A Bayesian precision profile for measuring the quality of immunoassay experiments

BY A. JAMES O’MALLEY *

Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115-5899, USA

A precision profile, the relationship between the concentration of a substance and its measured precision, is a convenient way of conveying the ability of an immunoassay to accurately measure the concentration of a substance in blood serum. A precision profile is characterized by the definition of precision. Historically, precision has been evaluated as the standard error of an estimator of the concentration in a sample conditional on the true concentration. In this paper, Bayesian predictive inference is used to develop a new measure of precision based on the accuracy with which an assay could infer the concentration in a hypothetical new sample. This leads to a natural procedure for evaluating a precision profile that avoids using approximations such as those inherent in traditional methods.

Keywords: Bayesian analysis; calibration; immunoassay; nonlinear regression; precision profile; predictive inference

1. Introduction

In an assay, independent responses are observed for a set of standards (samples containing known concentrations of test substance) together with specimens containing unknown concentrations of test substance. The purpose of the assay is to make inferences about the unknown concentrations. The standards are the primary source of information about the dose–response curve, the relationship between the responses and the concentrations, from which the unknown concentrations are inferred. Standard and unknown samples are processed identically. Usually, several responses are obtained for each standard, and there are many more unknown samples than standard samples. Historically, replicated responses were also obtained for the unknown samples, although recently there has been a trend towards a singleton measurement (Sadler & Smith 1990a).

A precision profile for an immunoassay is the relationship between the concentration of a substance in blood serum and its measured precision, where precision quantifies the uncertainty or noise in a measurement (incorporating

*omalley@hcp.med.harvard.edu

One contribution of 13 to a Theme Issue ‘Mathematical and statistical methods for diagnoses and therapies’.
baseline errors, errors in dispensing samples and reagents and errors counting antibodies at the completion of the reaction). The precision evaluates the performance of an assay across the full range of concentrations of interest, whereas other measures of assay performance, such as the minimal detectable concentration (MDC; Davidian et al. 1988; Brown et al. 1996; O’Malley & Deely 2003), tend to focus on a specific concentration (e.g. 0). As such, the precision profile is extremely useful to practitioners trying to determine which assay is best for a given application.

The use of precision profiles in assays dates back to Ekins et al. (1972) and Ekins (1978, 1983). On occasion, they have also been referred to as imprecision profiles (e.g. Sadler et al. 1988; Sadler & Smith 1990a,b). To cater for five different analytical requirements of assays, Ekins (1983) defines five distinct precision profiles. Analytical requirements may encompass one or multiple samples of the substance being tested from the same individual (intra-sample versus inter-sample), multiple repetitions of the assay (intra-assay versus inter-assay) and multiple laboratories (inter-laboratory). The intra-sample, intra-assay precision profile that incorporates the error in the randomness of the response counts and the error in the fitted dose–response curve is of interest here. A precision profile along these lines appears to be generally desired by practitioners (O’Connell et al. 1993). However, as elaborated on in §6, the introduction of modern machine assays that use only singleton measurements has led to precision profiles that do not account for the error in the fitted dose–response curve. By contrast, the methods developed in this paper encapsulated both sources of errors and can be adapted to correspond to a variety of analytical requirements.

Precision has been expressed in terms of the estimated variance, standard deviation or coefficient of variation of the sampling distribution of an estimator of the concentration in a sample by Ekins and co-workers (O’Connell et al. 1993; Sadler et al. 1988; Sadler & Smith 1990a). A concern with the use of measures based solely on variability is that the concept of ‘bias’ is not considered. For such a measure to be valid, bias must be negligible—if bias is not part of the evaluation, then constants become perfect estimators.

In this paper, a new measure of precision (and hence a new precision profile) is derived from a Bayesian model for the analysis of an immunoassay experiment. Although the use of Bayesian methods in the analysis of assays dates back at least to Ramsey (1972), Bayesian precision profiles have not been developed. However, Brown et al. (1996) and O’Malley & Deely (2003) incorporate Bayesian ideas in defining the MDC, and empirical Bayes methods have been used to characterize intra-assay variation (Giltinan & Davidian 1994). Similar to the Bayesian measures of MDC in O’Malley & Deely (2003), a Bayesian measure of precision is constructed using posterior predictive distributions.

