Journal of Food and Drug Analysis
1994. 2(3) : 161-176

藥物食品分析
第二卷　第三期

Validation of Analytical Methods : A Simple Model for HPLC Assay Methods

HSING-CHU HSU AND *CHUN-SHENG CHIEN

Department of Pharmaceutics, Chia Nan Junior College of Pharmacy, Tainan-hsien, Taiwan, R.O.C.
**National Laboratories of Foods and Drugs, Department of Health, Executive Yuan, R.O.C.*

ABSTRACT

A simple model for the validation of HPLC analytical methods, to be carried out in a typical laboratory within the constraints of time and cost, is described by reviewing currently available validation methods. The statistical techniques employed in assessing the accuracy, precision, selectivity, reproducibility, sensitivity and validity of a non-official assay method in comparison with the compendial method are reviewed and demonstrated.

Key words : Validation, HPLC, analytical model.

INTRODUCTION

A variety of reasons demand validation of analytical methods. However, two especially important reasons should be stressed. The first is that an analytical method should be scientifically rational and reliable. Since the method selected plays an integral role in obtaining the observed data, it must be accurate and free from bias. The objective of method validation is to produce the best analytical results possible. The second reason is that current good manufacturing practice regulations (CGMPs) require assay validation.

Regarding this regulatory issue, there was no explicit mention of validation in the 1971 CGMPs (U.S.A.). However, paragraph 133.11 states that laboratory controls should include the establishment of scientifically sound test procedures, and paragraph 133.11(f) mentions accuracy and precision in relation to laboratory test procedures[1]. The word "validation" is explicitly used in the 1979 CGMPs paragraph 211. 165(e); where regulations call for proof that each method used in testing meets proper standards of accuracy and reliability, and that the suitability of all testing methods should be verified under actual conditions of use. Regulations also require all analytical methods used for stability testing to be validated to demonstrate any indications of instability[2]. These guidelines are quite similar in Japan's GMPs.

Youden[3,4] and Mandel[5] have been acknowledged as the pioneering contributors to the literature of the design and analysis of assay validation experiments. Their investigations showed the effect on test results of variables such as analysts, time, equipment, reagents, laboratories etc. These papers gave examples of balanced

Correspondence to : Hsing-Chu Hsu　　*Accepted for publication : Nov. 4, 1993*

sampling designs and demonstrated the analysis of data, utilizing statistical methods. Two references targeted primarily for the pharmaceutical industry in validating the analytical methods were the Greenbrier Procedure[6] and the Rosemont Procedure[7]. The procedures were named for papers presented respectively at Greenbrier, North Carolina in 1973 and at Rosemont, Pennsylvania in 1982. The Greenbrier and Rosemont Procedures focused on validation of alternative assays in multiple system and assay precision involving interlaboratories. The methods depend largely on the asymptotic variance and at best are a large-sample approximation. However, the methods can be adopted for use in any assay validation program in a quality control laboratory.

In Europe, a validation guideline called "GVP : Good Validation Practice" was presented by FIP in 1980[8]. GVP did not focus on the validation of analytical methods specifically; however, it stated that during the development phase, the analytical procedures for chemical, galenical and biological tests must be validated, i.e. the methods had to be checked for correctness, accuracy and reproducibility and thus, for their reliability. Validation for the chromatography assay methods was discussed in a series of technical reports on validation in Japan in 1983[9]. The compendial issues in the experimental error and tests of assay validity were discussed in the biological assays section of the USP XX[10].

Since 1982 U.S. FDA officers have described and explained the requirements of assay method validation, documentation and laboratory evaluation of proposed NDA nad ANDA analytical methodology in FDA in a series of reports[11-16]. In 1987, FDA disclosed the guidelines for submitting samples and analytical data for methods validation[17]. A guidance note on analytical validation was also published by the Committee for Proprietary Medicinal Products, U.S.A. in 1988[18]. Validation of compendial assay methods is for the first time described in the general information section of the USP XXII (1990)[19]. Typical analytical performance parameters for assay validation such as precision, accuracy, limit of detection, limit of quantitation, selectivity, range, linearity, ruggedness are defined; the procedures to determine the system suitability are also briefly described in the USP

