ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
16 (1998) 1125–1141

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

Statistical approaches to determine analytical variability and specifications: application of experimental design and variance component analysis[1]

P.K.S. Tsang [a,*], J.S.A. Larew [b], L.A. Larew [a], T.W. Miyakawa [c], J.D. Hofer [c]

[a] *Lilly Research Laboratories, Eli Lilly and Company, PO Box 685, Lafayette, IN 47902, USA*
[b] *Tippecanoe Laboratories, Eli Lilly and Company, PO Box 685, Lafayette, IN 47902, USA*
[c] *Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, USA*

Received 12 May 1997; received in revised form 17 September 1997

Abstract

Assessment of analytical variability is recognized as an important factor for the establishment of specifications. Estimation of the variance for an analytical procedure can be accomplished using a variety of approaches. The approach of variance component analysis was applied retrospectively, as well as prospectively, to estimate analytical variance. The prospective approach also included the use of experimental design. Recent new drug substance examples illustrating these approaches are presented. In these examples, the analytical property of potency was evaluated. Factors examined in the experimental design include laboratory, day, analyst, instrument and column. Process variability can also be determined by variance component analysis. For a stable drug substance, combining the analytical and process variances provides an estimate on the total variance for the analytical property of potency. With the total variability statistically derived, an appropriate specification that is consistent with process and analytical capability can be established.

Keywords: Analytical variability; Experimental design; Variance component analysis; Statistically derived specifications

1. Introduction

Analytical variability is a generally recognized concept in the analysis of chemical goods [1,2], including pharmaceutical products [3–5]. Identifying the sources of analytical variability and the magnitude of the variability is critical in the establishment of product specifications. The need to consider analytical as well as manufacturing process variability is consistent with the present International Conference on Harmonization (ICH) effort on providing a guideline for establishing specifications of synthetic new drug substances and new drug products, Topic Q6A. One of the

* Corresponding author. Tel.: +1 765 4774602; fax: +1 765 4774887; e-mail: tsang_paul_k@lilly.com

[1] Presented at the Eighth International Symposium on Pharmaceutical and Biomedical Analysis (PBA '97), Orlando, FL, USA, 4–8 May, 1997.

0731-7085/98/$19.00
PII S0731-7085(97)00266-5

Table 1
Chromatographic conditions for potency determination of NDS A and NDS B

Condition	NDS A	NDS B
Mobile phase	Acetonitrile–sodium phosphate/SDS (pH 2.5; 50 mM/42 mM; 50:50 v/v)	Acetonitrile–potassium phosphate (Ph 2.5, 75 mM; 33:67 v/v)
Column	Zorbax® SB-C8; 15 cm × 4.6 mm i.d., 5 μm particle size	Zorbax® SB-C8; 15 cm × 4.6 mm i.d., 3.5 μm particle size
Column temperature (°C)	Ambient temperature (approximately 25)	35
Flow rate (ml min^{-1})	1.5	1.5
Injection volume (μl)	10	10
Detection	UV at 260 nm	UV at 280 nm

challenges in establishing specifications for a new product is the long-term commitment to regulatory specifications based on a limited size sample set of development, scale-up and production results.

Understanding the performance capability of the analytical method and manufacturing process provides a basis for the development and establishment of specifications. Statistical methods are often employed in gaining a better understanding of performance capability. In addition to data obtained from the analytical or process development areas, data from the quality control laboratories or manufacturing sites should be included in the statistical analysis. This becomes more critical when multiple laboratories or sites are involved in supplying the product. If multiple quality control laboratories use the analytical method, an assessment of the method performance can ensure each quality control laboratory will be able to appropriately control the product. While the ICH guidelines on method validation do not require reproducibility to be included in the method validation [6,7], the evaluation of method reproducibility has been advocated [4]. Furthermore, in the present paradigm of increasing the speed of new drug development, there may be diminishing opportunities to obtain extensive data from the quality control laboratory prior to the establishment of specifications. Similarly, a minimal number of lots may be available from the long-term manufacturing site.

The goal of this work is to illustrate the application of statistically-based experimental design and variance component analysis to efficiently assess the analytical and manufacturing process variabilities. Examples of new drug substances will be used to demonstrate the retrospective analysis of a small sample set and the prospective analysis of a larger sample set. The analytical property of potency was selected for the investigation based on the large absolute magnitude of the potency variability in comparison to the variability for other analytical properties (e.g. process related impurities). As the drug substances evaluated are stable compounds, variability associated with stability will not be discussed. Statistical analysis of stability results has been described [8,9].

2. Experimental

2.1. Chromatographic systems

Different HPLC systems were used at the three laboratories involved in the studies. The specifics are provided below. In all cases, an in-house custom-designed system was used as the data acquisition system. The chromatographic conditions for new drug substances A and B (NDS A and NDS B) are given in Table 1. The retention times for NDS A and NDS B are approximately 6.5 and 3.3 min, respectively.

Lab A: A Waters 600E multisolvent delivery system (Bedford, MA, USA) equipped with a column oven, an Alcott 728 autoinjector (Norcross, GA, USA) equipped with a fixed-loop

Valco injection valve (Houston, TX, USA) and an Applied Biosystems 757 or 759A variable wavelength detector (Ramsey, NJ, USA).

Lab B: A Waters 600E multisolvent delivery system equipped with a column oven, a Waters 717 WISP variable-loop autoinjector and a Waters 484 variable wavelength detector, as well as a Waters 616 multisolvent delivery system equipped with a column oven, a Waters 717 Plus WISP variable-loop autoinjector and a Waters 486 variable wavelength detector.