The rest of this paper is organized as follows. Model details and notation are presented in §2. Existing measures of precision are reviewed in §3, while in §4 a Bayesian measure of precision is developed. In §5, the traditional and the new Bayesian precision profiles are applied to data from a serum thyroxine radio-immunoassay. The results are compared between the estimators and the implications of using different precision profiles in practice are described. Section 6 presents the concluding remarks.
2. Assay model and notation

For simplicity, no distinction will be made between random variables and their realizations. Lower-case letters in bold-italic denote vectors, while upper-case letters in bold-italic denote matrices and concatenated vectors.

Let 
\[ y_i = (y_{i1}, \ldots, y_{ir_i}) \]
 denote the vector of \( r_i \ (i=1, \ldots, n) \) exchangeable responses at known concentrations \( x_i \ (i=1, \ldots, n_s) \) or at unknown concentrations \( \eta_i \ (i=n_s+1, \ldots, n_s+n_u) \), where \( n_s \) is the number of known specimens (standards); \( n_u \) is the number of unknown specimens; and \( n = n_s + n_u \) is the total number of specimens. The sets of indices for the standard and the unknown samples are denoted by \( S \) and \( U \), respectively. Let 
\[ x = (x_1, \ldots, x_{n_s}) \]
 and 
\[ \eta = (\eta_1, \ldots, \eta_{n_u}) \]
 denote the respective vectors of known and unknown concentrations, respectively. The sample mean of \( y_i \) is denoted by \( \bar{y}_i \).

The concatenated vector \( Y = (y_1, \ldots, y_n) \) denotes all the observed responses.

It will be convenient to discuss a ‘general’ or ‘future’ vector of \( m_f \) exchangeable responses. Such a response vector at a concentration \( x_f \) will be denoted by \( y_f \) and the associated mean (of the elements of \( y_f \)) by \( \bar{y}_f \).

Let the probability distribution of \( y_{ij} \) at concentration \( \eta_i \) be 
\[ p(y_{ij}|\eta_i, \beta, \theta) \]
 with associated mean \( m(\eta_i, \beta) \) and variance \( v(m_i, \theta) \), where \( \beta \) and \( \theta \) are unknown vectors of parameters. Traditionally, \( p(y_{ij}|\eta_i, \beta, \theta) \) is a normal density (Brown et al. 1996).

The mean function \( m(\eta_i, \beta) \) is assumed to be monotonic in \( \eta_i \) but may depend nonlinearly on \( \eta_i \) and \( \beta \). Two commonly used mean functions are the sigmoid function (Rodbard & Hutt 1974; Finney 1976)
\[ m(\eta_i, \beta) = \beta_4 + \frac{\beta_1 - \beta_4}{1 + (\frac{\eta_i}{\beta_3})^{\beta_2}} \]
and its affine-transformed version (O’Connell et al. 1993)
\[ m(\eta_i, \beta) = \frac{mt}{100} \left( \beta_4 + \frac{\beta_1 - \beta_4}{1 + (\frac{\eta_i}{\beta_3})^{\beta_2}} \right) + \text{nsb}, \]  \hspace{1cm} (2.1)
where ‘mt’ and ‘nsb’ are assay-specific constants. Equation (2.1) may be increasing or decreasing depending on the way in which the assay is performed (Davies 1994); in this paper, the decreasing case is assumed, implying \( \beta \geq 0 \).

The variance function is typically an increasing function of the mean response such as the power function
\[ v(m, \theta) = \theta_1 m^{\theta_2}, \quad \theta \geq 0. \]  \hspace{1cm} (2.2)
In most immunoassays, the variance of the response counts exceeds the variance of a Poisson-distributed random variable. This possibility is accommodated in (2.2) by allowing \( \theta_2 > 1 \).

The prior distribution of a parameter \( \lambda \) will be denoted by \( p(\lambda) \) and the resulting posterior distribution, given data \( D \), is denoted by \( p(\lambda|D) \). From Bayes’ theorem,
\[ p(\lambda|D) = \frac{p(D|\lambda)p(\lambda)}{p(D)}, \]  \hspace{1cm} (2.3)
where \( p(D) = \int \lambda p(D|\lambda) p(\lambda) \, d\lambda \) is the marginal distribution of \( D \). Non-informative prior distributions will be assumed in this paper (see §5). An estimate of \( \lambda \) is denoted by \( \hat{\lambda} \).