Table 1. Data elements required for assay validation

Analytical Performance Parameter	Assay Category I	Assay Category II		Assay Category III
		Quantitative	Limit Tests	
Precision	Yes	Yes	No	Yes
Accuracy	Yes	Yes	*	*
Limit of Detection	No	No	Yes	*
Limit of Quantitation	No	Yes	No	*
Selectivity	Yes	Yes	Yes	*
Range	Yes	Yes	*	*
Linearity	Yes	Yes	No	*
Ruggedness	Yes	Yes	Yes	Yes

* May be required, depending on the nature of specific test.

XXII. Analytical methods are divided into three categories : Category I is the method used in the quantitation of drug in bulk form and in finished products; Category II refers to the stability -indicating assay method; Category III is the method for determination of performance characteristics. The data elements required for assay validation for each category are specified, (Table 1).

These documents and guidelines, as it was pointed out by Carr and Wahlich[20], are often vague and rarely provide the detailed procedures required in the validation exercise because the analytical methods used in drug analysis are so diverse that general procedures suitable for validating all kinds of analytical methods are for pharmaceuticals practically impossible.

In general, procedures for analytical methods validation can be divided into two parts: single-system validation and dual- or multiple-system validation. Single-system validation is directed primarily to a single assay method. The validation is used to demonstrate that this method is accurate and reliable, but without reference to other methods. Multiple-system validation intends to compare (1) the standard or compendial and the alternative methods; (2) the same method tested in two or more different laboratories; (3) the original condition and the perturbed condition; (4) the same method conducted by two or more different analysts etc.

The validation of an analytical method is generally carried out by extensive statistical evaluation of the analytical methods performance characteristics. Duarte and Vest[21] reported a model for mixed factorial analysis of variance for obtaining detailed information about the precision and accuracy of methods being validated in comparison with the compendial method. Carey et al.[22] and Lawton et al.[23] also described a method using principal component analysis to obtain data concerning accuracy and precision. Koopmans et al.[24] reported the method to estimate the precision using confidence intervals method. Westgard and Hunt[25] demonstrated the use and interpretation of statistical tests in method-comparison studies. There are many descriptions in the literature of sophisticated statistical methods and their application in the evaluation of method accuracy and precision. However, most of these methods require advanced techniques and a large numbers of assay determinations which are usually beyond the capability of a typical laboratory. The main purpose of the present work is to abstract the essence of the current state-of-the-art validation methodology, and to propose a simple model focusing on the HPLC assay method for modeling the collection, analysis and interpretation of the valid data. This simple model of method validation can be performed in a typical laboratory within the constraints of time and cost, and complies with both the CGMPs and the compendial requirements. It should be pointed out that when this proposed simple validation method is applied, it is not possible to obtain all of the relevant information which would be desirable to learn working of an assay. Nevertheless, these procedures may provide a scientific rationale for an assay method.

ASPECTS OF VALIDATION

I. *System Suitability*[19,26]

It is necessary to determine the adequacy and suitability of an HPLC analytical system during method development and, after the method has been validated the established system suitability specifies the adequate conditions for band retention, selectivity, resolution, measured plate number and tailing factor for a proposed HPLC assay method.

(I). *Capacity Factor* (k')

Capacity factor (k') is defined as

$$k' = (t_r - t_o)/t_o$$

where t_o is column dead-time and t_r is the retention time of the analyte.

In general, k' value is preferably between 2 to 8 for the band of analyte. In case of trace

measurement, $1<k'<3$ is preferable, if resolution from possible interference permits. For the stability-indicating assay, $k'>4$ may be adequate. The constant k' can be taken as the indication of column equilibrium, and is the system parameter to be checked first.

(II). *Selectivity* (α)

Selectivity (α) is defined as

$$\alpha = \frac{k'_2}{k'_1}$$

where k'_1 and k'_2 are the capacity factors of analyte 1 and 2, respectively. The adequate α value is between 1.05 to 2.0.

