Add the following:

▲⟨1210⟩ STATISTICAL TOOLS FOR PROCEDURE VALIDATION

1. INTRODUCTION

This chapter describes utilization of statistical approaches in procedure validation as described in *Validation of Compendial Procedures* ⟨1225⟩. For the purposes of this chapter, "procedure validation" refers to the analytical procedure qualification stage of the method life cycle, following design and development and prior to testing.

Chapter ⟨1225⟩ explains that capabilities of an analytical procedure must be validated based on the intended use of the analytical procedure. Chapter ⟨1225⟩ also describes common types of uses and suggests procedure categories (I, II, III, or IV) based on the collection of performance parameters appropriate for these uses. Performance parameters that may need to be established during validation include accuracy, precision, specificity, detection limit [limit of detection, (LOD)], quantitation limit, linearity, and range. In some situations (e.g., biological assay), relative accuracy takes the place of accuracy. This chapter focuses on how to establish analytical performance characteristics of accuracy, precision, and LOD. For quantitative analytical procedures, accuracy can only be assessed if a true or accepted reference value is available. In some cases, it will be necessary to assess relative accuracy. In many analytical procedures, precision can be assessed even if accuracy cannot be assessed. The section addressing LOD can be applied to limit tests in Category II.

The other analytical performance characteristics noted in ⟨1225⟩, which include specificity, robustness, and linearity, are out of scope for this chapter.

Because validation must provide evidence of a procedure's fitness for use, the statistical hypothesis testing paradigm is commonly used to conduct validation consistent with ⟨1225⟩. Although some statistical interval examples are provided in *3. Accuracy and Precision*, these methods are not intended to represent the only approach for data analysis, nor to imply that alternative methods are inadequate.

Table 1 provides terminology used to describe an analytical procedure in this chapter. The definitions for individual determination and reportable value are in alignment with *General Notices, 7.10 Interpretation of Requirements*.

Table 1. Analytical Procedure Validation Terminology

Terminology	Description
Laboratory sample	The material received by the laboratory
Analytical sample	Material created by any physical manipulation of the laboratory sample, such as crushing or grinding
Test portion	The quantity (aliquot) of material taken from the analytical sample for testing
Test solution	The solution resulting from chemical manipulation of the test portion such as chemical derivatization of the analyte in the test portion or dissolution of the test portion
Individual determination (ID)	The measured numerical value from a single unit of test solution
Reportable value	Average value of readings from one or more units of a test solution

Not all analytical procedures have all stages shown in *Table 1*. For example, liquid laboratory samples that require no further manipulations immediately progress to the test solution stage. Demonstration that a reportable value is fit for a particular use is the focus of analytical validation.

Table 2 provides an example of the *Table 1* terminology for a solid oral dosage form.

Table 2. Example for Coated Tablets

Terminology	Description	
Laboratory sample	100 coated tablets	
Analytical sample	20 tablets are removed from the laboratory sample and are crushed in a mortar and pestle	
Test portion	Replicate 1: 1 g of crushed powder aliquot from the analytical sample	Replicate 2: 1 g of crushed powder aliquot from the analytical sample

2

Table 2. Example for Coated Tablets (continued)

Terminology	Description			
Test solution	Replicate 1: Test portion is dissolved in 1 L of solvent		Replicate 2: Test portion is dissolved in 1 L of solvent	
ID	ID 1 of replicate 1: test solution	ID 2 of replicate 1: test solution	ID 1 of replicate 2: test solution	ID 2 of replicate 2: test solution
Reportable value	Average value of four readings			

2. CONSIDERATIONS PRIOR TO VALIDATION

Procedure validation is a cornerstone in the process of establishing an analytical procedure. The aim of procedure validation is to demonstrate that the procedure, when run under standard conditions, will satisfy the requirement of being fit for use. To maximize the likelihood of a successful validation, it is imperative that all aspects of the procedure be well understood prior to the validation. Surprising discoveries (whether "good" or "bad") during validation should be carefully evaluated to determine whether the procedure was adequately developed. Moreover, pre-validation work can reveal suitable approaches to reduce the total size of the validation experiment without increasing the risk of drawing the wrong conclusion. General principles and plans for sample preparation, general principles, experimental design, data collection, statistical evaluation, and choice of acceptance criteria should be documented in a validation experimental protocol signed before initiation of the formal validation.

Questions considered prior to validation may include the following:
- What are the allowable ranges for operational parameters, such as temperature and time, that impact the performance of the analytical procedure?
  - Robustness of these ranges can be determined using a statistical design of experiments (DOE).
- What are the ruggedness factors that impact precision?
  - Factors such as analyst, day, reagent lot, reagent supplier, and instrument that impact the precision of a test procedure are called ruggedness factors. When ruggedness factors impact precision, reportable values within the same ruggedness grouping (e.g., analyst) are correlated. Depending on the strength of the correlation, a statistical analysis that appropriately accounts for this dependence may be necessary. Ruggedness factors can be identified empirically during pre-validation or based on a risk assessment.
- Are statistical assumptions regarding data analysis reasonably satisfied?
  - These assumptions may include such factors as normality, homogeneity of variance, and independence. It is useful during pre-validation to employ statistical tests or visual representations to help answer these questions. *Analytical Data—Interpretation and Treatment* ⟨1010⟩ provides information on this topic.
- What is the required range for the procedure?
  - The range of an analytical procedure is the interval between the upper and lower levels of an analyte that has been demonstrated to be determined with a suitable level of precision, accuracy, and linearity using the procedure as written.
- Do accepted reference values or results from an established procedure exist for validation of accuracy?
  - If not, as stated in International Council for Harmonisation (ICH) Q2, accuracy may be inferred once precision, linearity, and specificity have been established.
- How many individual determinations will compose the reportable value, and how will they be aggregated?
  - To answer this question, it is necessary to understand the contributors to the procedure variance and the ultimate purpose of the procedure. Estimation of variance components during pre-validation provides useful information for making this decision.
- What are appropriate validation acceptance criteria?
  - The validation succeeds when there is statistical evidence that the assay is no worse than certain pre-specified levels for each relevant validation parameter.
  - What defines the assay as fit for use, and how does this relate to acceptance criteria?
- How large a validation experiment is necessary?
  - Validation experiments should be properly powered to ensure that there are sufficient data to conclude that the accuracy and precision can meet pre-specified acceptance criteria. Computer simulation is a useful tool for performing power calculations.
  - Efficiencies (both cost and statistical) can be gained if assessment of linearity, accuracy, and precision can be combined.

