at $6.00/sample and 3 samples/pen, and $6.00/diet.

3. Results

The Counter-Intuitive Method results showed that the Test Sample results were not different from Standard Sample results from 51 to 136 mg/kg of diet (Table 2). The Intuitive and Abductive methods estimated the same mean Zn value for this sample, but the CI for the Abductive method was much smaller (Table 3). The mean estimate using the Sophistic Method was less than the others, because this Test Sample had responses above the average for the Zn levels in the calibration curve (Table 3, Fig. 4). The CI for the calibration curve was smallest at the average standard level, resulting in different SD's of the predicted values using the Abductive Method (Fig. 5), but not the methods.

When the test sample values were adjusted (by adding the same value to each replicate) to have the same mean response as the average of the Zn standards, the predicted Zn values were identical for the Intuitive and Abductive methods (83.859; Table 4). This is the single point where the calibration curves cross (Figs. 1 and 4). When the test sample values were adjusted to have the same mean response as the upper extreme of the Zn standards (worst case scenario), the SEM of the test sample by the Abductive Method increased compared to the best case scenario (6.764 versus 5.926), but was still much smaller than for the Intuitive Method (6.764 versus 17.009; Table 4). Descriptive statistics not dependent on the SD of the test sample were the same for the best case scenario (average test sample response at mean of the Zn Standards) versus worst case scenario (average test sample response at the extreme of the Zn Standards) for Intuitive and Abductive Methods.

The experimental power estimations showed different responses if the samples were centered in the calibration curve or at the ends of the standard curve (Table 5). The SDs and SEMs exhibited a diminishing returns function to increasing number of replicates: Adding one to the number of test sample replicates from 2 to 3 decreased the SEM from 28.82 to 17.61 ($\Delta = 11.21$), increasing replicates from 3 to 4 decreased the SEM from 17.61 to 13.10 ($\Delta = 4.51$), etc.

4. Discussion

4.1. Statistical theorist perspective

The Intuitive Method which gives Eq. (11) and the Sophistic Reverse Regression Method which gives Eq. (19) are used extensively in scientific practice. However, these are mathematically and statistically incorrect. No amount of scientific usage, no matter how extensive that usage might be, can ever justify the use of inaccurate/incorrect results. The underlying assumptions of the experiment itself dictate how the predicted responses and model parameters (β_0, β_1, σ^2) are to be estimated. These values in turn mathematically dictate how to estimate a response $\widehat{Y}_0$ for a given X_0, or how to estimate an unknown X_0 for m given responses Y_0 at that X_0. These allow for the derivation of the relevant variances and hence the CIs. For the particular problem motivating the present work, there are m dependent values for which the single independent X_0 value is unknown. Therefore, the CI is given by Eq. (23).

4.1.1. The importance of applying the right theoretical math to each problem?

Calibration curve experiments should be designed with the premise that the responses of the unknown test samples are dependent on levels of the independent predictor variable (that is known to be without error). If the values of X have error then it is a different problem and neither line in Fig. 1 is appropriate to use (Tellinghuisen, 2010). Improved methods of fitting regressions to data with variation in both X and Y have been developed (Tellinghuisen, 2010) and may prove helpful for estimating CI's in the future for certain types of regression problems.

4.1.2. To minimally understand the methods advocated by the statistical theorist

There are 2 fundamental mathematical/statistical concepts that the practical researcher needs to firmly grasp to apply the proper analyses to any research outcome. The first concept is the difference between continuous and discrete variables. When dealing with calibration curves, there is no need to test if any test sample is different from any of the standard samples as is done with discrete variables and one-way ANOVA. The standard curve should be in a linear range of X and Y and each level of X results in a different level of Y (extensions to non-linear models) should follow the same principles. There is simply no need to determine if Y_1 is significantly different from Y_2, their difference is an assumption of regression. The second concept is that the predictor, or independent variable (the standards), is assumed to be without error. Only the response variables have error as in any standard regression model (see Analytical Methods Committee (2006) for clarifying examples).