(III). *Resolution* (R_s)

Resolution (R_s) is defined as

$$R_s = \frac{2(t_2 - t_1)}{W_1 + W_2} = \frac{1.18(t_2 - t_1)}{[W_{1(1/2)}] + [W_{2(1/2)}]}$$

where t_1 and t_2 are the retention time; W_1 and W_2 are the bandwidth; $W_1(1/2)$ and $W_2(1/2)$ are the bandwidth at half-height for band 1 and band 2, respectively.

A resolution (R_s) of not less than 1.5 is strongly recommended for quantitative analysis in the baseline resolution of the interest band from the adjacent noise bands and the bands of its related substances (impurities, degradants, metabolites etc). Nevertheless, resolution not less than 1.2 may be tolerable where assaying samples such as biological specimens. Specificity requires the analytical system to provide separation of the analyte from process impurities, degradants, excipients, packaging extractables, and analytical artifacts. The specificity of the method can be delineated by the resolution of the integrity bands. The integity of a band determined by multiple UV wavelength detection (e.g. photodiode array detector or ratio between two wavelengths) may not always be reliable, since many degradants and metabolites may have chromophores similar to the intact drug. Band purity can be assessed with a higher degree of certainty only by band collection, followed by analysis using significantly different chromatographic procedures and using IR, NMR and Mass spectrometry. However, this is often difficult if not impracticable in typical laboratories at present. In the simplest approach, the normal phase preparative TLC is carried out and each band is collected respectively to run a reverse phase HPLC to determine the integrity of a band by the ratio between two wavelengths. This approach is suitable for analysis of the samples with known constituents.

For cases where bands overlap moderately, a convenient measurement of R_s has been proposed by Snyder and Kirkland[27]. This approach is based on the height of the valley between the two bands (h_r) as described in Figure 1 and Table 2. Quantitation with $R_s<1$ is generally less precise, particularly since R_s will vary from day to day and from lab to lab.

(IV). *Measured Plate Number* (N)

Measured plate number (N) is defined as

$$N = 16\left(\frac{t_r}{W}\right)^2 = 5.54\left(\frac{t_r}{W_{1/2}}\right)^2$$

where t_r is the retention time, W and $W_{1/2}$ are the bandwidth and the bandwidth at half-height for the band of the analyte. N value should be measured under the analytical condition for the analyte band and can be used as an

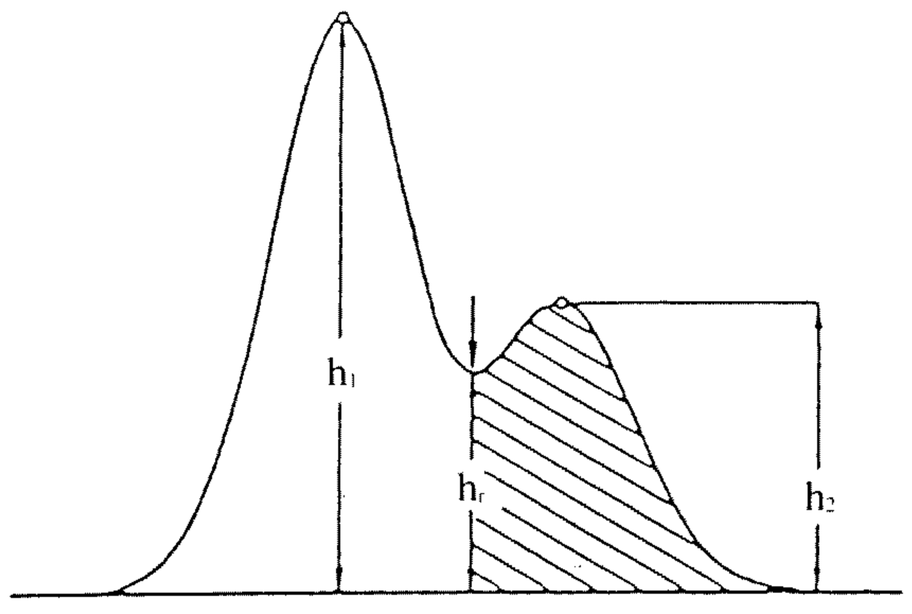

Figure 1. Measurement of height of the valley between two overlapping bands.

indicator of the column efficiency in routine analytical practice. If the N value decreased significantly after a certain time period of analytical practice, decrease of the column efficency should be suspected. When significant tailing of a band exists, then the plate number measured by the above equation will be artificially high. The Dorsey-Foley equation should be used to estimate the plate number : $N=41.7(t_r/W_{0.1})^2/(A_s+1.25)$, where $W_{0.1}$ and A_s are defined in the following section (tailing factor).