(a) Estimation of unknown concentrations

Owing to the limited range of \( m(\eta_i, \beta) \), the estimation of \( \eta_i \) for \( i \in U \) must be examined carefully. Classical estimators of \( \eta_i \) are typically determined by back-fitting the mean response \( \bar{y}_i \) to the corresponding concentration \( \eta_i \). The mapping in the case of (2.1) is given by

\[
m^{-1}(\bar{y}_i, \hat{\beta}) = \hat{\beta}_3 \left( \frac{\hat{\beta}_1 - \bar{y}_{i, \text{affine}}}{\bar{y}_{i, \text{affine}} - \hat{\beta}_4} \right)^{1/\hat{\beta}_2},
\]

where \( \bar{y}_{i, \text{affine}} = (100/\text{int})(\bar{y}_i - \text{nsb}) \) and, by assumption, \( \beta > 0 \). In order for the mapping to be well defined, the inverse function \( \eta(\bar{y}_i, \beta) = m^{-1}(\bar{y}_i, \beta) \) needs to be defined over the full range of the data, which in the case of (2.1) requires

\[
\hat{\beta}_4 < \bar{y}_{i, \text{affine}} < \hat{\beta}_1 \quad \text{for all} \quad i \in U.
\]

That is, the affine-transformed mean response \( \bar{y}_{i, \text{affine}} \) must lie between the expected counts at zero (\( \hat{\beta}_1 \)) and infinite (\( \hat{\beta}_4 \)) concentrations. In practice, the standards are typically designed so that their concentrations span the concentrations observed in most (or almost all) clinical specimens. However, very rare cases of a mean response \( \bar{y}_{i, \text{affine}} > \hat{\beta}_1 \) occur in practice, in which case a zero (or a smaller value than \( \hat{L} \), where \( \hat{L} \) is an estimate of the smallest concentration that can be distinguished from zero) concentration is assigned since interpolation is precluded. Similarly, if a mean response is less than the fitted response of the highest concentration standard, then a larger than \( \hat{M} \) concentration is reported, where \( \hat{M} \) is the largest concentration that could reasonably be expected in practice. When using Bayesian inference, a prior distribution that specifies \( \eta_i \geq 0 \) with a probability of 1 will ensure that only non-negative values of \( \eta_i \) can be inferred, eliminating the need for ad hoc fixes and potentially yielding more efficient inferences.

Traditional estimators of \( \eta_i \) typically depend only on the data from sample \( i \) through \( \bar{y}_i \) exclusively and on the data from other samples through \( \hat{\beta} \) exclusively; \( \eta(\bar{y}_i, \hat{\beta}) \) denotes such an estimator. However, because Bayesian inferences about \( \eta_i \) also incorporate information from the dependence of the variance function \( v(m, \theta) \) on the mean \( m \), the more general notation \( \eta(y_i, Y, X) \) is used for Bayesian estimators.

3. Current method of calculating precision

Historically, the precision of an assay at concentration \( x_f \) (the concentration in some hypothetical future sample) is based on the variance or coefficient of variation of the sampling distribution of \( \eta(\bar{y}_i, \hat{\beta}) \), where \( y_i \) is a response measurement associated with \( x_i \). This calculation cannot, in general, be performed in a closed form. In this section, we describe two methods of approximating the sampling variance of \( \eta(\bar{y}_i, \hat{\beta}) \): the analytical method and the empirical method.
(a) Analytical method

When \( \hat{\eta}_t = \eta(\bar{y}_t, \hat{\beta}) \) is a closed-form function of \( \bar{y}_t \) and \( \hat{\beta} \), an approximation based on a Taylor series expansion (TSE) may be used to estimate the precision of an assay. The inverse estimator in (2.3), which corresponds to the pseudo-maximum-likelihood estimator of \( \eta \) for fixed \( \theta \), has the required form. However, when the variance function depends on \( \eta_t \) directly or through the mean as in (2.2), the likelihood equations cannot, in general, be solved for \( \eta_t \), implying that Taylor series methods may not be applicable to the full maximum-likelihood estimator.