4.2. Practical researcher perspective

Practical researchers will apply the most familiar statistical models to their data that they understand. They commonly apply analysis of variance and mean separation techniques to their experimental data when comparing the responses to various inputs like environmental qualities including dietary nutrient levels, genetic differences, and so on. When faced with a calibration curve problem, some naturally want to apply some form of multiple range testing to see which standards and unknown test samples are not significantly different. When bar graphs are used to represent the standard curve with individual CI's on each bar, it is clear that the practical researcher views the appropriate model to be a one-way

Table 2
One-way analysis of variance for the standard curve and an unknown zinc content test sample for estimating the zinc contents of broiler chicken feed (mg/kg).

Dietary zinc	Mean tibia zinc	Standard deviation	Standard error	Duncan grouping[1]
32	367.0	13.2	5.0	c
51	427.8	20.4	7.7	b
74	425.5	13.0	4.9	b
97	435.4	11.1	4.2	ab
114	431.8	15.0	5.7	ab
136	450.1	15.4	5.8	a
Test sample	437.0	24.8	9.4	ab

[1] Duncan's New Multiple Range Test ($P < 0.05$).

Table 3
The descriptive statistics for the amount of zinc in a test sample estimated by 4 different statistical methods (mg/kg).

Method	Characterization	Upper 95% CL	Mean	Lower 95% CL	Confidence interval	SD	SEM
Counter-Intuitive	Multiple Range Test	?	?	?	>32	?	?
Intuitive	Classic Regression & Inverse Prediction	191.67	107.47	23.26	168.41	41.66	15.75
Sophistic	Reverse Regression & Direct Prediction	?	96.02	?	?	?	?
Abductive	Graybill's Equation	137.71	107.47	77.23	60.48	14.96	6.11

CL = confidence limit; SD = standard deviation; SEM = standard error of the mean.
Question marks indicate the values are unknowable because they are not defined in the models by which they appear.

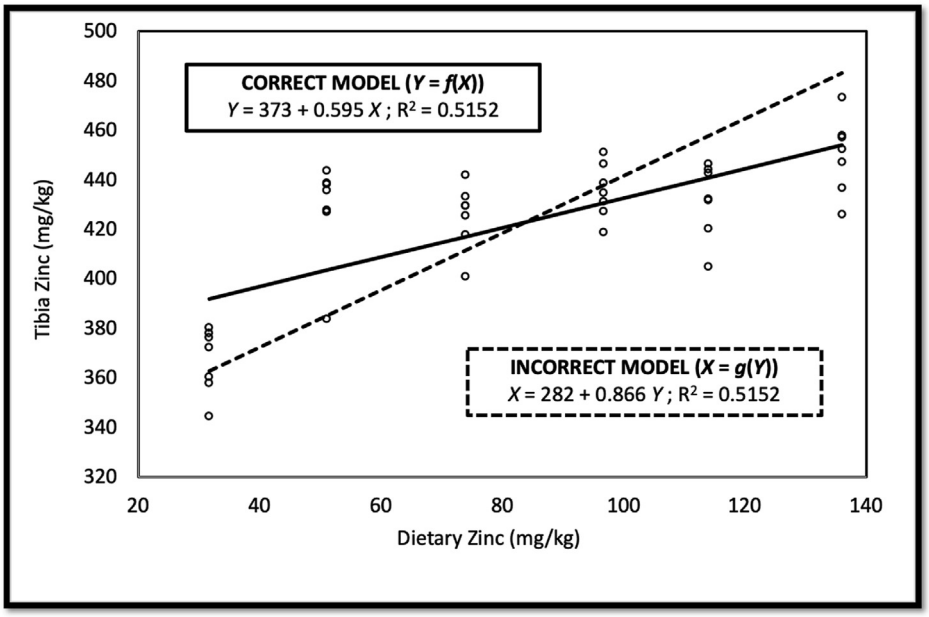

Fig. 4. Tibia zinc standard curve from feeding 6 levels of Zn with the Ordinary Least Squares fits of $Y = b_1 X + b_0$ in Eq. (1) and $X = b_1 Y + b_0$ in Eq. (15).