(V). *Asymmetry Factor or Tailing Factor* (T or A_s)

Asymmetry Factor, T is defined in the USP XXII as

$$T=W_{0.05}/2f$$

where $W_{0.05}$ and f are defined in Figure 2.

A majority of the T values for the HPLC analysis of drugs complied in the USP XXII are 1.5 to 2.0. The T values of a few drugs are specified as less than 1.5 (e.g. Flunisolide, $T<1.0$), whereas some other drugs are allowed to have relatively large T values (e.g. Famotidine, $T<3.0$; Ciprofloxacine HCl, $T<4.0$). Another measurement of band asymmetry, the tailing factor (A_s) is used in the journals. The A_s is defined as

Table 2. Estimating resolution from the height of the valley between two adjacent bands

h_r/h_2 (%)	R_s for Indicated Ratio of Band Size (h_r/h_2)			
	1/1	2/1	4/1	8/1
10	1.22	1.26	1.30	
20	1.07	1.13	1.17	1.31
30	0.97	1.05	1.10	1.22
40	0.89	0.99	1.05	1.13
50	0.83	0.92	1.01	1.07
60	0.78	0.86	0.96	1.01
70	0.74	0.82	0.91	0.96
80	0.70	0.77	0.86	0.92
90	0.66	0.74	0.82	0.89

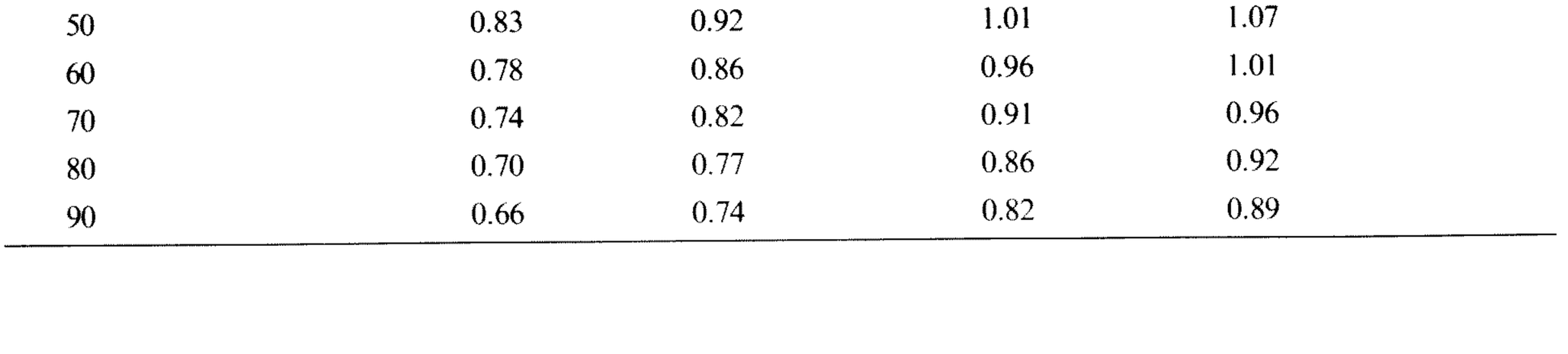

Figure 2. Asymmetrical chromatographic peak

$A_s = BC/AC$ (Figure. 2). A_s is preferably between 0.9 to 1.3, to ensure accuracy and precision of analytical results. A_s will increase under inadequate HPLC conditions and with decreasing column efficiency. If the assay is performed under a value of A_s larger than 1.5, it is very difficult to meet requirements of accuracy and precision.