Ekins (1983) uses a first-order TSE of \( \eta(\bar{y}_t, \hat{\beta}) \) about \( E_{\bar{y}_t} = m(x_t, \beta) \) to approximate the mean and variance of \( \hat{\eta}_t \). This ignores the variability in \( \beta \) (i.e. treats \( \beta \) as known) and so does not incorporate all of the uncertainty in the estimated concentrations. In §5, it is shown that this method yields smaller values of precision than those of other methods. The approach of O’Connell et al. (1993) is based on the TSE of \( \eta(\bar{y}_t, \hat{\beta}) \) about \( E_{\bar{y}_t} \) and \( E_\beta = E[\beta|x] \), which is given by

\[
\eta(E_{\bar{y}_t}, E_\beta) + \eta_{\bar{y}_t}(E_{\bar{y}_t}, E_\beta)(\bar{y}_t - E_{\bar{y}_t}) + \eta_\beta(E_{\bar{y}_t}, E_\beta)'(\hat{\beta} - E_\beta),
\]

where \( \eta_z \) denotes the partial derivative (or gradient function in the case of a vector argument) of \( \eta(z, w) \) with respect to \( z \). The first-order approximations of the mean and variance of \( \eta(\bar{y}_t, \beta) \) are thus

\[
E[\hat{\eta}_t|x_t, x] \approx \eta(E_{\bar{y}_t}, E_\beta) \approx x_t
\]

and

\[
V(\hat{\eta}_t|x_t, x) \approx \eta_{\bar{y}_t}(E_{\bar{y}_t}, E_\beta)^2 v(E_{\bar{y}_t}, \hat{\theta}) + \eta_\beta(E_{\bar{y}_t}, E_\beta)' \text{cov}(\hat{\beta}) \eta_\beta(E_{\bar{y}_t}, E_\beta),
\]

where \( \text{cov}(\hat{\beta}) \) is the covariance matrix of \( \hat{\beta} \). To enable (3.3) to be evaluated, parameter estimates are substituted for the unknown parameters in \( E_{\bar{y}_t}, E_\beta, v(E_{\bar{y}_t}, \hat{\theta}) \) and \( \text{cov}(\hat{\beta}) \).

The estimated precision at concentration \( x_t \) is \( S(\hat{\eta}_t|x_t, x) = V(\hat{\eta}_t|x_t, x)^{1/2} \) and \( S(\hat{\eta}_t|x_t, x)/x_t \) when precision is defined in terms of the standard deviation and the coefficient of variation, respectively. Equations (3.2) and (3.3) do not depend on the distribution of the response counts.

Approximations based on second-order (or higher) TSEs are computed analogously. Although these have the potential to be more accurate, the resulting expressions tend to be unwieldy, even if the assumption that the response counts are normally distributed is invoked, and so are not displayed here. However, in §5, precision profiles based on first- and second-order TSEs are plotted.

(b) Empirical method

An alternative approach to deriving the precision profile is to fit a model to a dataset of estimated concentrations. Let \( \hat{\eta}_{ij} \) denote the concentration back-fitted from \( \bar{y}_{ij} \), \( i = 1, \ldots, n \) and \( j = 1, \ldots, r_i \). A variance function of \( \hat{\eta}_{ij} \) in terms of the true concentration \( \eta_i \) is then estimated using a parametric model, such as

\[
\hat{\eta}_{ij} \sim N(\eta_i, (\alpha_1 + \alpha_2 \eta_i)^{\alpha_3}),
\]

where \( \alpha \) is a vector of unknown parameters and \( N(\mu, \sigma^2) \) denotes the density of a normal distribution with mean \( \mu \) and variance \( \sigma^2 \). We recommend the use of restricted maximum likelihood to fit the model. The efficiency of \( \hat{\alpha} \) can be improved by using the fact that \( \eta_i = x_i \), for \( i \in S \) (O’Malley et al. in press).
Expressions for $E[\hat{\eta}_i|\eta_i, x]$ and $V(\hat{\eta}_i|\eta_i, x)$ are derived from the fitted model. For example, if $r_i=2$, the fitted variance function associated with (3.4) is given by

$$V(\hat{\eta}_i|\eta_i = x_i, x) = 2^{-1}(\hat{\alpha}_1 + \hat{\alpha}_2 x_i)\hat{\beta}_i.$$  

The empirical method implicitly assumes that the fitted values of the model parameters are the true values or, in other words, that the variability of the response counts is the only source of uncertainty in the unknown concentrations that is of interest. This viewpoint is equivalent to that of the analytical method based on a TSE in $\tilde{y}_i$ alone. However, the empirical method involves fewer assumptions because it does not rely on the distribution, mean function or variance function of the response counts.