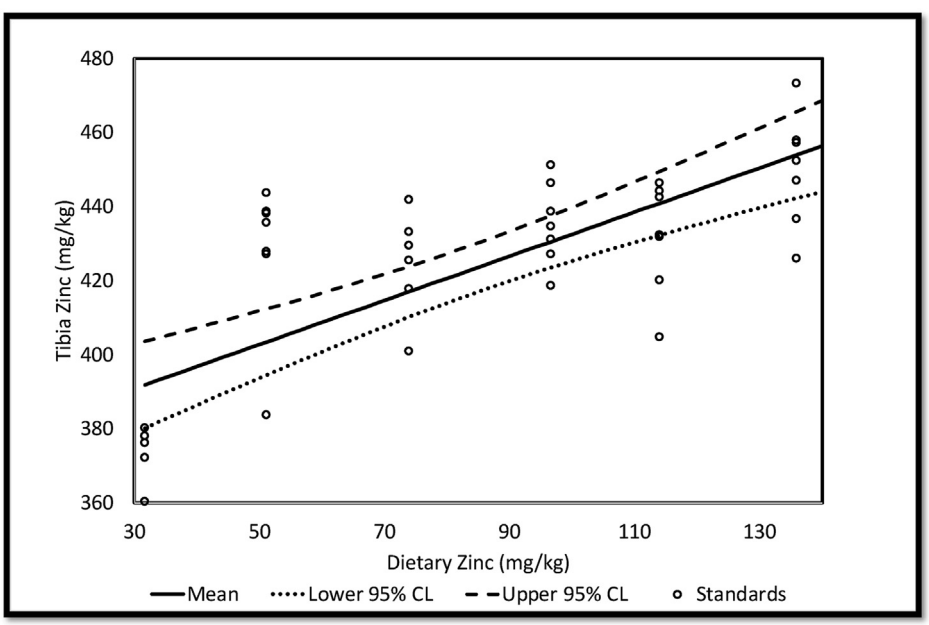

Fig. 5. Tibia zinc standard curve from feeding 6 levels of Zn with the Ordinary Least Squares fit of $Y = b_1 X + b_0$ and 95% confidence limits (CL).

ANOVA. Unfortunately, this approach yields measures of neither the mean bioavailability nor its CI. Although they know that they would like to determine a value for their unknown test samples, their primary tool is often a multiple range test and they apply it universally. For the example in Table 2, all the analyst can determine is that the test sample contains more than 32 mg Zn/kg.

Table 4
Descriptive statistics for 3 methods of interpreting bioavailability data from calibration curve experiments.[1]

Parameter	Symbol/formula	Best case scenario		Worst case scenario	
		Intuitive	Abductive	Intuitive	Abductive
Calibration (standard) curve	b_1	0.5951	0.5951	0.5951	0.5951
	b_0	373.03	373.03	373.03	373.03
	R^2	0.5152	0.5152	0.5152	0.5152
Test sample replicates	n_u	7	7	7	7
Mean test sample response	y_0	422.93	422.93	453.92	453.92
Predicted test sample Zn	x_0	83.859	83.859	135.931	135.931
SD of x_0	s_{x0}	41.664	14.516	41.664	16.568
CV	$(s_{x0}/x_0) \times 100$	49.683	17.310	30.651	12.189
SEM of x_0	$s_{x0}/(n_u - 1)^{-2}$	17.009	5.926	17.009	6.764
SD about regression	$s_{x/y}$	21.161	21.161	21.161	21.161
SD of calibration slope	s_{b1}	0.0913	0.0913	0.0913	0.0913
SD of calibration intercept	s_{b0}	8.3219	8.3219	8.3219	8.3219
LOD	$3s_{x/y}/b_1$	106.67	106.67	106.67	106.67
LOQ	$10s_{x/y}/b_1$	355.58	355.58	355.58	355.58

SD = standard deviation; CV = coefficient of variation; SEM = standard error of the mean; LOD = lower limit of detection; LOQ = lower limit of quantification.
[1] For the best case scenario, the average test sample responses were at the center of the standard curve. For the worst case scenario, the average test sample responses were at the upper extreme of the calibration curve.