Lab C: A Hitachi (Tokyo, Japan) L-6200A solvent delivery system equipped with a column oven, a Hitachi AS-2000 variable-loop autoinjector and a Hitachi L-4200 variable wavelength detector, as well as a Thermo Separation Products (Spectra-Physics; San Jose, CA, USA) SP8800 solvent delivery system equipped with a column oven, an Alcott 728 autoinjector equipped with a fixed-loop Valco injection valve and an Applied Biosystems 757 variable wavelength detector.

2.2. Reagents and materials

Acetonitrile purchased from EM Science (Gibbstown, NJ, USA) was HPLC grade. Water for the mobile phase was purified with a Milli-Q system from Millipore (Milford, MA, USA). Phosphoric acid (85% wt/wt), potassium phosphate monobasic and sodium phosphate monobasic from EM Science, were of analytical reagent grade. Sodium dodecyl sulfate (SDS), of greater than 98% purity, was obtained from Kodak. Drug substances (NDS A and NDS B) used in the studies were from Eli Lilly and Company (Indianapolis, IN, USA).

2.3. Assay conditions

The standards for NDS A were prepared at 0.05, 0.10 and 0.15 mg ml^{-1} and the samples were prepared as duplicates at 0.10 mg ml^{-1}. The standards and samples for NDS B were prepared as duplicates at 0.05 mg ml^{-1}. The appropriate mobile phase was used as the diluent for solution preparation. For NDS A, the potency result for a sample was determined versus the standard calibration curve. For NDS B, the potency result for a sample was determined versus the average of the standards.

2.4. Experimental design (for NDS B potency)

Three laboratories participated in the determination of method reproducibility for the potency of NDS B. These laboratories are coded in this work as Lab A, B and C. Different experimental designs were employed for each laboratory. The experimental design for each laboratory and the rationale for the selection of the parameter are described in detail. A common control sample of NDS B was used in the studies.

2.4.1. Lab A

The precision of the method was initially evaluated in Lab A, an analytical development laboratory, by completing a full factorial design for three factors at two levels. The factors evaluated were analyst, instrument and column. The analysts selected were ones that normally performed the method. The instruments and columns were those typically used to provide analytical support. Two of the experiments were repeated and served as replicates. A total of ten experiments were performed on eight days during a 7-week period. The design is illustrated in Fig. 1.

2.4.2. Lab B and C

In the quality control laboratories, Lab B and C, the number of analysts, instruments and columns evaluated was expanded and experimental designs involving both nested and cross classified factors were used. The selection of the specifics of each design was guided by the objective to model the long-term performance of the method in the quality control laboratories. Additional details are provided below.

Day: The Lab B design was performed on one day. Potential day-to-day changes were evaluated by running the design at Lab C on two different days. Approximately one month separated the two days selected.

Analyst: Prior experience suggests the analyst may be a critical factor in the performance of a method. Hence, emphasis was placed on including as many different analysts as possible in the de-

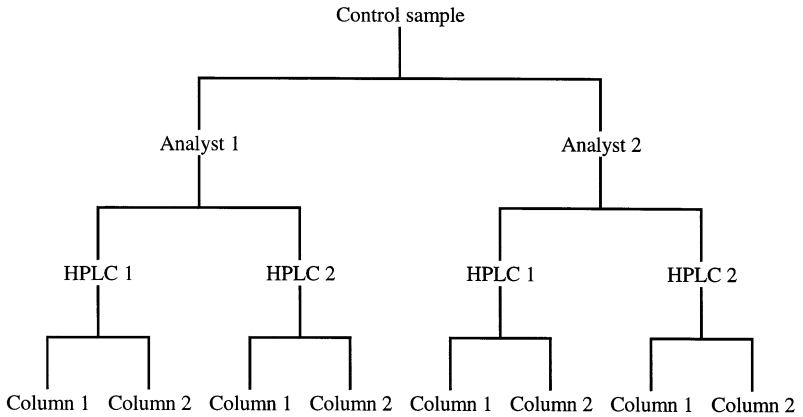

Fig. 1. Experimental design for Lab A.

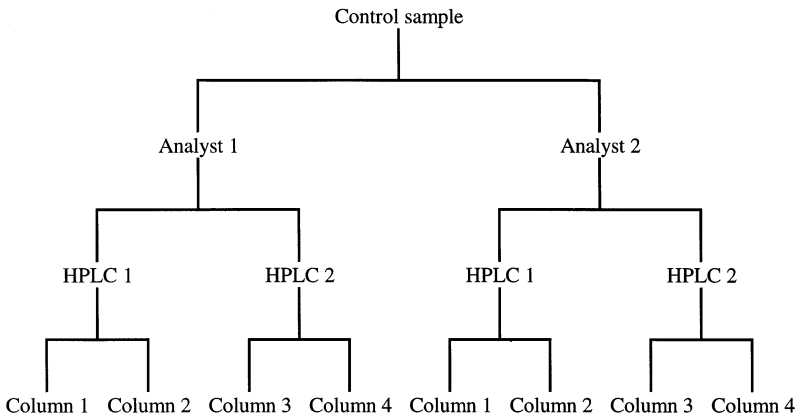

Fig. 2. Experimental design for Lab B.

sign. The analysts were selected to reflect the range of general HPLC experience. Specific experience with the method was not deemed to be as critical as general HPLC expertise. The number of analysts selected for the design is proportional to the number of personnel in the laboratory. A total of two analysts, one analyst each from Lab A and B, participated in the work carried out in Lab B. A total of four analysts, two analysts each from Lab A and C, participated in the work conducted in Lab C.

Instrument (HPLC): HPLC systems were selected to represent the instruments that could be used to support the potency work in the future. Only one HPLC manufacturer exists in Lab B and two typical instruments were selected. For Lab C, two different manufacturers were selected.