More extensive descriptions of the analytical method are given in Sadler et al. (1988) and Sadler & Smith (1990a,b). A key point is that the empirical method was not intended for use on data from a single batch of samples since, in order to compute the variability of the estimated concentrations, it relies on replicated samples (thus making it impossible to apply to modern machine assays). Rather, the empirical method has historically been used in situations where the assay includes replication of some sort. For example, if the same assay is calibrated at the start of the day (thus, separate batches of samples are obtained each day), a precision profile can be determined from the standards and any other samples that are repeatedly analysed. However, for comparison purposes, the empirical method is included in the analysis.

4. Bayesian measure of precision

The key terms in a Bayesian analysis of an assay are described in the following. For the model described in §2, the likelihood function is given by

$$p(Y|x, \eta, \beta, \theta) = \prod_{i=1}^{n} \prod_{j=1}^{r_i} p(y_{ij}|x_i, \beta, \theta)^{I(i\in S)} p(y_{ij}|\eta_i, \beta, \theta)^{I(i\in U)},$$

where $I(\text{event})=1$ if ‘event’ is true and 0 otherwise. The joint prior distribution of the unknown parameters is denoted by $p(\beta, \theta, \eta)$. Using Bayes’ theorem, the posterior distribution of the model parameters ($\beta, \theta$) is computed as

$$p(\beta, \theta|Y, x) = \frac{\int_{\eta} p(Y|x, \eta, \beta, \theta) p(\beta, \theta, \eta) \, d\eta}{\int_{\eta} \int_{\beta} \int_{\theta} p(Y|x, \eta, \beta, \theta) p(\beta, \theta, \eta) \, d\eta \, d\beta \, d\theta} = \int_{\eta} p(\beta, \theta, \eta|Y, x) \, d\eta.$$

The posterior distribution $p(\beta, \theta|Y, x)$ contains all the information in the assay about ($\beta, \theta$).

The posterior distribution conditional on ($y_i, \beta, \theta$) of the unknown concentration $\eta_i$ in a new sample is given by

$$p(\eta_i|y_i, \beta, \theta) = \frac{p(y_i|\eta_i, \beta, \theta) p(\eta_i|\beta, \theta)}{\int_{\eta} p(y_i|\eta_i, \beta, \theta) p(\eta_i|\beta, \theta) \, d\eta_i},$$

(4.1)

where the prior $p(\eta_i|\beta, \theta) = p(\eta_i)$ if $\eta_i$ is a priori independent of ($\beta, \theta$). To minimize the influence of the prior on the analysis of the serum thyroxine radioimmunoassay in §5, we set $p(\eta_i) = U(0, M)$, where $U(0, M)$ denotes the uniform distribution on the interval $(0, M)$; as defined earlier, $M$ is the largest concentration that could realistically be observed in practice. There is no point in

---

Phil. Trans. R. Soc. A (2008)
using a value larger than \( M \) as an upper limit because such parameter values would never be supported by the data. In practice, the distribution of the concentration \( \eta \) of a substance in a test population is often well known. However, we chose not to incorporate such a priori information in this analysis as we desired to compare the procedures in terms of experimental information alone.

If \( y_i \) was observed, it would contain information about \((\beta, \theta)\), and hence the posterior of \( \eta \) would be computed as

\[
p(\eta_i | y_i, Y, x) = \int_\theta \int_\beta p(\eta_i | y_i, \beta, \theta) p(\beta, \theta | y_i, Y, x) \, d\beta \, d\theta.
\]  

(4.2)

(a) Concentration distribution of assay noise precision

We desire to describe the precision (or level of uncertainty) in an assay’s measurement of an unknown concentration \( \eta \) when the true concentration is \( x \). From the perspective of evaluating the precision of an assay conditional on the observed data \((Y, x)\), \( y_i \) has not been observed and so does not represent ‘data’. Therefore, \( y_i \) should not be used to update the posterior distribution of \((\beta, \theta)\), and hence (4.2) is not appropriate for evaluating precision (see O’Malley & Deely 2003 for further comments). In the following, we overcome this problem.