On the basis of the answers to these and similar questions, one can design a suitable validation experimental protocol.

3. ACCURACY AND PRECISION

A useful model for representing a reportable value is:

$$Y = \tau + \beta + E \ (1)$$

Y = a reportable value
τ = true or accepted reference value
β = systematic bias of the procedure

E = random measurement error

Both τ (tau) and β (beta) are fixed statistical parameters, and E is a normal random variable with a mean of zero and standard deviation σ (sigma). The magnitude of σ depends on the number of individual readings averaged to obtain the reportable value.

Accuracy of an analytical procedure expresses the closeness of agreement between τ and Y. Closeness is expressed as the long-run average of $(Y - \tau)$. This long-run average is called the systematic bias and is represented with β. To estimate β, it is necessary to know the true value, τ. Chapter ⟨1225⟩ notes that a reference standard or a well-characterized orthogonal procedure can be used to assign the value of τ. Accuracy should be established across the required range of the procedure.

Precision of an analytical procedure is the degree of agreement among reportable values when the procedure is applied repeatedly (possibly under different conditions) to multiple test portions of a given analytical sample. The most common precision metric is the standard deviation σ. This is denoted in *Equation 4* with the variable S. The term σ^2 is called the variance. Precision improves as σ decreases. Many commonly used statistical procedures rely on the assumption of the normal distribution, for which σ is a natural descriptor of variability.

Change to read:

3.1 Methods for Estimating Accuracy and Precision

An example is provided to demonstrate the test procedure for lot release using statistical analysis. This example uses high-performance liquid chromatography (HPLC). The measured drug substance (DS) is a *USP* compendial substance, so information concerning τ is available (*1*). Three different quantities of reference standard were weighted to correspond to three different percentages of the test concentrations: 50%, 100%, and 150%. The unit of measurement on each reportable value is the mass fraction of DS expressed in units of mg/g and does not change as the level of concentration varies. The value of τ is 1000 mg/g for all three concentrations. The computed statistics from the validation data set include the sample mean ($\overline{Y}$), the sample standard deviation (S), and the number of reportable values (n). *Table 3* presents the $n = 9$ reportable values and the computed statistics.

Table 3. Reportable Values for Experiment

Test Concentration (%)	Test Solution	Reportable Value (mg/g)
50	1	996.07
50	2	988.43
50	3	995.90
100	4	987.22
100	5	990.53
100	6	999.39
150	7	996.33
150	8	993.67
150	9	987.76
Sample mean ($\overline{Y}$)		992.81
Sample standard deviation (S)		4.44

Several assumptions are made for purposes of this example, which allows analysis of the combined data set in *Table 3*:
1. All $n = 9$ reportable values are independent.
2. The standard deviation of the reportable value is constant across all three concentration levels. If this condition is not met, data transformations may still allow combination of all the data in *Table 3* (pooling). If transformations are not successful, each concentration level must be validated for precision separately.
3. The average reportable value is equal across concentration levels. If this condition does not hold, it is necessary to employ an analysis of variance model and validate accuracy for each concentration level separately.

The point estimator for unknown bias β is:

$$\beta = \overline{Y} - \tau \ (2)$$

β = systematic bias
$\overline{Y}$ = sample mean
τ = true or accepted reference value

where

Printed on: Tue Apr 12 2022, 04:29:24 PM(EST)   Official Status: Currently Official on 12-Apr-2022   DocId: 1_GUID-13ED4BEB-4086-43B5-A7D7-994A02AF25C8_6_en-US
Printed by: Lucas Croffi   Official Date: Official as of 01-May-2018   Document Type: GENERAL CHAPTER   @2022 USPC
                                                       DOI Ref: 8kmx6   DOI: https://doi.org/10.31003/USPNF_M8646_06_01
4

$$\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \qquad (3)$$

$\overline{Y}$ = sample mean
Y_i = individual values
n = number of reportable values

The point estimator for the unknown value of σ is

$$S = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2}{n-1}} \qquad (4)$$

S = point estimator for the unknown value of σ
Y_i = individual values
$\overline{Y}$ = sample mean
n = number of reportable values

Because point estimators have uncertainty associated with them, best practice requires calculation of a statistical confidence interval to quantify the uncertainty. Statistical confidence intervals provide a range of plausible values for β and σ for a given level of confidence. A 100(1 − 2α)% two-sided interval for the bias β is

$$(\overline{Y} - \tau) \pm t_{1-\alpha:n-1} \times \frac{S}{\sqrt{n}} \qquad (5)$$

β   = 100(1 − 2α)% two-sided confidence interval of bias
$t_{1-\alpha:n-1}$ = percentile of central t-distribution with area 1 − α to the left and (n − 1) degrees of freedom
S = result found from *Equation 4*
n = number of reportable values

For example, with α = 0.05 and n = 9, $t_{0.95:8}$ = 1.860 provides a 100(1 − 2 × ▲0.05▲ (ERR 1-Mar-2018))% = 90% two-sided confidence interval for β. Using the example data in *Table 3* with τ = 1000 mg/g, the 90% confidence interval on β is

$$(\overline{Y} - \tau) \pm t_{1-\alpha:n-1} \times \frac{S}{\sqrt{n}}$$

$$(992.81 - 1000) \pm 1.86 \times \frac{4.44}{\sqrt{9}} \qquad (6)$$

$$[-9.94 \text{ to} -4.44] \text{ mg/g}$$

β   = 100(1 − 2α)% two-sided confidence interval of bias
$t_{1-\alpha:n-1}$ = percentile of central t-distribution with area 1 − α to the left and (n − 1) degrees of freedom
S = result found from *Equation 4*