Column: Multiple columns will be used throughout the life of the method. The best simulation of this parameter was to obtain as many different batches of columns as possible from the supplier. When the experiments were designed and performed, only three batches of Zorbax SB-C8, 3.5 μm particle size, silica had been manufactured by the supplier. Only two of the three batches were readily available. Thus, two columns from each batch were randomly selected to give four columns for Lab B and C. Both new and aged columns were used in the studies.

The experimental design employed at Lab B is presented in Fig. 2. Eight sets of replicates of the NDS B control sample were obtained from Lab B. The experimental design used at Lab C is shown in Fig. 3. A total of 16 sets of replicates of

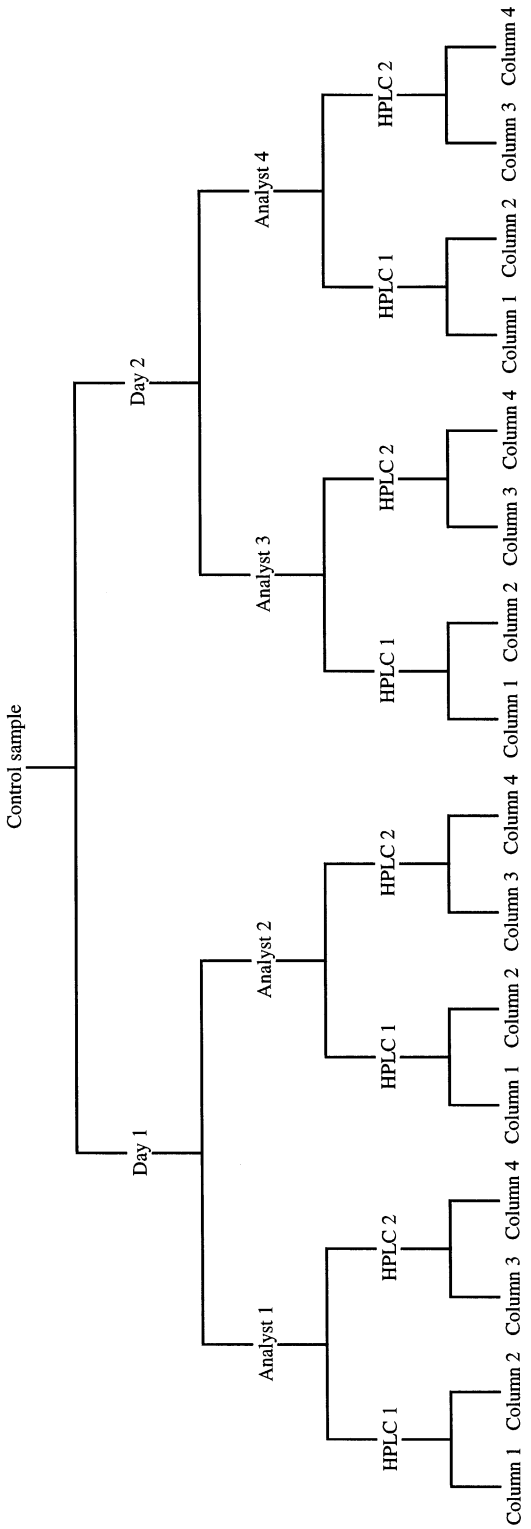

Fig. 3. Experimental design for Lab C.

Table 2
Statistical summary of within-run precision for NDS A

Data set	Type of data	Lab	df	Standard deviation (%)
1	Control sample	A	16	0.643
2	Repeatability	A	12	0.350
3	Replicates from manufactured lots	A	5	0.305
4	Replicates from manufactured lots	B	4	0.494
Pooled standard deviation				0.507

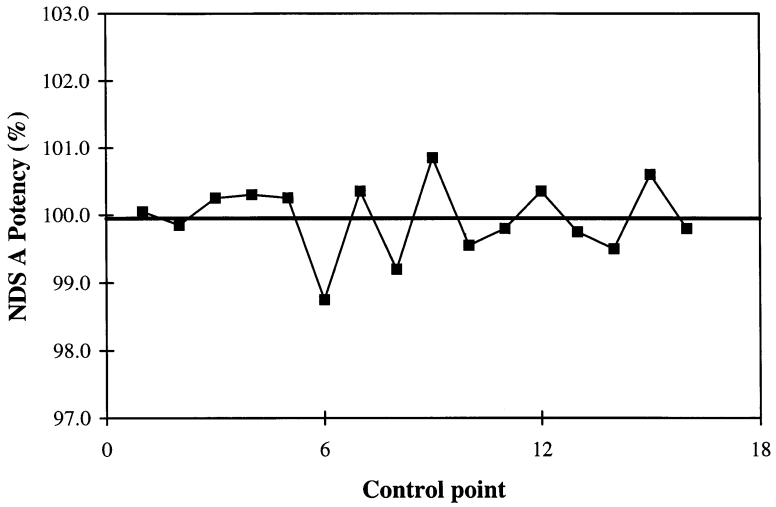

Fig. 4. Plot of potency results (mean, 100.0%) for NDS A control sample.

the NDS B control sample were generated in Lab C. Hence, a total of 24 pairs of potency results for the NDS B control sample were generated in the quality control laboratories that will implement this method in support of manufacturing. The combined data of Lab B and C were obtained on three different days by six analysts on four instruments and six columns. Each pair of replicate results represents a new preparation of the standard and sample solutions. Previous Plackett–Burman and response surface analyses have shown the mobile phase composition, buffer pH and buffer concentration do not have significant effect on the method. Hence, a new mobile phase was not prepared for each part of the experimental designs.