Table 5
Power analysis showing how increasing the number of test sample replicates is expected to influence variations in the estimated level of zinc in feed from tibia zinc.[1]

Number of sample replicates	Centered results (best-case scenario)		Ends of range results (worst-case scenario)	
	Standard deviation	Standard error	Standard deviation	Standard error
1	36.16		38.25	
2	25.99	25.99	28.82	28.82
3	21.56	15.24	24.90	17.61
4	18.96	10.94	22.68	13.10
5	17.21	8.60	21.24	10.62
6	15.94	7.13	20.23	9.05
7	14.96	6.11	19.47	7.95
8	14.19	5.36	18.88	7.14
9	13.55	4.79	18.41	6.51
10	13.03	4.34	18.02	6.01
11	12.58	3.98	17.70	5.60
12	12.19	3.68	17.43	5.26

[1] Centered results fall in the middle of the standard curve zinc levels. End of range results fall at the extremes of the standard curve zinc levels.

Multiple range test results are not helpful in finding a relative bioavailability value, so the technique is really counter-intuitive (Table 2).

Practical researchers, especially in biological fields (as opposed to chemistry) may not be aware of theoretical problems with the Intuitive Method (Figs. 1 and 2 versus 3). Grabill's abductive method of estimating CIs in Eq. (23) is superior to the others, because it is theoretically sound, and considers both the error in the standard calibration curve and the different replicates of the unknown test sample (Fig. 3). The 2 important features of Graybill's equation in Eq. (23) are that: 1) the mean is predicted from the appropriate equation; and 2) that test samples near the average of the standard X values will have the smallest CIs, because that is where there is the most confidence in the calibration curve (Figs. 3 and 5), and conversely.

The method of Graybill has achieved general acceptance, as evidenced by the Royal Society of Chemists Technical Brief 22 (Thompson, 2006), and is commonly taught to analytical chemistry students (Harvey, 2019). The important question is "How often is this equation used in and outside of analytical chemistry?" There are excellent web resources explaining the use of Graybill's method on the internet (Prichard and Barwick, 2003). Multiple estimates of X_0 likely leave practical researchers to apply the simple statistics that they understand and regularly use the incorrect Eq. (11) and not the more appropriate abductive method correct Eq. (23).

Refereed journals do not generally require methods of calculating simple statistics like the CI to be reported. It cannot be known how often Eq. (23) is applied to calibration curve problems in biology and agriculture. Our experience is that neither Graybill's, nor any other, abductive method, is routinely taught to students outside of analytical chemistry classes. We suspect that the appropriate methods have rarely been applied to biological and agricultural calibration curve problems. The obvious, potentially large exception, is when practical researchers derive standard curves and unknown test samples from automated laboratory equipment that uses Graybill's method without the user even knowing it.

The CI's from the Sophistic, or reverse regression, method are not presented, because they are not appropriate for calibration curve bioassays. There continues to be interest in using the Sophistic method, finding X as a function of Y in Eq. (18) for calibration curve problems (Krutchkoff, 1967; Parker et al., 2010; Demidenko et al., 2013; Watters and LaMotte, 2020) despite compelling reasons why it is not appropriate (Eisenhart, 1939). Evaluating simulations not consistent with theory is irrational. From the very beginning of simulations seemingly supporting alternative approaches to the calibration curve dilemma, Krutchkoff (1967) should have questioned his simulations or explained what was wrong with current theory (Eisenhart, 1939). Instead, he advocated using a method contrary to the idea that Y should be dependent on X. If the purpose of an experiment is to estimate some property of

test samples by comparing them to responses of known quantities, the X's, then Eqs. (1) and (23) are the appropriate ones to use to determine means and CI's of the quantity Y. That is, the solid line in Fig. 1 is the correct one to use to estimate the amount of X in the unknown test sample Y.

When estimates of the test sample compositions' variation are desired, the reverse regression or Sophistic Method is still often considered (Halperin, 1970; Kannan et al., 2007; Parker et al., 2010; Demidenko et al., 2013; Watters and LaMotte, 2020). Parker et al. (2010) concluded from a series of simulations that both Intuitive and Sophistic methods have bias. When the R^2 of the $Y = f(X)$ and $X = f(Y)$ regressions are very high, there is little difference in predictions, and mean estimates are very similar, or practically indistinguishable, but still different except for one point, the intersection of the red and blue lines in Fig. 1. For many applications in analytical chemistry, the calibration curve has little associated error making discrimination between simulation methods very difficult. However, when R^2 values are lower, as in Figs. 4 and 5 and Tables 3 and 4, the predicted means will be clearly different. The predicted CI is for any method not accurately predicting the mean, because incorrect methods/models were used, are irrelevant.