II. *Linearity*

Linearity is conducted to justify the single point ratio calculation method (SPRC method) of which the analyst is assuming a linear response with zero intercept. In some methods, e.g. fluorescence or atomic absorption spectrometry or the detectors used in the GC, this assumption cannot be true. It should be shown that the calibration curve has a satisfactory slope sensitivity over the range of interest.

To validate linearity between the response and the concentration, a series of standard preparations are used to construct the standard calibration line which serves as evidence that the analytical system is adequate, and that systematic and proportional biases are absent. In addition, tests of linearity using spiked placebo samples must be performed to show that the response is directly proportional to the concentration of the analyte in the presence of excipients, since the objectives of the assay are the samples containing excipient matrix. Table 3. depicts the Design of Experiment and the equations relevant to the linear regression.

The regression model and regression equation are:

$$Y = a + bX + e, \quad e = N(0, \sigma_e^2)$$

$$\text{where } a = \frac{[\Sigma Y\ \Sigma X^2 - \Sigma X\ \Sigma XY]}{[mn\ \Sigma X^2 - (\Sigma X)^2]}$$

$$b = \frac{[mn\ \Sigma XY - \Sigma X\ \Sigma Y]}{[mn\ \Sigma X^2 - (\Sigma X)^2]}$$

It should be pointed out that the simple linear regression using nonweighted-least-squares method is adequate only for the concentration spans around a 100-fold range in absorbance spectrometry (in case of fluorometry, this may be a few 10-fold ranges), in which the homogeneity of variance at each concentration may be assumed. For the linear regression in which the concentration spans a 1000-fold range or more, or the heteroscedasticity of data is suggested, weighted regression should be carried out.

The test of homoscedasticity can be performed to the test of null hypothesis, $\sigma^2 = \sigma_{yx}^2$ (equivalency of variances for a single level and the regression line). The test statistic F_o for the hypothesis is:

$$F_o = \frac{S_{yx}^2}{S^2} \sim F_{(n_1-1),(n_2-1),0.05}$$

$$\text{where } S_{y\cdot x}^2 = \frac{\Sigma Y^2 - \frac{(\Sigma Y^2)}{n_1} - b[\Sigma XY - \frac{\Sigma X \Sigma Y}{n_1}]}{n_1 - 2}$$

$$S^2 = \frac{\Sigma Y^2 - \frac{(\Sigma Y)^2}{n_2}}{n_2 - 1}$$

In case heteroscedasticity is suggested, the weighted regression equation would be:

$$Y = a + bX + e, \quad e = N(0, \sigma_e^2)$$

$$a = \frac{[\Sigma WY^2\ \Sigma WY - \Sigma WX\ \Sigma WXY]}{[\Sigma W\ \Sigma WX^2 - (\Sigma WX)^2]}$$

$$b = \frac{[\Sigma W\ \Sigma WXY - \Sigma WX\ \Sigma WY]}{[\Sigma W\ \Sigma WX^2 - (\Sigma WX)^2]}$$

Table 3. Design of experiment for the test of linearity

Cause (X)	$X_1 \cdots\cdots Xi \cdots\cdots X_m$
Response (Y)	$Y_{11} \cdots\cdots Y_{i1} \cdots\cdots Y_{m1}$
	$\vdots \quad \vdots \quad \vdots$
	$Y_{1n} \cdots\cdots Y_{in} \cdots\cdots Y_{mn}$
	$\Sigma Y_1 \cdots\cdots \Sigma Y_i \cdots\cdots \Sigma Y_m$

$n \geqq 5, m \geqq 6$

where W: weight $= \frac{1}{\sigma^2}$

It is important to note the coefficient of correlation (r) does not test the linearity of two variables. Although it has barely been practiced correctly in the pharmaceutical literatures, the misuse of the coefficient of correlation has been criticized for some time[20,28-32]. The proper test for linearity is the Lack-Of-Fit test[33]; another convenient alternative test is the Quadratic Regression test[34].