Let \((y_i)\) and \((x_i)\) indicate that \( y_i \) and \( x_i \) are not used to update \( p(\beta, \theta | Y, x) \) in a given computation. To extend (4.1) to account for the variability in \( y_i \) at concentration \( x_i \), compute

\[
p(\eta_i | x_i, \beta, \theta) = E_{y_i}[p(\eta_i | y_i, \beta, \theta)] = \int_{y_i} p(\eta_i | y_i, \beta, \theta) p(y_i | x_i, \beta, \theta) \, dy_i,
\]

the posterior distribution of \( \eta_i \) conditional on \((x_i, \beta, \theta)\). Finally, to account for the uncertainty in \((\beta, \theta)\), compute

\[
p(\eta_i | x_i, Y, x) = E_{\beta, \theta}[p(\eta_i | x_i, \beta, \theta) | Y, x] = \int_\theta \int_\beta p(\eta_i | x_i, \beta, \theta) p(\beta, \theta | Y, x) \, d\beta \, d\theta.
\]  

(4.3)

(b) CDAN precision and CDAN precision profile

The final step in the evaluation of a Bayesian measure of precision is to reduce \( \text{CDAN}(x_i) = p(\eta_i | x_i, Y, x) \) to a summary statistic to be plotted against \( x_i \) to yield the precision profile. The expected root mean squared error (RMSE), or relative RMSE, about \( x_i \) of a random variable drawn from \( \text{CDAN}(x_i) \) is a natural choice. The mean squared error (MSE) of \( \text{CDAN}(x_i) \) is the posterior predictive value of \( E[(\eta_i - x_i)^2 | y_i, Y, x] \). In this regard, the MSE of \( \text{CDAN}(x_i) \) is directly related to the quantity that would be reported if, having observed \( y_i \), \( \eta_i \) was estimated by minimizing the expected squared-error loss. An advantage of RMSE relative to the standard deviation or the

\[\text{Phil. Trans. R. Soc. A} (2008)\]
coefficient of variation is that it accounts for the bias or accuracy of the assay in addition to its variability. For the general loss function, denoted by $d(\eta_t, x_t)$, the CDAN precision is evaluated as $E[d(\eta_t, x_t) | x_t, Y, x]$. For example, if we desire to compute $pr(\eta_t > \eta_0 | x_t, Y, x)$, we would let $d(\eta_t, x_t) = I(\eta_t > \eta_0)$; in the special case where $\eta_0 = 0$, this yields a measure of the MDC similar to the measures derived in O’Malley & Deely (2003).

Measures of the CDAN precision may be evaluated using Monte Carlo averages involving a large sample of $\eta_t$ drawn from CDAN$(x_t)$. The values of $\eta_t \sim$ CDAN$(x_t)$ are obtained using a Markov chain Monte Carlo (MCMC) algorithm, involving Gibbs sampler (Geman & Geman 1984; Gelfand & Smith 1990; Casella & George 1992) and Metropolis–Hastings (Metropolis et al. 1953; Hastings 1970) steps, to draw values of $(\beta, \theta)$ from the joint posterior distribution. Next, conditional on each draw of $(\beta, \theta)$, we generate a hypothetical new response $y_t$ at concentration $x_t$. Finally, we generate $\eta_t$ from the posterior distribution of $\eta_t$ given $(y_t, \beta, \theta)$. After many draws of $\eta_t$ (e.g. 10 000 or more), the precision of the assay at $x_t$ can be accurately approximated. If interpolation is used to directly link the simulated values of precision, the resulting precision profile will depict the level of simulation error in the computation. To obtain a smooth precision profile, a parametric family of functions (e.g. such as the variance function in (3.4)) or a non-parametric regression function may be fitted to the simulated values of precision.