For the standard deviation, one is concerned with only the $100(1 - \alpha)\%$ upper confidence bound since typically, it needs to be shown that the standard deviation is not too large. An upper $100(1 - \alpha)\%$ confidence bound for σ is

$$U = S\sqrt{\frac{n-1}{\chi^2_{\alpha:n-1}}} \qquad (7)$$

U	= an upper $100(1 - \alpha)\%$ confidence bound for σ
S	= result found from *Equation 4*
n	= number of reportable values
$\chi^2_{\alpha:n-1}$	= a percentile of a central chi-squared distribution with area α to the left and $(n - 1)$ degrees of freedom

For example, if $\alpha = 0.05$ and $n = 9$, then $\chi^2_{0.05:8} = 2.73$. Using the data in *Table 3*,

$$U = S\sqrt{\frac{n-1}{\chi^2_{\alpha:n-1}}}$$

$$U = 4.44\sqrt{\frac{9-1}{2.73}} = 7.60 \text{ mg/g} \qquad (8)$$

The confidence intervals in *Equations 5* and *7* can be used to perform statistical tests against criteria included in the validation protocol. Use of point estimates only does not provide the required scientific rigor. In particular, the two-sided confidence interval in *Equation 5* can be used to perform a two one-sided test (TOST) of statistical equivalence (*2*). Assume in the present example that the accuracy requirement is validated if evidence demonstrates that the absolute value of β is NMT 15 mg/g. Since the computed confidence interval from -9.94 to -4.44 mg/g falls entirely within the range from -15 to $+15$ mg/g, the bias criterion is satisfied. Most typically, the TOST employs a type I error rate of $\alpha = 0.05$. This error rate represents the maximum risk of declaring that the acceptance criterion is satisfied, when in truth it is not satisfied. Thus, with $\alpha = 0.05$, the two-sided confidence interval in *Equation 5* is $100(1 - 2\alpha)\% = 90\%$.

The upper bound in *Equation 7* is used to validate precision. Suppose the pre-defined acceptance criterion for precision requires σ to be <20 mg/g. The computed upper bound of 7.60 mg/g in *Equation 8* represents the largest value we expect for σ with 95% confidence. Since 7.60 mg/g is <20 mg/g, precision has been successfully validated with a confidence of 95%.

Change to read:

3.2 Combined Validation of Accuracy and Precision

When assessing whether an analytical procedure is fit for its intended purpose, it is often useful to consider the combined impact of bias and precision. The degree to which β impacts the usefulness of an analytical procedure depends in part on σ. That is, a procedure with a relatively small value of σ can accommodate a relatively greater value of β than a procedure with a greater value of σ. For this reason, it is useful to establish a single criterion that can be used to simultaneously validate both accuracy and precision. One such criterion is proposed in a series of articles by Hubert et al. (*3–5*) and seeks to ensure that

$$Pr(-\lambda < Y - \tau < \lambda) \geq P, \text{ or}$$

$$Pr(-\lambda + \tau < Y < \lambda + \tau) \geq P \quad (9)$$

Pr	= reference probability
λ	= acceptable limit
Y	= a reportable value
τ	= true or accepted reference value
P	= desired probability value

Equation 9 has a dual interpretation. It can be interpreted as either (i) the probability that the next reportable value falls in the range from $(-\lambda + \tau)$ to $(\lambda + \tau)$ is $\geq P$, or (ii) the proportion of all future reportable values falling between $(-\lambda + \tau)$ and $(\lambda + \tau)$ is $\geq P$. Accordingly, two statistical intervals have been proposed to demonstrate that *Equation 9* is true:
1. A prediction interval (also referred to as an expectation tolerance interval) is used to demonstrate (i).
2. A tolerance interval (also referred to as a content tolerance interval) is used to demonstrate (ii).

Hahn and Meeker (*6*) note that the prediction interval is also referred to as an expectation tolerance interval and that the tolerance interval is also referred to as a content tolerance interval. Because the inference associated with the tolerance interval

6

concerns a larger set of values, the tolerance interval is always wider than the prediction interval. Selection of an interval will depend on the desire to validate ▲▲ (ERR 1-May-2018) either (i) or (ii) and a company's risk profile.

Either interval can be used in the following manner to evaluate accuracy and precision simultaneously through *Equation 9*.

1. Compute the appropriate statistical interval using *Equation 10* for the prediction interval and *Equation 11* for the tolerance interval.
2. If the computed interval falls completely in the range from $(-\lambda + \tau)$ to $(\lambda + \tau)$, the criterion in *Equation 9* is satisfied, and the procedure is validated for both accuracy and precision.

The prediction interval used to validate *Equation 9* is

$$\overline{Y} \pm t_{(1+P)/2:n-1} \times S\sqrt{1+\frac{1}{n}} \qquad (10)$$

$\overline{Y}$ = sample mean
$t_{(1+P)/2:n-1}$ = percentile of a central *t*-distribution with area $(1 + P)/2$ to the left and $(n - 1)$ degrees of freedom
S = result found from *Equation 4*

The $100(1 - \alpha)\%$ tolerance interval used to validate *Equation 9* is

$$\overline{Y} \pm K \times S$$

$$K = \sqrt{\frac{Z^2_{(1+P)/2} \times (n-1)}{\chi^2_{\alpha:n-1}} \times \left(1+\frac{1}{n}\right)} \qquad (11)$$

$\overline{Y}$ = sample mean
K = result found from *Equation 11*
S = result found from *Equation 4*
$Z^2_{(1+P)/2}$ = ▲the square of the standard normal▲ (ERR 1-May-2018) percentile with area $(1 + P)/2$ to the left
n = number of reportable values
$\chi^2_{\alpha:n-1}$ = a chi-squared percentile with area α to the left and $(n - 1)$ degrees of freedom

The formula for *K* is based on an approximation by Howe (*7*), although exact tabled values can be found in several sources. The approximation works well in practical situations if exact values are not available.