(I). *Lack-Of-Fit Test*

When a set of data X and Y are fitted to a regression model, there is always danger of using a regression model which is a poor approximation of the true functional relationship and is thus inadequate, the Lack-Of-Fit test may be used to support the adequacy of the regression model. The hypothesis to be tested is:

H_o:The linear model adequately fits the data.

H_1:The linear model does not fit the data.

The test involves partitioning the residual sum of squares (S_E) into the sum of squares of pure experimental error (S_{PE}), and the sum of squares of the Lack-Of-Fit to the regression model (S_{LOF}). The test statistic for Lack-Of-Fit, F_o is:

$$\hat{Y}_i = a + bX_i$$

$$S_E = \Sigma(Y_i - \hat{Y}_i)^2$$

$$S_{PE} = \sum_{i=1}^{m} \sum_{u=1}^{n} (Y_{iu} - \bar{Y}_i)^2$$

$$S_{LOF} = S_E - S_{PE}$$

$$F = \frac{S_{LOF}/(m-2)}{S_{PE}/(nm-m)} \sim F_{(m-2),(nm-m),\alpha}$$

If the null hypothesis of model adequacy is rejected, then the model must be abandoned since no linear raltionship exists between two varibales. If the null hypothesis cannot be rejceted ($p>\alpha$), then there is no apparent reason to doubt the adequacy of the regression model, so that the linearity is implied.

(II). *Quadratic Regression Test*

The quadratic regression test is proposed by the NCCLS (National Committee for Clinical Laboratory Standards)[34]. The data are fitted to a quadratic model first and if the results show a poor fitting in comparison with a first order model, the adequacy of the linear model is suggested. The manual algorithms for the quadratic regression were described respectively by Cardone et al,[32] and Burnett[35].

For a set of data of X and Y

X	$X_1 \cdots\cdots X_i \cdots\cdots X_n$
Y	$Y_1 \cdots\cdots Y_i \cdots\cdots Y_n$

Premise

$$Y = b_o + b_1X + b_2X^2 + e,$$

$$e = N(0,\sigma_e^2)$$

$$b_o = Y - b_1X - b_2[(X^2)/n]$$

$$b_1 = \frac{\Sigma[X^2-\Sigma(X^2)/n]^2\Sigma(X-\bar{X})(Y-\bar{Y})-\Sigma[X-\bar{X})\Sigma[X^2-\Sigma(X^2)/n][\Sigma[X^2-\Sigma(X^2)/n][Y-\bar{Y}]}{\Sigma(X-\bar{X})^2\Sigma[X^2-\Sigma(X^2)/n]^2-\Sigma[X-\bar{X})\Sigma[X^2-\Sigma(X^2)/n]^2}$$

$$b_2 = \frac{\Sigma[X-\bar{X})^2\Sigma[X^2-\Sigma(X^2)/n][\Sigma[Y-\bar{Y}]-\Sigma[X-\bar{X})\Sigma[X^2-\Sigma(X^2)/n]\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma[X-\bar{X})^2\Sigma[X^2-\Sigma(X^2)/n]^2-[\Sigma[X-\bar{X}]\Sigma[X^2-\Sigma(X^2)/n]^2}$$

$$\hat{Y}_i = b_o + b_1X_i + b_2X_i^2$$

$$(S_{b2})^2 = \frac{\Sigma(Y_i-\hat{Y}_i)^2/(n-3)}{\Sigma X_i^4-(\Sigma X_i^2)^2/n-[n^2(\Sigma X_i^3)^2-2n\Sigma X_i\cdot\Sigma X_i^2\cdot\Sigma X_i^3+(\Sigma X_i)^2(\Sigma X_i^2)^2/[n^2\Sigma X_i^2-n(\Sigma X)^2]}$$

The test statistic for $b_2=0$, F_o is:

$$F_o = (b_2)^2/(S_{b2})^2 \sim F_{1,n-3,0.05}$$

In a statistical sense, the Lack-Of-Fit test is more useful than the quadratic regression test because a showing of no significant Lack-Of-Fit indicates that the model is adequate to fit the data. On the other hand, the quadratic re-

Table 4. Design of experiment in the evaluation of recovery

% Amount Found (Y)	% Amount Added (X)
Y_{11} • Y_{1n}	50
Y_{21} • Y_{2n}	80
Y_{31} • Y_{3n}	100
Y_{41} • Y_{4n}	120
Y_{51} • Y_{5n}	150
$n \geqq 5$	

gression test merely shows that the quadratic model is not significantly better than the first order model, whereas both may be inadequate. However, sets of replicate measurements areneeded for the Lack-Of-Fit test, whereas the quadratic regression test may utilize the data of single determination to select the regression order.