2.5. Statistical analysis

The data were evaluated using the statistical analysis software JMP® (version 3 for the Macintosh, SAS Institute, Cary, NC). Additional details in the derivation of the experimental designs and statistical analysis used in this work can be found in general statistics references [10,11] and the JMP user guide [12]. Some of the general statistical terminology used in this report are briefly discussed below with respect to the experimental design implemented in Lab C (Fig. 3).

Since the two columns used for HPLC 1 are not the same ones used for HPLC 2, the factor of Column is considered to be nested within the factor of HPLC. The nested term has a notation of 'Column [HPLC]'. By the same logic, the two

analysts that conducted the experiments on Day 1 are different from those on Day 2, the factor of Analyst is nested within the factor of Day and has a notation of 'Analyst [Day]'. As the factors of Day and HPLC are main effects and not nested, they can have interaction terms such as 'HPLC * Day' and 'HPLC * Analyst [Day]'. The '*' indicates that the two factors are crossed. These and similar notations are used in the statistical summaries.

3. Results

3.1. NDS A

The variance associated with the analytical property of potency may be expressed in terms of the sum of the process and analytical variances:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{process}} + \sigma^2_{\text{analytical}} \quad (1)$$

Similarly, the analytical variance can be broken into the long-term (run-to-run) and short-term (within-run) contributions, where n is the number of within-run replicates in the potency determination:

$$\sigma^2_{\text{analytical}} = \sigma^2_{\text{long-term}} + \sigma^2_{\text{short-term}} \quad (2)$$

$$\sigma^2_{\text{analytical}} = \sigma^2_{\text{run-to-run}} + \frac{\sigma^2_{\text{within-run}}}{n} \quad (3)$$

In ICH terminology, the short-term component is the repeatability and the long-term component is the combination of intermediate precision and reproducibility [6].

The short-term analytical variance (0.258) for NDS A potency was estimated based on four different data sets: (a) a control sample analyzed 16 times in duplicate by Lab A; (b) repeatability study performed by Lab A; (c) replicates from manufactured lot data analyzed by Lab A; and (d) replicates from manufactured lot data analyzed by Lab B. The statistical summary is given in Table 2.

A control sample of NDS A was analyzed over a 9-month period by two analysts on three instruments. The average of each pair of the 16 sets of results for the control sample (referred to as control data) was calculated and graphically presented in Fig. 4. Variance component analysis, summarized in Table 3, was performed to provide an estimate on the long-term analytical variance (0.082). Therefore, for duplicate analysis of each sample, the total analytical variance estimate is 0.211.

For NDS A, the historical potency results for the manufactured lots were used to estimate the total variance, which is a combination of process and analytical variances. The process variance for NDS A was then obtained by taking the difference between the total and analytical variances (Eq. (1)). Nine lots of manufactured NDS A were used to estimate the process variability and the manufactured lot results are shown in Fig. 5. In this figure, Sites A and B refer to the manufacturing locations of the new drug substance.

The five lots manufactured at Site A have a combined process and long-term analytical variance estimate of 0.250. Based on a long-term analytical variance of 0.082, the process variance for Site A was estimated to be 0.169. The four lots manufactured at Site B were analyzed in one setup and have a total variance estimate of 0.110, which was a composite of the process and short-term analytical variances. Based on the short-term analytical variance estimate of 0.129 (for duplicate analysis of sample), the process variance was estimated to be nearly zero. When the process variance estimates from Sites A and B are pooled, the process variance estimate is 0.094. The NDS A from the two sites was found to be not statistically different ($P > 0.05$).

3.2. NDS B

Statistical analysis of the full factorial design data generated from Lab A is summarized in Table 4. The data from the full factorial design were pooled with additional control data (14 sets of duplicates) generated from Lab A to provide estimates of the short- and long-term analytical variances. The variance component analysis on Lab A data was performed on the pooled data set and is summarized in Table 5. As the additional control data were not obtained from statistically designed experiments, specific factors (e.g. analyst, HPLC) could not be analyzed.

Variance component analyses of the data generated from the experimental designs of Lab B and

Table 3
Summary of variance component analysis for NDS A control data obtained in Lab A

Summary of fit	
Coefficient of determination (r^2)	0.567
Root mean square error	0.643
Mean	99.95
Observations	32

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	15	8.65	0.58	1.40	0.26
Error	16	6.61	0.41		
Total	31	15.26			

Variance component estimates

Component	Variance
Run-to-run	0.082
Residual (within-run)	0.413

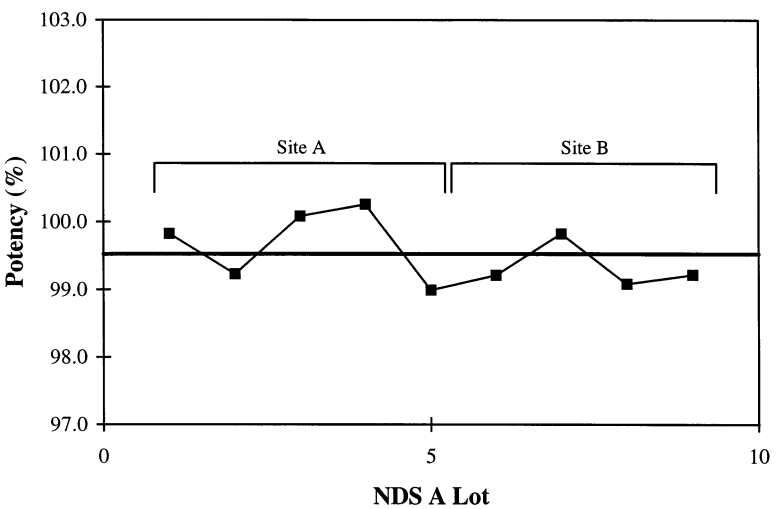

Fig. 5. Plot of potency results (mean, 99.5%) for nine lots of NDS A.