For the data in *Table 3* with $P = 0.90$, the interval for *Equation 10* is computed as

$$\overline{Y} \pm t_{(1+P)/2,n-1} \times S\sqrt{1+\frac{1}{n}}$$

$$992.81 \pm 1.86 \times 4.44\sqrt{1+\frac{1}{9}} \qquad (12)$$

$$[984.1 \text{ to } 1001.5] \text{ mg/g}$$

The interval for *Equation 11* with $1 - \alpha = 0.90$ and $P = 0.90$ is

$$K = \sqrt{\frac{Z_{(1+P)/2}^2 \times (n-1)}{\chi_{\alpha:n-1}^2} \times \left(1 + \frac{1}{n}\right)}$$

$$K = \sqrt{\frac{(1.64)^2 \times (9-1)}{3.49} \times \left(1 + \frac{1}{9}\right)} = 2.63$$

$$\overline{Y} \pm K \times S$$

$$992.81 \pm 2.63 \times 4.44$$

$$[981.2 \; to \; 1004.5] \; mg/g \tag{13}$$

The exact value for K is 2.637, and the approximation is seen to work quite well. As predicted earlier, the interval for *Equation 13* is wider than the interval for *Equation 12*.

Suppose the criterion for *Equation 9* is designed to ensure that the difference between Y and τ is <2% of τ with a probability NLT $P = 0.90$. Thus,

$$-\lambda + \tau = \tau(-0.02 + 1) = 1000(0.98) = 980 \; mg/g$$

$$\tau + \lambda = \tau(1 + 0.02) = 1000(1.02) = 1020 \; mg/g \tag{14}$$

Since both *Equations 12* and *13* fall in the range from 980 to 1020 mg/g, the procedure is validated using either interval.

It is also possible to estimate $Pr(-\lambda < Y - \tau < \lambda)$ in *Equation 9* directly using either the confidence interval described by Mee (*8*) or a Bayesian approach. The validation criterion is thus satisfied if this estimated probability exceeds P. A Bayesian tolerance interval is provided in Wolfinger (*9*) and can be computed using the statistical software package WinBUGS (*10,11*). Bayesian analyses can be challenging, and the aid of an experienced statistician is recommended.

4. LIMITS OF DETECTION AND QUANTITATION

The LOD and limit of quantitation (LOQ) are two related quantities determined in the validation of Category II procedures of ⟨1225⟩. These are procedures for the determination of impurities or degradation products in DS and finished pharmaceutical products. Only one is needed for each use: LOQ for quantitative tests and LOD for qualitative limit tests. These limits are also known under other names, including detection limit (DL) for LOD and lower limit of quantitation (LLOQ) for LOQ.

The following definitions are consistent with ⟨1225⟩ and ICH Q2:
- The LOD is the lowest amount of analyte in a sample that can be detected, but not necessarily quantitated, under the stated experimental conditions.
- The LOQ is the lowest amount of analyte in a sample that can be determined with acceptable precision and accuracy under the stated experimental conditions.

Candidate values for LOD or LOQ are examined during pre-validation or based on a risk assessment. The candidate values must then be verified. This is particularly important for LOQ, since the formulas for determining candidate values do not address the acceptable accuracy and precision requirement. Verification of the candidate values is performed as part of the validation protocol.

4.1 Estimation of LOD

The basic approach to estimating LOD is based on an alternative definition adopted by the International Union of Pure and Applied Chemistry (IUPAC) and the International Organization for Standardization (ISO). This definition introduces the notion of false-positive and false-negative decisions, thus recognizing the risk elements in using the LOD for decision making, and the definition makes clear that these values are dependent on laboratory capability.

The IUPAC/ISO definition of LOD is based on the underlying concept of a critical value (R_C), defined as the signal readout exceeded with probability α when no analyte is present. That is,

$$R_C = B + Z_{1-\alpha}\sigma_E \tag{15}$$

R_C = signal readout exceeded with probability α when no analyte is present
B = estimated mean readout for blanks
$Z_{1-\alpha}$ = a standard normal quantile with area $1 - \alpha$ to the left
σ_E = true repeatability deviation

Figure 1 presents this relationship graphically.

Printed on: Tue Apr 12 2022, 04:29:24 PM(EST)   Official Status: Currently Official on 12-Apr-2022   DocId: 1_GUID-13ED4BEB-4086-43B5-A7D7-994A02AF25C8_6_en-US
Printed by: Lucas Croffi   Official Date: Official as of 01-May-2018   Document Type: GENERAL CHAPTER   @2022 USPC
DOI Ref: 8kmx6   DOI: https://doi.org/10.31003/USPNF_M8646_06_01
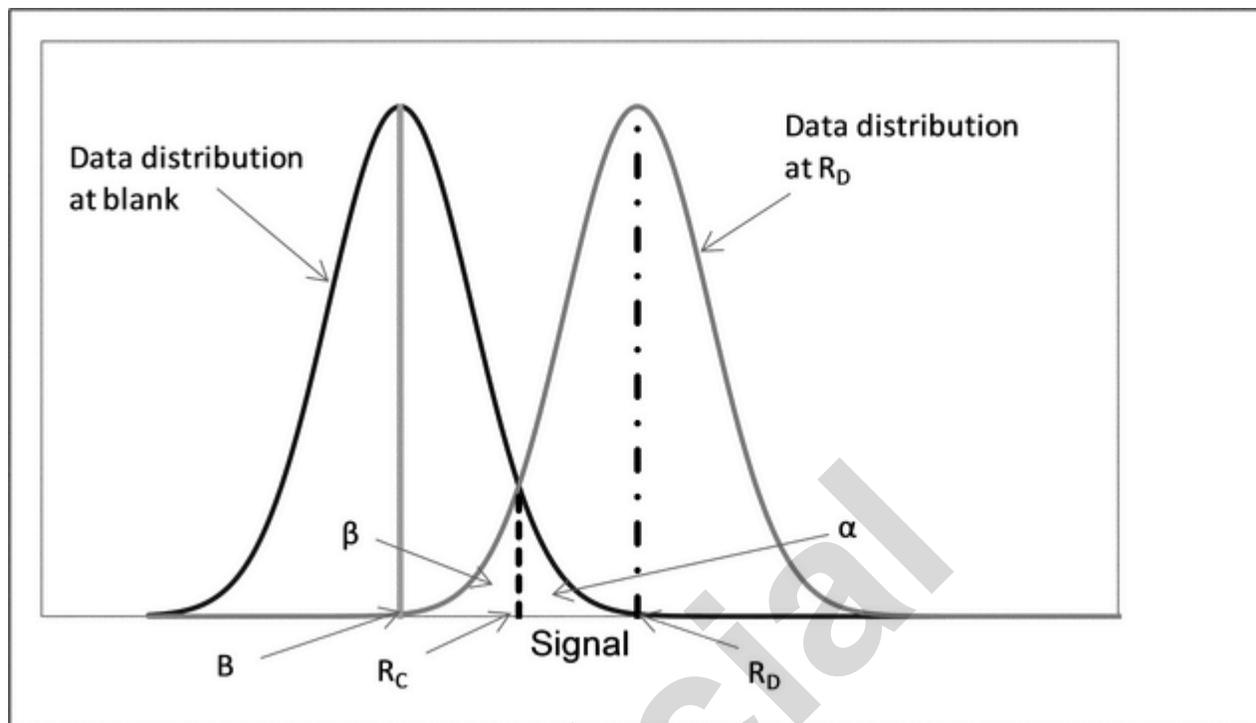8