5. Analysis of radioimmunoassay data

In this section, the Bayesian CDAN precision profile is evaluated and compared with traditional precision profiles using data from a serum thyroxine radioimmunoassay performed in the Department of Nuclear Medicine at Christchurch Public Hospital (NMCPH), New Zealand. Thyroxine is a hormone secreted by the thyroid gland and its measurement is a primary tool in the investigation of thyroid function. The radioimmunoassay uses seven standard samples with thyroxine concentrations of 0, 35, 70, 140, 210, 280 and 350 nmol l$^{-1}$, and it has a reference range (encompasses 95% of subjects free of thyroid disease) of 55–140 nmol l$^{-1}$. In these assays, the zero standard was replicated four times, while all other samples were measured in duplicate. The total number of unknown samples was 72, bringing the total number of distinct samples in the assay to 79. For further description of such assays and the data they generate, see O’Malley & Deely (2003) and O’Malley et al. (in press).

Because the response counts are large (ranging from 3000 to 20 000), the conditional distribution of the response count, given the concentration in the sample, is likely to be closely approximated by a normal distribution (see O’Malley et al. (in press) for justification). Therefore, we assume that the response counts follow a normal distribution with the mean and variance functions in (2.1) and (2.2), respectively. The non-informative prior $\pi(\beta, \theta, \eta) \propto 1/\theta_1^{1/2}$ is used for Bayesian analysis; this is Jeffrey’s prior in the case when $\theta_2 = 0$ and it is assumed that $\beta$, $\theta$ and $\eta$ are independent, suggesting that the data will have close to maximal weight on the ensuing inferences. For brevity, in the following we omit details of the parameter estimates and posterior distribution of $(\beta, \theta)$ and, instead, focus on the precision profiles derived from the fitted model.
Figure 1 presents six different precision profiles for this serum thyroxine radioimmunoassay. Figure 1a, b shows the precision profiles based on the first- and second-order TSE in (a) $y$ only and (b) $y$ and $\beta$. Figure 1c is the precision profile determined via the empirical approach. Figure 1d is the Bayesian precision profile based on the CDAN. Precision is evaluated for two replicates using the coefficient of variation for the traditional definitions and the relative RMSE of the CDAN for the new Bayesian definition.

All of the precision profiles are decreasing functions for an initial range of concentrations and then are increasing functions over larger concentrations. Thus, each profile identifies a point of maximal relative precision and yields finite-width intervals of the concentrations that can be measured within various thresholds of precision (e.g. the 0.05 threshold is depicted in figure 1).

The first- and second-order analytical precision profiles in figure 1a, b are almost indistinguishable, implying that the addition of the second-order terms has little impact on the evaluated precision. Thus, results for the first- and second-order analytical approaches may be discussed as one.

The analytical precision profile in figure 1a and the empirical precision profile in figure 1c are close for $x_t \leq 150$ nmol l$^{-1}$ after which the trajectory of the empirical precision profile is flatter. To study this finding in more detail, note that the

*Phil. Trans. R. Soc. A* (2008)
estimated parameters for the model in (3.4) are $\hat{\alpha} = (1.0573, 0.1365, 2.6910)$, implying that the empirical precision profile has the form $2^{-1/2}x_f^{-1}(1.0573 + 0.1365x_f)^{1.3455}$. Therefore, as $x_f$ becomes large, precision becomes proportional to $x_f^{0.3455}$. Hence, the empirical precision profile is a slowly increasing function in $x_f$.

The precision profiles in figure 1a,c incorporate a translated form of the error in the response counts but ignore the variability in $\beta$, and hence are, not surprisingly, the lowest profiles. A precision profile that ignores the uncertainty in the fitted assay model would potentially be of interest when the statistical design (the number and concentration of the standard samples, and the number of replicated measurements performed on each sample) is not considered an attribute of the assay. In this case, precision reflects only the variability in the response counts. However, if as in this paper, the objective of the analysis is to quantify an assay’s ability to measure the concentration in an unknown sample, the uncertainty in the fitted model should be incorporated in the precision profile.

When the variability in $\hat{\beta}$ is accounted for, the analytical precision profiles (figure 1b) attain much higher values of precision, especially over larger concentrations, than the corresponding precision profiles that consider only the uncertainty in $\hat{y}_f$ (figure 1a,c). However, for $x_f \geq 80 \text{ nmol l}^{-1}$, the precision profile in figure 1b is well above the Bayesian precision profile in figure 1d. Thus, for a large range of concentrations, the analytical precision profile gives a more pessimistic assessment of the quality of the assay than the Bayesian (CDAN) precision profile. The discrepancy arises because the Bayesian precision profile uses all of the information in the assay and the impact of (differences in the level of) information is magnified when $x_f$ is large. The analytical method culls information by invoking simplifying assumptions and ignoring the dependence of $v$ on $m$ and the information about $\theta$ in replicated unknown samples.