The origin of misunderstanding of the relationship of $Y = f(X)$ and $X = g(Y)$ may be introductory texts that inadequately explain the fitting of linear equations. The excellent text of Kutner et al. (2005) explains and "proves" the Gauss-Markov theorem: "*Under the conditions of regression model (1.1), the least squares estimators b_0 and b_1 in (1.10) are unbiased, and have minimum variance among all unbiased linear estimators*". This relationship can be proven, but only when the $R^2 = 1$, a condition never achieved when dealing with biological systems. This misconception naturally leads researchers to assume that $Y = f(X)$ and $X = g(Y)$ will give identical results, which they actually do not (Fig. 1).

Using the Intuitive Method has a drawback: the curve itself is correct and the average X_0 is predicted correctly. Near the average of the range of standards, where the lines cross in Fig. 1, the CI will be over estimated, but near the extremes of the standard levels the CI will be underestimated (Tables 3 and 4). Nonetheless, the results do not appear unreasonable to most researchers, and the method is rarely questioned beyond analytical chemistry. The variation in the results (X_0's) is highly proportional to the variation in the responses (Y_0's) for most assays with relatively high R^2 values, so the results appear reasonable, even if theoretically unsustainable.

Abductive methods are the only ones to approximate different CI's for the calibration curve at different levels of X, the true situation. They should therefore be the most appropriate to use (Table 3). The regression parameter estimates are exactly the same for the Intuitive and Abductive Methods, as are the average predicted test sample Zn values. The observed variation (SD's) from the Abductive Method are only about 1/3 of those from the Intuitive Method for the worst case scenario. This demonstrates that researchers using the Intuitive Methods are overestimating the variation in their results. The degree of overestimation depends on the error in the standard curve. When the R^2 is 1.000, the difference in the predictions is zero.

4.3. Power analysis

Power analysis is usually associated with hypothesis testing (Berndtson, 1991; Demétrio et al., 2013; Shim and Pesti, 2014). It is an estimate of the number of replicate observations of each treatment that are necessary to find a specified mean difference at some probability β (usually 0.80), if there truly is a difference between treatments. Experimental power for the calibration curve problem is different, because hypothesis testing is not the objective. The objective is only to find the mean and CI of some property of an unknown test sample. The importance of having a small CI depends on the value of the property being quantitated.

Experimental power for the calibration curve problems is complicated, because the CI for X_0 will have different ranges depending on the value estimated for X_0. If the value of an unknown test sample X_0 is at the average of the standard values making up calibration curve, it will have the smallest CI (best case scenario). If the value of X_0 is nearer the extremes of the calibration curve, the CI will be wider (worst case scenario).

The starting point to estimate the CIs for different numbers of replicates for a future experiment is Eq. (23) with the estimated σ^2, β_0, β_1, etc. from a previous experiment. Then the values of n and m can be varied to produce the dashed line in Fig. 6. For the Zn