Figure 1. Determination of R_C and R_D.

For example, if $\alpha = 0.05$, $1 - \alpha = 0.95$ and $Z_{0.95} = 1.645$. This determination depends on the distribution of values obtained when analyzing blanks. The LOD in the signal space (R_D) is defined as that value, which if true, is such that R_C is exceeded with probability $1 - \beta$. That is,

$$R_D = R_C + Z_{1 - \beta}\sigma_E \text{ (16)}$$

R_D = LOD in the signal space
R_C = critical value using IUPAC/ISO definition of LOD
$Z_{1 - \beta}$ = standard normal quantile with area $1 - \beta$ to the left
σ_E = true repeatability deviation

Solving *Equations 15* and *16* for R_D, we have

$$R_D = B + (Z_{1 - \alpha} + Z_{1 - \beta})\sigma_E \text{ (17)}$$

R_D = LOD in the signal space
B = estimated mean readout for blanks
$Z_{1 - \alpha}$ = a standard normal quantile with area $1 - \alpha$ to the left
$Z_{1 - \beta}$ = a standard normal quantile with area $1 - \beta$ to the left
σ_E = true repeatability deviation

Note that this definition allows for two values to be selected by the laboratory: α and β (which need not be equal). The α represents the type I or false-positive error rate, and β represents the type II or false-negative error rate. In *Figure 1*, R_C and R_D are illustrated with $\alpha = \beta = 0.05$ for normally distributed data so that $Z_{1 - \alpha} = Z_{1 - \beta} = 1.645$. Although the values of α and β need not be equal, this choice leads to a common rule for R_D, namely $B + 3.3\sigma_E$ ($3.3 \cong 2 \times 1.645$).
   The LOD on the concentration scale is then found by converting the value in the signal scale, R_D, to one in the concentration scale, LOD, as shown in *Figure 2*.
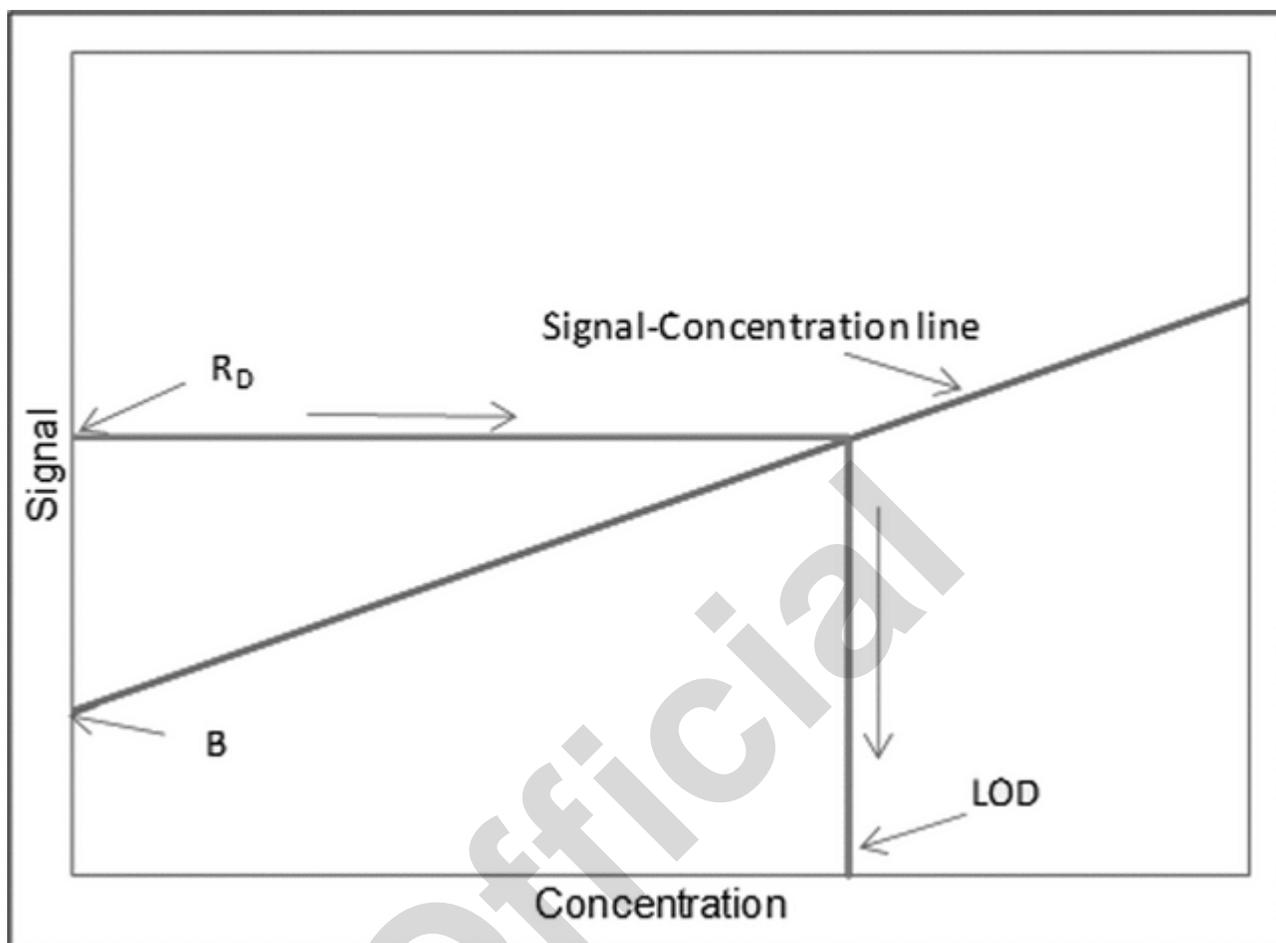
Figure 2. Determination of LOD from R_D.