C are summarized in Tables 6 and 7, respectively. The statistical analyses (Tables 4, 6 and 7) of the experimental designs performed in Lab A, B and C indicate the factors evaluated in each laboratory are not statistically significant ($P > 0.05$). The pooled data for Lab A, B and C were also analyzed and the statistical summary is given in Table 8. The analysis indicates that the laboratory is not a significant factor ($P > 0.05$). A summary of the individual laboratory and pooled variances is provided in Table 9.

A total of 48 pairs of control sample data were generated in three laboratories by seven analysts on seven different HPLCs and seven columns over a 6-month period. The average for each pair of the 48 sets of data obtained from the three laboratories was calculated and plotted in Fig. 6.

Table 4
Summary of statistical analysis for full factorial of NDS B potency performed in Lab A

Summary of fit	
Coefficient of determination (r^2)	0.707
Root mean square error	0.861
Mean	99.54
Observations	10

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	7	3.58	0.51	0.69	0.703
Error	2	1.48	0.74		
Total	9	5.06			

Effect test

Source	$P>F$
Analyst	0.863
HPLC	0.330
Column	0.749
Analyst * HPLC	0.750
Analyst * Column	0.525
HPLC * Column	0.477
Analyst * HPLC * Column	0.335

Table 5
Summary of variance component analysis for NDS B control data obtained in Lab A

Summary of fit	
Coefficient of determination (r^2)	0.730
Root mean square error	0.502
Mean	99.49
Observations	48

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	23	16.32	0.71	2.82	0.007
Error	24	6.04	0.25		
Total	47	22.36			

Variance component estimates

Component	Variance
Run-to-run	0.229
Residual (within-run)	0.252

A total of 21 lots of manufactured NDS B were used to provide an estimate on the process variance. All 21 lots, 11 lots from Site A and ten lots from Site B, were analyzed in duplicate on one

Table 6
Summary of variance component analysis for the experimental design of NDS B potency performed in Lab B

Summary of fit	
Coefficient of determination (r^2)	0.624
Root mean square error	0.824
Mean	99.58
Observations	16

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	7	9.06	1.29	1.89	0.195
Error	8	5.47	0.68		
Total	15	14.53			

Tests with respect to random effects

Source	$P>F$
Analyst	0.491
HPLC	0.613
Analyst * HPLC	0.402
Column [HPLC]	0.480
Analyst * Column [HPLC]	0.225

Variance component estimates

Component	Variance
Run-to-run	0.306
Residual (within-run)	0.684

potency run at Lab A. The NDS B from the two sites was found to be not statistically different ($P > 0.05$). A plot of the lot averages is shown in Fig. 7. The variance component analysis is summarized in Table 10.

4. Discussion

4.1. NDS A

Based on existing data readily available from the analytical method validation and manufactured lot data, variance component analysis was retrospectively applied to determine the short- and long-term contributions to the analytical variability (Tables 2 and 3) as well as the process variability of NDS A. The short-term variability (standard deviation, 0.36%; duplicate analysis of sample) for the NDS A potency is slightly higher than the long-term variability (standard deviation, 0.29%). One generally expects the long-term variability to be greater than the short-term variability. This suggests that the sample preparation for NDS A has a greater impact on the method precision than other long-term factors such as analyst, instrument and column. Thus, a laboratory analyst can exercise greater care in the standard and sample preparation for this method to improve the overall method precision.

A common approach to reduce analytical variability is to increase the number of replicates performed to reduce the short-term component of the analytical variability. Given that the overall analytical variation is inversely related to the number of replicates (Eq. (3)) and the overall analytical variation is relatively small, increasing

Table 7
Summary of variance component analysis for the experimental design of NDS B potency performed in Lab C

Summary of fit	
Coefficient of determination (r^2)	0.748
Root mean square error	0.507
Mean	99.99
Observations	32

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	15	12.18	0.81	3.16	0.014
Error	16	4.11	0.26		
Total	31	16.29			

Tests with respect to random effects

Source	$P>F$
Day	0.585
Analyst [Day]	0.468
HPLC	0.874
HPLC * Day	0.392
HPLC * Analyst [Day]	0.525
Column [HPLC]	0.801
Day * Column [HPLC]	0.441
Analyst [Day] * Column [HPLC]	0.054

Variance component estimates

Component	Variance
Run-to-run	0.278
Residual (within-run)	0.257

the number of replicates in this case does not yield any significant improvement in the method precision. For example, doubling the number of replicates from two to four only results in decreasing the total analytical variability from 0.46 to 0.38%.

Pooling the process and analytical variances (0.094 and 0.211, respectively) gives a total standard deviation estimate of 0.55%. Thus, this statistical analysis suggests that specification limits of 97.5 to 101.0% would be consistent with the process mean (99.5%) $\pm 3\sigma$. As the NDS A sample set contains a small number of data points, uncertainty in the variance estimates could be considered by using confidence interval, tolerance limit, or other statistical methods to account for the degrees of freedom used to derive the estimate.

4.2. NDS B

The pooled analytical data from all three laboratories provide a total analytical standard deviation estimate of 0.67% for the NDS B potency method (when each sample is analyzed in duplicate). Although this multiple-laboratory study is not as extensive as a formal Association of Official Analytical Chemists (AOAC) collaborative study, the data generated in this study give a useful assessment of the method reproducibility.