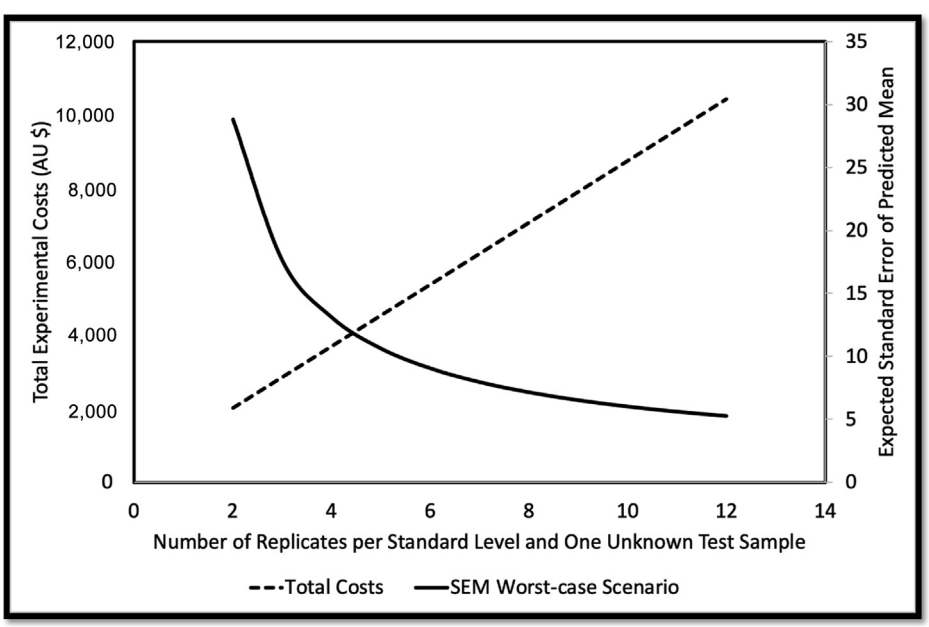

Fig. 6. Predicted experimental power from expected SEM of one unknown test sample (solid line) conducted with different numbers of experimental observations versus the total costs of the experiment (dashed line).

bioavailability example (Table 4), the SD of worst case scenario is only about 6% greater than the best case scenario with one test sample replicate (38.25 versus 36.16), but almost 43% greater with 12 test sample replicates (5.25 versus 3.68; Table 4). Using the worst case scenario for experimental power considerations for calibration curves should result in smaller than expected CI's, because all the unknown test samples should have values closer to average of the standards (the best case scenario).

Research designers must decide on the best use of resources in determining the number of replicates (Fig. 6). There is no number of replicates that give the best or optimum level of confidence in the results. A value judgment must be made to balance expenses with expected outcomes. Increasing replication from 2 to 4 pens per standard level and test sample results in a much greater decrease in the expected SEM than increasing from 8 to 10 replicates per standard level and sample. For calibration curve experiments when relatively low R^2 values are expected, using the worst case scenario to determine replication needs should help prevent disappointment (and repeatability problems) in research results. Power analyses add some level of objectivity to experimental planning since costs versus expected outcomes can be quantified. The researcher must still balance expenses with needed outcomes (e.g. the ability to provide enough of the quantity being measured). Cost and benefit analyses can be greatly improved by choosing the correct method to predict outcomes.

5. Conclusions

Using Graybill's Abductive Method for calculating confidence intervals in calibration curve experiments is prudent, based on: 1) its adherence to the underlying assumptions of the experimental situation and hence to proper statistical theory; and 2) its ability to predict accurately sample error along the calibration curve. There are clear reasons not to use the Counter-intuitive, Intuitive, or Sophistic methods, not least because they are mathematically and therefore scientifically incorrect. Graybill's method (Eq. (23)) is scientifically sound and has wide acceptance among analytical chemists. It should be even more helpful to biologists who often deal with higher levels of variation in their experiments than do chemists. With modern computers practically ubiquitous, there is no extra effort (programming) involved in using Graybill's Abductive Method. Practical researchers need not understand the finer points of mathematical theory, any more than mathematical theorists need appreciate the finer points of biologists' practical problems. It is important for the 2 groups to learn some fundamental aspects of each other's discipline to have meaningful communication.

Author contributions

Gene M. Pesti: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing – Original Draft; **Lynne Billard**: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing – Original Draft; **Shu-Biao Wu**: Resources, Supervision, Writing – Review & Editing; **Robert A. Swick**: Resources, Supervision, Writing – Review & Editing; **Thi Thanh Hoai Nguyen**: Investigation, Writing – Original Draft Writing – Review & Editing; **Natalie Morgan**: Resources, Writing – Review & Editing.