This step requires that the signal (R) versus concentration (X) line, $R = B + mX$, as well as σ_E, be known exactly. The formulation provided in this section assumes the regression measurements are independent.

The LOD on the concentration scale is then calculated as

$$LOD = \frac{R_D - B}{m} = \frac{(Z_{1-\alpha} + Z_{1-\beta})\sigma_E}{m} \tag{18}$$

LOD   = limit of detection
R_D = LOD in the signal space
B = estimated mean readout for blanks
$Z_{1-\alpha}$ = standard normal quantile with area $1 - \alpha$ to the left
$Z_{1-\beta}$ = standard normal quantile with area $1 - \beta$ to the left
σ_E = true repeatability deviation
m = slope

As a statistical procedure, the LOD definition in *Equation 18* is unsatisfactory for two reasons. First, since σ_E is generally unknown, it must be determined how to best estimate this parameter. This is complicated because σ_E is typically concentration dependent. Two common estimates are (i) the standard deviation of the blank responses and (ii) the standard deviation obtained from deviations about the regression line of signal on concentration. The choice needs to be the value that best represents σ_E in the neighborhood of the LOD. Laboratories will often pick a worst-case value for σ_E. If the LOD is still suitable for its intended use, the laboratories are protected against understating the LOD. Understatement of the LOD results in an inflated type II error rate (β) and a deflated type I error rate (α).

The second statistical concern with *Equation 18* is how to incorporate uncertainty due to the fact that the exact slope of the regression line of signal on concentration is unknown. Because the regression line is estimated, the definition of R_D in *Equation*

17 is itself an estimate. This is corrected by using a statistical prediction interval that takes into account the uncertainty in the estimated line as well as the variability associated with a future observation. The expanded formula for the critical value, R_C, originally defined in *Equation 15* that accounts for this uncertainty is

$$R_c = B + t_{1-\alpha:n-2} \times S \sqrt{1 + \frac{1}{n} + \frac{\overline{X}^2}{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}} \qquad (19)$$

$$S = \sqrt{\frac{\sum\limits_{i=1}^{n}(R_i - B - mX_i)^2}{n-2}}$$

R_C　　= critical value using the IUPAC/ISO definition of LOD
B　　= intercept of the fitted calibration line
$t_{1-\alpha:n-2}$ = percentile of a central *t*-distribution with area $1 - \alpha$ to the left and $(n - 2)$ degrees of freedom
S　　= standard error of regression line
$\overline{X}$　　= average concentration
n　　= number of observations used in the regression analysis
X_i　　= concentration value used in determining the line
R_i　　= reference interval
m　　= slope

Equation 19 differs from *Equation 15* because the *t*-distribution is used instead of the normal distribution for the multiplier, and two additional terms appear in the square root to capture the uncertainty of the regression line.

A second equation for R_C answers the question, "Above which concentration can we be confident that we will obtain signals that are distinguishable from background?" This question is answered by using the lower $100(1 - \beta)\%$ prediction bound of the calibration curve as shown in *Figure 3*.
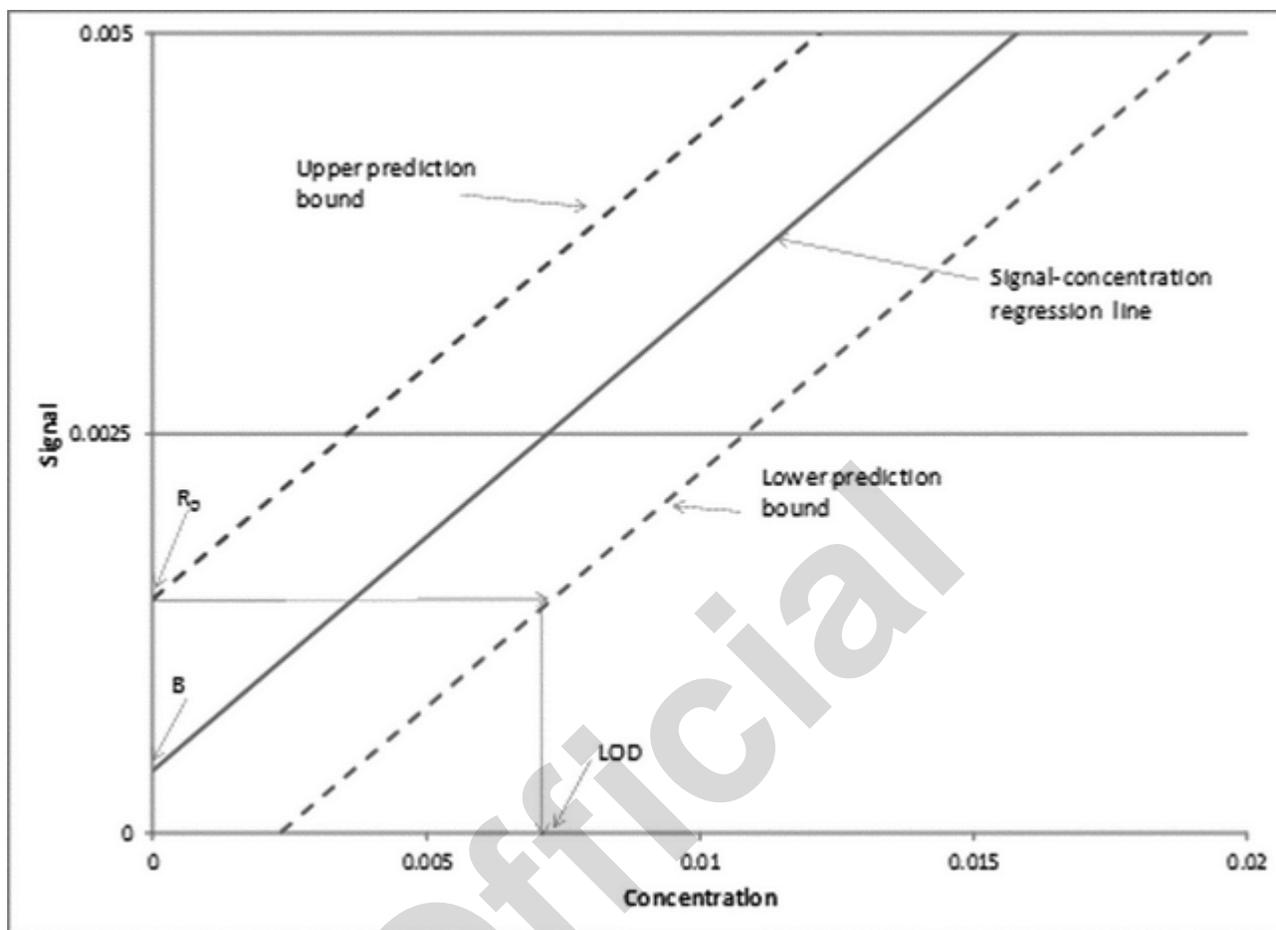
Figure 3. Determination of LOD using prediction bounds.