The statistical analyses also revealed potential sources of variance contributing to the analytical variability. While the long-term standard devia-

Table 8
Summary of variance component analysis for combined Lab A, B and C data

Summary of fit	
Coefficient of determination (r^2)	0.731
Root mean square error	0.570
Mean	99.67
Observations	96

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	47	42.49	0.90	2.78	0.0003
Error	48	15.62	0.33		
Total	95	58.11			

Tests with respect to random effects

Source	$P>F$
Lab	0.063
Run [Lab]	0.001

Variance component estimates

Component	Variance
Lab	0.056
Run [Lab]	0.255
Residual	0.325

Table 9
Summary of Lab A, B, and C standard deviations and the pooled values for NDS B potency

Standard deviation (%)	Lab A	Lab B	Lab C	Pooled
Run-to-run	0.478	0.553	0.527	0.538
Within-run (duplicate)	0.355	0.585	0.358	0.403
Total (run-to-run and within-run)	0.596	0.805	0.637	0.672

tion for both quality control laboratories are comparable (Lab B, 0.55%; Lab C, 0.53%), the short-term standard deviation at Lab B (0.59%) is larger than that for Lab C (0.36%). This difference may be attributed to the variable-loop injector used on the HPLCs in Lab B. The Lab C study included both fixed-loop and variable-loop injectors. The variable-loop injector in Lab C has been demonstrated to have comparable within-run precision to the fixed-loop injector and better within-run precision than the variable-loop injector used at Lab B.

Due to a higher NDS B analytical variance relative to the total variance for the historical potency results for manufactured lots of NDS B, an approach different from that taken for NDS A was applied to estimate NDS B process variance. The low total variance is attributed to the analyses performed by experienced analysts on a limited number of instruments and columns in Lab A and the confounding of factors due to the analysis of multiple lots within a determination. To obtain an estimate on the process variance, 21 lots of NDS B were analyzed in a single potency determi-

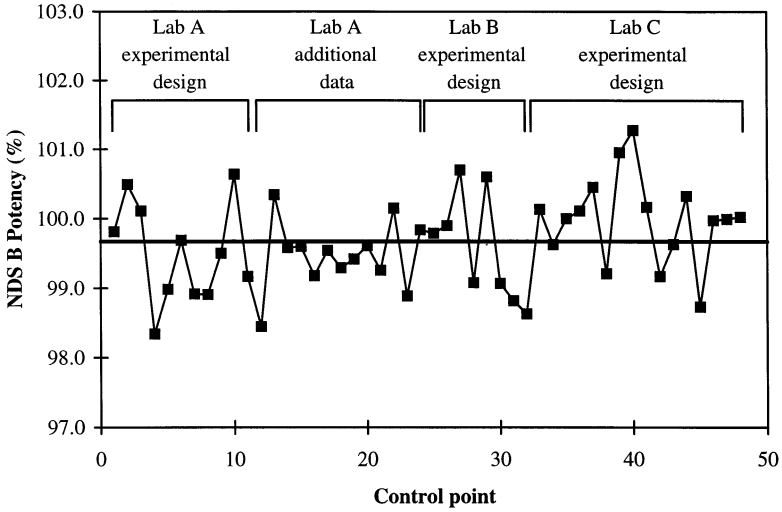

Fig. 6. Plot of potency results (mean, 99.7%) for NDS B control sample.

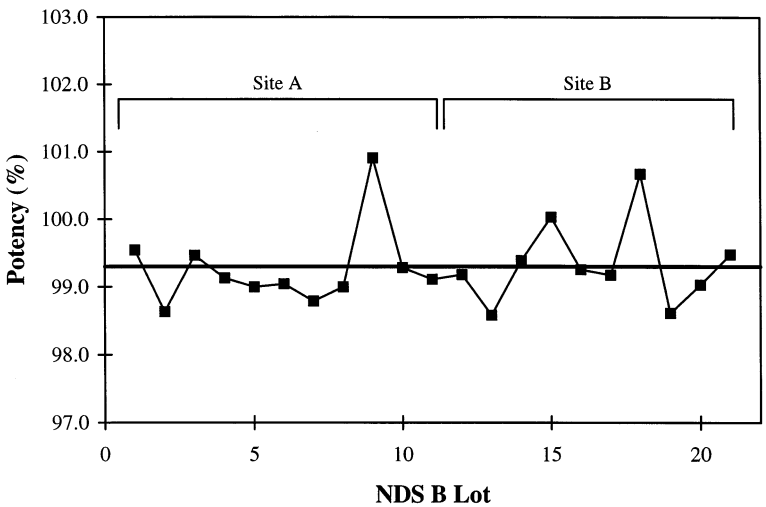

Fig. 7. Plot of potency results (mean, 99.3%) for 21 lots of NDS B.

nation. After adjusting for the within-run variance, a process variance estimate of 0.22 was obtained (Table 10). The analytical variance and process variance were then pooled to provide an estimate of the overall variance and standard deviation of 0.81% for the analytical property of potency for NDS B. Specification limits of 97.0 to 102.0%, calculated from the process mean $(99.3\%) \pm 3\sigma$, are consistent with the analytical and process capability.

4.3. Analytical and process capability

In addition to the two examples above, the process and analytical variability for two additional synthetic drug substances, NDS C and NDS D, were derived by retrospective analyses analogous to that applied for NDS A. The process and analytical (short- and long-term) variability, expressed as standard deviations, for all four drug substances are summarized in Table 11.