A reassuring feature of the Bayesian precision profile is that it is a compromise between the precision profiles in figure 1a–c. It is similar in shape to the plots in figure 1a but, by virtue of accounting for the uncertainty in the fitted assay model, shifts upwards and has a greater curvature (but as noted above, these adjustments are not as great as for the analytical precision profile in figure 1b).

(a) CDAN for the NMCPH data

To study the precision of the assay in a broader sense, CDAN($x_f$) is plotted for $x_f = 0, 2, 5, 10, 100$ and $300 \text{ nmol l}^{-1}$ (figure 2a–f, respectively). As $x_f$ increases, the spread of CDAN($x_f$) increases. As expected, CDAN($x_f$) is skewed to the right when $x_f \approx 0$ and CDAN($x_f$) is nearly symmetrical when $x_f$ is large.

6. Conclusion

In this paper, Bayesian inference was used to develop a new precision profile of an assay. An advantage of the Bayesian CDAN precision profile is that mathematical approximations are avoided by using exact Bayesian inference (to the accuracy of the finite-length MCMC simulation) to evaluate moments of CDAN($x_f$) at each $x_f$. Therefore, the new Bayesian precision profile takes a full account of the information in both the standard and the unknown samples.
The Bayesian precision profile would appear to yield a more informative representation of the quality of an assay than traditional methods that rely on simplified forms of estimators or other mathematical approximations.

The new Bayesian precision profile was compared with existing procedures using a serum thyroxine radioimmunoassay. The differences in the results are non-trivial, particularly over medium to large concentrations, implying that the use of traditional precision profiles may give misleading representations of the assay quality.

The Bayesian method developed in this paper can be easily extended to accommodate analytical requirements involving multiple samples or laboratories by specifying a hierarchical model that allows certain parameters to vary between samples, assays or laboratories but linking them through a common underlying distribution that describes the differences between them. If the same assay is repeated on multiple occasions (e.g. each day of the year), it is possible that the concentration of the standard samples will degrade over time. To accommodate drift in the concentration of the standards, the model can be extended by allowing the parameters to change over time (a description of the importance of modelling carry-over and drift is given in Daniel 1975).

Figure 2. The concentration distribution of assay noise (CDAN) at $x_f = (a) 0$ (mean 0.641 and s.d. 0.869), (b) 2 (mean 1.916 and s.d. 1.312), (c) 5 (mean 4.978 and s.d. 1.513), (d) 10 (mean 10.02 and s.d. 1.53), (e) 100 (mean 100.2 and s.d. 4.3) and (f) 300 (mean 300.2 and s.d. 11.9) nmol l$^{-1}$ for a serum thyroxine radioimmunoassay performed at the NMCPH. The dotted vertical lines are the means of CDAN($x_f$). Note that the distributions are purposely not smoothed to emphasize that the model was fitted using simulation methods.

The Bayesian precision profile would appear to yield a more informative representation of the quality of an assay than traditional methods that rely on simplified forms of estimators or other mathematical approximations.

The new Bayesian precision profile was compared with existing procedures using a serum thyroxine radioimmunoassay. The differences in the results are non-trivial, particularly over medium to large concentrations, implying that the use of traditional precision profiles may give misleading representations of the assay quality.

The Bayesian method developed in this paper can be easily extended to accommodate analytical requirements involving multiple samples or laboratories by specifying a hierarchical model that allows certain parameters to vary between samples, assays or laboratories but linking them through a common underlying distribution that describes the differences between them. If the same assay is repeated on multiple occasions (e.g. each day of the year), it is possible that the concentration of the standard samples will degrade over time. To accommodate drift in the concentration of the standards, the model can be extended by allowing the parameters to change over time (a description of the importance of modelling carry-over and drift is given in Daniel 1975).
Appreciation is extended to William A. Sadler for supplying data from radioimmunoassay experiments at the NMCPH and to two anonymous reviewers for their valuable comments on an earlier version of the manuscript.

References


Phil. Trans. R. Soc. A (2008)