Figure 3 is similar to *Figure 2*, but uses two dashed curves instead of the solid calibration line. Here

$$R_C = B + LOD \times m - t_{1-\beta:n-2} \times S \sqrt{1 + \frac{1}{n} + \frac{(LOD - \overline{X})^2}{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}} \qquad (20)$$

R_C = critical value using the IUPAC/ISO definition of LOD
B = estimated intercept of the fitted calibration line
LOD = limit of detection
m = slope
$t_{1-\beta:n-2}$ = percentile of a central *t*-distribution with area $1 - \beta$ to the left and $(n - 2)$ degrees of freedom
S = standard error of regression line
$\overline{X}$ = average concentration value squared
n = number of observations used in the regression analysis
X_i = concentration value used in determining the line

After equating *Equation 19* and *Equation 20*, and cancelling the *B* terms,

$$t_{1-\alpha:n-2} \times S \sqrt{1 + \frac{1}{n} + \frac{\overline{X}^2}{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}} = LOD \times m - t_{1-\beta:n-2} \times S \sqrt{1 + \frac{1}{n} + \frac{(LOD - \overline{X})^2}{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}} \qquad (21)$$

$t_{1-\alpha:n-2}$ = percentile of a central *t*-distribution with area $1 - \alpha$ to the left and $(n - 2)$ degrees of freedom

Printed on: Tue Apr 12 2022, 04:29:24 PM(EST)   Official Status: Currently Official on 12-Apr-2022   DocId: 1_GUID-13ED4BEB-4086-43B5-A7D7-994A02AF25C8_6_en-US

Printed by: Lucas Croffi   Official Date: Official as of 01-May-2018   Document Type: GENERAL CHAPTER   @2022 USPC

      DOI Ref: 8kmx6   DOI: https://doi.org/10.31003/USPNF_M8646_06_01

12

S = standard error of regression line

$\overline{X}^2$ = average concentration value squared

LOD   = limit of detection

m = slope

$t_{1-\beta:n-2}$ = percentile of a central t-distribution with area $1 - \beta$ to the left and $(n - 2)$ degrees of freedom

n = number of observations used in the regression analysis

X_i = concentration value used in determining the line

$\overline{X}$ = average concentration value

Equation 21 is a quadratic equation for LOD that can be solved exactly or by using iterative search tools available in spreadsheets. A slightly conservative (overly large) approximation for LOD that does not require a quadratic solution is obtained by assuming that LOD is negligible compared to $\overline{X}$ [i.e., $(LOD - \overline{X})^2$ is replaced with $\overline{X}^2$]. The resulting equation under this simplification is

$$LOD = (t_{1-\alpha:n-2} + t_{1-\beta:n-2}) \times \frac{S}{m} \sqrt{1 + \frac{1}{n} + \frac{\overline{X}^2}{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}} \qquad (22)$$

LOD   = limit of detection

$t_{1-\alpha:n-2}$ = percentile of a central t-distribution with area $1 - \alpha$ to the left and $(n - 2)$ degrees of freedom

$t_{1-\beta:n-2}$ = percentile of a central t-distribution with area $1 - \beta$ to the left and $(n - 2)$ degrees of freedom

S = standard error of regression line

$\overline{X}$ = average concentration value

m = slope

n = number of observations used in the regression analysis

X_i = concentration value used in determining the line

which is similar in form to *Equation 18*. *Equations 18* and *22* both allow the two error probabilities, α and β, to differ. Often they are both taken as equal to 0.05.

The data in *Table 4* are used to demonstrate calculation of the LOD.

Table 4. Data for LOD Example

Concentration X (mg/mL)	Area (signal)
0.01	0.00331
0.02	0.00602
0.05	0.01547
0.10	0.03078
0.15	0.04576
0.25	0.07592

Fitting the linear regression to these data yields the regression line:

Area = 0.000235 + 0.3032 × Concentration (23)

so that $m = 0.3032$ and $B = 0.000235$. Values needed to compute LOD shown in *Equation 22* with $\alpha = \beta = 0.05$ are provided in *Table 5*.