Table 10
Summary of variance component analysis for potency results of NDS B lots

Summary of fit	
Coefficient of determination (r^2)	0.705
Root mean square error	0.537
Mean	99.30
Observations	42

Analysis of variance

Source	df	Sum of squares	Mean square	F ratio	$P>F$
Model	20	14.43	0.72	2.51	0.021
Error	21	6.05	0.29		
Total	41	20.48			

Variance component estimates

Component	Variance
Lot	0.217
Residual	0.288

Table 11
Summary of variability, expressed as standard deviation (%), for the analytical property of potency in the determination of drug substances

				Analytical	
NDS	Total	Process	Analytical	Long-term	Short-term
A	0.55	0.31	0.46	0.29	0.36
B	0.81	0.45	0.67	0.54	0.40
C	0.57	0.42	0.38	0.31	0.22
D	0.65	0.44	0.48	0.37	0.30

Despite the varying complexity of the four different synthetic drug substance processes, the process variability is relatively constant at 0.4% and is comparable to the analytical variability. Review of various chromatographic analyses of formulated drug products has shown that methods typically have repeatability and reproducibility of approximately 1 and 2% RSD, respectively [3]. The lower repeatability and reproducibility values for the four drug substances reported in this work are likely to be due to a combination of simpler matrix of drug substance relative to drug product, improved HPLC instrumentation and small sample sets relative to the AOAC-type collaborative studies in the review. Qualitatively, the repeatability and reproducibility are comparable, suggesting that the laboratories are performing the analyses in a consistent manner [3].

The compendial potency specification for a synthetic drug substance has been suggested to be not less than 98.0% and not more than 102.0% [13], a specification range of 4%. For a proposed potency specification of process mean $\pm 3\sigma$, a specification range of 4% translates to a total variability, expressed as standard deviation, of 0.67%. Based on the total variability estimates in Table 11, the suggested total variability threshold of 0.67% would be inappropriate for NDS B.

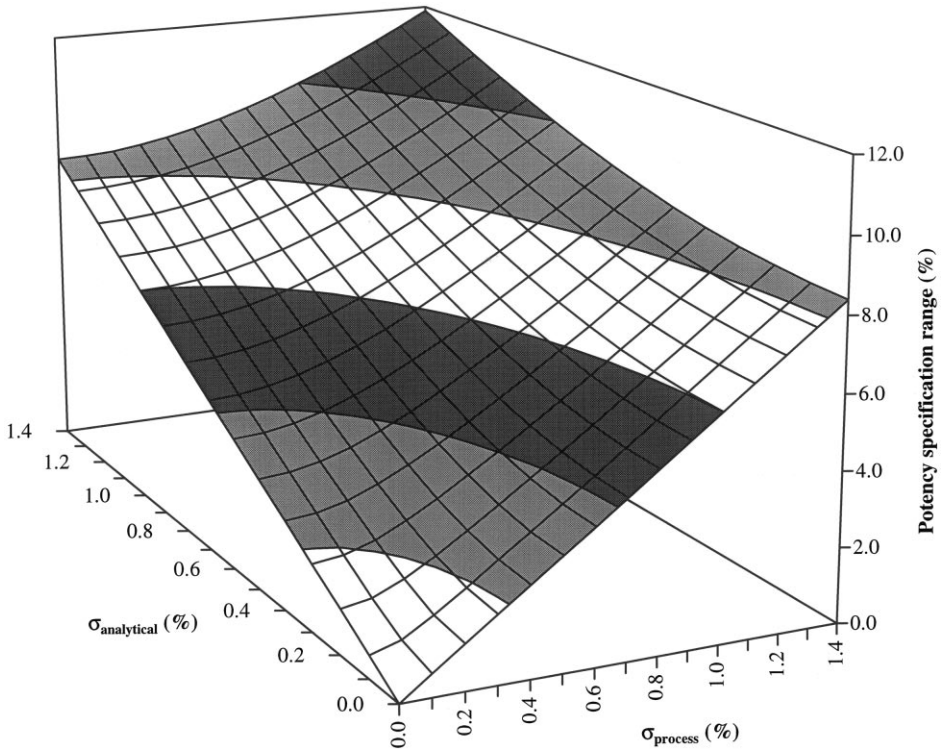

Fig. 8. Plot of potency specification range as a function of analytical and process variability.

In accordance with the ICH effort on specifications, specifications should be established based on the manufacturing process and analytical capability [14]. The prospective approach for assessment of analytical variability by modeling the operation of the eventual quality control laboratory and obtaining a large sample set, can provide a reasonable estimate of the analytical variability. Extending this approach to evaluate process variability by generating a large sample set in production-scale equipment prior to regulatory submission and approval can be unreasonable and costly. A more likely scenario is that data from a limited number of lots in production-scale equipment would be available.

An acceptable estimate of the process variability can still be derived by considering the analytical variability, which can be experimentally ascertained, and Eq. (1). If the potency specification is established based on using a process mean $\pm 3\sigma$ approach, then:

$$\text{Potency specification range} = 6 \times \sigma_{\text{total}} \quad (4)$$

Eq. (1) and Eq. (4) can be combined to give:

$$\text{Potency specification range} = 6 \times \sqrt{\sigma^2_{\text{process}} + \sigma^2_{\text{analytical}}} \quad (5)$$

Whether one should use process mean $\pm 2\sigma$ or process mean $\pm 3\sigma$ is debatable. However, if the potency specification is established based on process mean $\pm 2\sigma$, approximately 5% of lots that are within statistical process control and of acceptable quality would not meet the narrower specification. Graphical representation of Eq. (5) gives a simple illustration of the relationship between the potency specification range and its process and analytical variability components (Fig. 8). The four drug substances discussed in this report have an analytical variability range of 0.4–0.7%. If this range is representative for well-behaved HPLC methods for the determination of potency of drug substances, then for an analytical