Declaration of competing interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, and there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the content of this paper.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Analytical Methods Committee. Uncertainties in concentrations estimated from calibration experiments. In: Thompson M, editor. AMC Technical Brief No 22; 2006. http://www.rsc.org/membership/networking/interestgroups/analytical/amc/technicalbriefs.asp. [Accessed 6 February 2021].
Aviagen. ROSS Nutrition Specifications. Scotland: Aviagen; 2019. https://en.aviagen.com/assets/Tech_Center/Ross_Broiler/RossBroilerNutritionSpecs2019-EN.pdf. [Accessed 29 May 2021].
Berndtson WE. A simple and reliable method for selecting or assessing the number of replicates for animal experiments. J Anim Sci 1991:6967–76.
Berkson J. Estimation of a linear function for a calibration line; Consideration of a recent proposal. Technometrics 1969;11:649–60.
Casella G, Berger RL. Statistical inference. 2nd ed. Pacific Grove CA: Duxbury Press; 2002.
Demétrio CGB, Menten JFM, Leandro RA, Brien C. Experimental power considerations – justifying replication for animal care and use committees. Poult Sci 2013;92:2490–7. https://doi.org/10.3382/ps.2012-02731.
Demidenko E, Williams B, Flood AB, Swartz HM. Standard error of inverse prediction for dose–response relationship: approximate and exact statistical inference. Stat Med 2013;32:2048–61. https://doi.org/10.1002/sim.5668.
Draper N, Smith H. Applied regression analysis. 2nd ed. New York: Wiley; 1981.
Eisenhart C. The interpretation of certain regression models and their use in biological and industrial research. Ann Math Stat 1939;10:162–86. https://doi.org/10.1214/aoms/1177732214.
Graybill FA. An introduction to linear statistical models, vol. 1. New York: McGraw-Hill; 1961.
Graybill FA. Theory and application of the linear model. North Scituate, MA: Duxbury Press; 1976.
Halperin M. On inverse estimation in linear regression. Technometrics 1970;12:727–36.
Harvey D. Section 5.4 Linear regression and calibration curves. In: Analytical chemistry 2.1. Libretext; 2019. https://chem.libretexts.org/bookshelves/analytical_chemistry/book%3a_analytical_chemistry_2.1_(harvey)/05%3a_standardizing_analytical_methods/5.04%3a_linear_regression_and_calibration_curves. [Accessed 6 February 2021].
Kannan N, Keating JP, Mason RL. A comparison of classical and inverse estimators in the calibration problem. Commun Stat-Theor M 2007;36:83–95.
Kleinbaum DG, Kupper LL, Muller KE. Applied regression analysis and other multivariate methods. 2nd ed. Boston: Pws-Ken; 1988.
Krutchkoff RG. Classical and inverse regression methods of calibration. Technometrics 1967;9:425–39. https://doi.org/10.2307/1266511.
Kutner MH, Nachtsheim CJ, Neter J, Li W. Applied linear statistical models. 5th ed. Boston: McGraw-Hill Irwin; 2005.
Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. 3rd ed. Hoboken: Wiley-Interscience; 2001.
Miller JN. Basic statistical methods for analytical chemistry, Part 2. Calibration and regression methods, A review. Analyst 1991;116:4–14.
Osborne C. Statistical calibration: a review. Int Stat Rev 1991;59:309–36. https://doi.org/10.2307/1403690.
Parker PA, Geoffrey G, Vining G, Wilson SR, Szarka III JL, Johnson NG. The prediction properties of classical and inverse regression for the simple linear calibration problem. J Qual Technol 2010;42:332–47. https://doi.org/10.1080/00224065.2010.11917831.
Prichard L, Barwick VR. Preparation of calibration curves. A guide to best practice. ResearchGate 2003. https://doi.org/10.13140/rg.2.2.36338.76488.
SAS Institute. SAS Institute SAS user's guide: statistics. 4th ed. Cary, NC (2008): SAS Institute. Inc.; 2008. Version 9.
Shim MY, Pesti GM. The use of a pen-size optimization workbook for experiment research design using the Visual Basic for Applications in Excel for poultry. J Appl Poult Res 2014;23:315–22.
Tellinghuisen J. Least squares in calibration: dealing with uncertainty in x. Analyst 2010;135:1961–9.
Watters CG, LaMotte LR. A comparison of four methods of inverse prediction. Commun Stat-Simul C 2020;49:709–16. https://doi.org/10.1080/03610918.2019.1601217.
Williams EJ. A note on regression methods in calibration. Technometrics 1969;11:189–92. https://doi.org/10.2307/1266774.