Table 5. Statistics Needed to Compute LOD in Concentration Units

Statistic	Value
n	6
m (slope)	0.3032
S	0.00019
$t_{1-\alpha:n-2} = t_{0.95:4}$	2.132
$t_{1-\beta:n-2} = t_{0.95:4}$	2.132
$\overline{X}$	0.0967

Table 5. Statistics Needed to Compute LOD in Concentration Units *(continued)*

Statistic	Value
$$\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$	0.0419

The value of LOD computed from *Equation 22* is

$$LOD = (t_{1-\alpha:n-2} + t_{1-\beta:n-2}) \times \frac{S}{m}\sqrt{1 + \frac{1}{n} + \frac{\overline{X}^2}{\Sigma_{i=1}^{n}(X_i - \overline{X})^2}}$$

$$LOD = (2.132 + 2.132) \times \frac{0.00019}{0.3032}\sqrt{1 + \frac{1}{6} + \frac{0.0967^2}{0.0419}} \qquad (24)$$

$$LOD = 0.0032 \text{ mg/mL}$$

LOD = limit of detection
$t_{1-\alpha:n-2}$ = percentile of a central *t*-distribution with area $1 - \alpha$ to the left and $(n - 2)$ degrees of freedom
$t_{1-\beta:n-2}$ = percentile of a central *t*-distribution with area $1 - \beta$ to the left and $(n - 2)$ degrees of freedom
S = residual standard error found from *Equation 4*
$\overline{X}^2$ = average concentration squared
m = slope
n = number of observations used in the regression analysis
X_i = concentration value used in determining the line

4.2 Estimation of LOQ

As previously discussed, the important consideration in determining the LOQ is the estimation of what LOQ is required based on the intended use. The validation is designed to validate accuracy and precision in the neighborhood of the required LOQ.

In the absence of such knowledge, or where the laboratory wants to determine how low the LOQ might be (e.g., for potential other uses), then the laboratory can start with potential LOQ values greater than but near the LOD. Alternatively, methods for determining the LOD can be adapted to the LOQ as candidate starting values. Essentially, the formula used to compute LOD in *Equation 22* can be used to compute LOQ by replacing $(t_{1-\alpha:n-2} + t_{1-\beta:n-2})$ with 10. Values other than 10 can be used if justified. Once candidate values are obtained (typically during pre-validation), accuracy and precision are validated at these values.

All of the methods presented in this section are based on two assumptions: linearity and homogeneity of variance across the range of concentrations used in determining the calibration curve. Neither is a necessary condition. The calibration curve may be nonlinear, and a weighted least squares approach can be used to account for a lack of homogeneity. If the curve is nonlinear or the concentration variances vary greatly in the range of the LOD and LOQ, it is best to seek expert statistical help in defining LOD and LOQ. If variability about a straight line exists but is not large, an unweighted regression of the calibration curve will provide an average variability that can be used in the LOD and LOQ formulas.

Procedures other than those described above, such as signal-to-noise ratios, can be used to estimate LOD and LOQ. In either case, analysts should consider these values as preliminary and proceed to verify them, particularly if they fall below the concentration values used in determining the calibration curve. Verification involves analyzing samples with concentrations near the preliminary LOD and LOQ. Consideration should be given to how low must an LOD and LOQ be such that the procedure is suitable. For example, if data are already available at a level below the required LOD and a signal was detectable at that lower value, then that lower value may be taken as a verified LOD. There is little value in further verification—given the current requirement. There could still be value in verification of a lower value in case the requirement changes.

5. CONCLUDING REMARKS

This chapter presented some simple statistical methods that can be used in procedure validation as described in ⟨1225⟩. These methods may not be applied in all situations, and other statistical approaches, both more and less sophisticated, may be appropriate for any particular situation.

Re-evaluation of a procedure should be considered whenever use of the procedure changes. For example, if a new product strength is introduced, the procedure is transferred to a new lab, samples are to be tested following a new type of stress test, or specifications change, a re-validation is most likely appropriate. In some situations, a re-assessment of existing data to revised acceptance limits is sufficient.

Finally, although not part of procedure validation, it is recommended that some type of statistical process control be used to monitor the performance of the procedure. Such a process provides early warning of "drift" in the analytical performance characteristics of accuracy and precision. Such changes in performance are not uncommon, and often occur as a result of worn equipment, change of routines, or aging reagents.

REFERENCES

1. Weitzel MLJ. The estimation and use of measurement uncertainty for a drug substance test procedure validated according to USP ⟨1225⟩. *Accred Qual Assur.* 2012;17:139–146.
2. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Biopharmacokinet Biopharmaceut.* 1987;15:657–680.
3. Hubert P, Nguyen-Huu JJ, Boulanger B, Chapuzet E, Chiap P, Cohen N, et al. Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal—part I. *J Pharm Biomed Anal.* 2004;36(3):579–586.
4. Hubert P, Nguyen-Huu JJ, Boulanger B, Chapuzet E, Chiap P, Cohen N, et al. Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal—part II. *J Pharm Biomed Anal.* 2007;45(1):70–81.
5. Hubert P, Nguyen-Huu JJ, Boulanger B, Chapuzet E, Chiap P, Cohen N, et al. Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal—part III. *J Pharm Biomed Anal.* 2007;45(1):82–96.
6. Hahn GJ, Meeker WQ. *Statistical Intervals: A Guide for Practitioners.* New York: John Wiley & Sons; 1991.
7. Howe WG. Two-sided tolerance limits for normal populations—Some improvements. *J Am Stat Assoc.* 1969;64(326):610–620.
8. Mee RW. Estimation of the percentage of a normal distribution lying outside a specified interval. *Commun Stat Theory Methods.* 1988;17(5):1465–1479.
9. Wolfinger RD. Tolerance intervals for variance component models using Bayesian simulation. *J Qual Technol.* 1998;30(1):18–32.
10. Ntzoufras I. *Bayesian Modeling Using WinBUGS.* New York: John Wiley & Sons; 2009.
11. Spiegelhalter D, Thomas A, Best N, Gilks W. *BUGS 0.5 Examples: Volume 1 (version i).* Cambridge, UK: MRC Biostatistics Unit; 1996. http://users.aims.ac.za/~mackay/BUGS/Manual05/Examples1/bugs.html. Accessed 12 Jul 2016.

▲ (USP41)