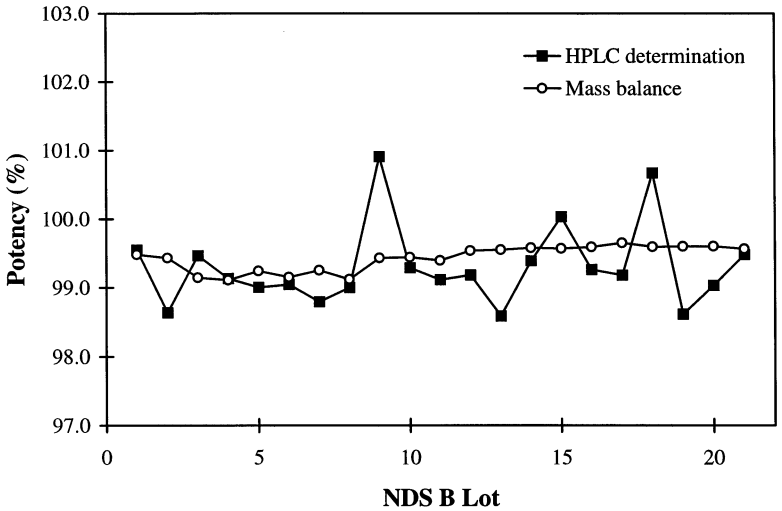

Fig. 9. Plots of NDS B potency by HPLC (mean, 99.3%) and by mass balance (mean, 99.4%).

variability of 0.6%, a potency specification range of 4% (e.g. 98.0–102.0%) would allow the process to have a maximum variability of 0.3%. Based on the observed process variability (Table 11), the proposed specification range of 4% cannot be met. However, allowing the potency specification range to be 5% (e.g. 97.5–102.5%), the process variability may then range from 0.4 to 0.7%, which is a more realistic process range relative to that of the analytical variability. A drug substance specification range of 5% is still smaller than the generally accepted drug product potency specifications of 95–105% or wider [4,5].

Another consideration in establishing a potency specification is that the potency specification is part of a total control strategy and not a lone indicator of purity or quality. A recent study has suggested that a pure drug substance will routinely give potency results within specification, but the exact value will not be indicative of the quality [15]. Support for such a proposition can be found when the HPLC potency results for 21 lots of NDS B, analyzed in duplicate on one potency run (Fig. 7), are compared to the purity of the same lots by subtracting impurity levels from 100% (i.e. mass balance). The comparison is shown in Fig. 9 and indicates that the potency by mass balance (standard deviation, 0.18%) has less variability than potency by HPLC determination (standard deviation, 0.60%). The lower variability of potency determination by mass balance can be attributed to the lower variability of the impurity methods and is consistent with the empirical relationship derived by Horwitz for the concentration of analyte and the method precision [16]. The relationship can be represented as:

$$\%\mathrm{RSD} = 2^{(1 - 0.5 \log C)} \quad (6)$$

where C is the concentration of the analyte expressed in powers of ten. Based on Eq. (6), a method for the determination of an impurity at 0.5% should have a %RSD of 4.4 and a standard deviation of 0.022%.

5. Conclusions

Retrospective application of variance component analysis on potency data for NDS A reduced the uncertainty associated with establishing the potency specification for NDS A. The unique combination of variance component analysis and experimental design was used to estimate analytical variability for NDS B. The described prospective approach provides a predictive tool and offers even greater confidence in the establishment of

specifications. Given the observed analytical variability for four different drug substances, the suggested compendial specification range of 4% for potency determination by HPLC does not allow for reasonable variation in the process. Alternatively, a potency specification range of 5% (e.g. 97.5–102.5%) is regarded as being more consistent with the data presented. Furthermore, a potency result alone is not an adequate indicator of drug substance purity. The quality of a drug substance can be better assessed when potency data are evaluated in combination with impurity results and other analytical data.

Acknowledgements

The authors are grateful for the critical contribution by the quality control laboratory component of Eli Lilly and Company to the studies conducted in this work. The authors also acknowledge and thank Dr Bernard Olsen and Dr John Murphy for technical discussion and comments.

References

[1] J. Gardner, S. Coleman, S. Farrow, Anal. Proc. 30 (1993) 183–185.
[2] J. Hansen, W. Zirk, L. D'Esposito, Pharm. Forum 22 (1996) 2931–2933.
[3] W. Horwitz, J. Assoc. Off. Anal. Chem. 60 (1977) 1355–1363.
[4] T. Layloff, P. Motise, Pharm. Technol. 16 (9) (1992) 122–132.
[5] G. Davis, J. Murphy, D. Weisman, S. Andersen, J. Hofer, Pharm. Technol. 20 (9) (1996) 110–118.
[6] International Conference on Harmonization, Guideline on Validation of Analytical Procedures: Definitions and Terminology, Federal Register 60 (1995) 11260–11262.
[7] International Conference on Harmonization, Draft Guideline on the Validation of Analytical Procedures: Methodology, Federal Register 61 (1996) 9316–9319.
[8] International Conference on Harmonization, Stability Testing of New Drug Substances and Products; Guideline, Federal Register 59 (1994) 48754–48759.
[9] J. Grimes, L. Foust, Proc. Biopharm. Sect. Am. Stat. Assoc. 1994, 498–502.
[10] R. Anderson, Practical Statistics for Analytical Chemists, Van Nostrand Reinhold, New York, 1987.
[11] J.C. Miller, J.N. Miller, Statistics for Analytical Chemistry, Ellis Horwood, Chichester, UK, 1988.
[12] JMP Statistics and Graphics Guide, Version 3, SAS Institute, 1994.
[13] W. Paul, Pharm. Technol. 15 (3) (1991) 130–141.
[14] Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances, Draft 5.0, Step 2 of the ICH process, April 1, 1997.
[15] R. Bunnell, Pharm. Res. 14 (1997) 156–163.
[16] W. Horwitz, Anal. Chem. 54 (1982) 67A–76A.

.