III. *Regression Line Passing Through the Origin*

In LC procedure employing absorbance spectrometry, plots that pass through the origin should be the general rule. Namely, the response must be zero, if the concentration of analyte is zero. A nonzero intercept suggests the presence of matrix effects (interfering substances) in the sample being analyzed and constant bias may exist. A statistical test using the null hypothesis: (the intercept is zero) can be done to suggest a zero intercept.

The test statistic for the null hypothesis, t_o is:

$$t_o = \frac{a}{S_a} \sim t_{(nm-2,\alpha/2)}$$

$$S_a^2 = \frac{\Sigma(Y-\bar{Y})^2}{nm-2}\left[\frac{1}{nm}+\frac{\bar{X}^2}{\Sigma(X-\bar{X})^2}\right]$$

If t_o is smaller than Student's $t_{(nm-2,\alpha/2)}$ value, it is likely that the apparent Y-intercept is an artifact of experimental imprecision; then the regression; equation passing through the origin may be expressed in the form of $Y=[\Sigma XY/\Sigma X^2]X$. If the test statistic t_o is larger than $t_{(nm-2,\alpha/2)}$, the real nonzero intercept may exist. Thus, depending on the condition of the sample when injected (degree of cleanup, composition of excipients--etc), the size of the intercept varies. Therefore, this may lead to the reduced precision and should be avoided if possible. If it is impossible to improve the analytical system to obtain a zero intercept, the method of standard additions (MOSA) is useful to confirm that matrix effect are not important. The method is performed with the addition of known amounts of the analyte to samples to that a predictable increase in analyte concentration, as assayed by the HPLC procedure, would result. The test is applied to several different samples to make sure that the matrix effects are the same for a representative group of samples. The test of zero intercept employing Student's t-test may not be reliable sometimes, since the test is based on the magnitude of the stadard error. If the standard error appears to be relatively large, then relative variation of the regression may provide the precision of the assay method.

III. *Accuracy*

Accuracy refers to the closeness of an individual observation or mean to the true value. The accuracy of an assay method can be evaluated by the test of recovery. Recovery is established by the analysis of spiked placebo at sub- and super-potent amounts. It is important to make sure the drug is incorporated into the placebo matrix just the same as into the real sample product. Bohidar and Peace[36] recommen-

ded an experimental design as shown in Table 4 to test a 100% recovery. The standard labeled amount of the analyte is taken as 100% and the samples containing theoretical amounts of 50 to 150% are assayed. The percent amount found by the HPLC assay and the percent amount added are subjected to the statistical equality test. Recoveries can be determined by either external or internal standard methods. The external standard method is the straightforward method but the precise control of injection volume is mandatory. If it is difficult to control the injection error to a degree not greater than 0.2%, then the internal standard method may be preferred. An internal standard must be completely resolved from all other bands in the chromatogram, have similar solubility behavior and acidity with the interest analyte. It is desirable that the internal stadard elutes near the interest analyte and the detector response are similar to the interest analyte. The internal standard enables compensation for sample losses during extraction, cleanup and injection volume errors.

A linear regression model of $Y=a+bX+e$, $e=N(0,\sigma_e^2)$ is employed to fit the data in Table 4. The joint hypothesis to test a 100% recovery are:

$$H_o: \begin{cases}\hat{a}=0\\ \hat{b}=1\end{cases} \qquad H_1: \begin{cases}\hat{a}\neq 0\\ \hat{b}\neq 1\end{cases}$$

The test statistic F_o for a 100% recovery is:

$$F_o=[M(a-\hat{a})^2+\Sigma X^2(b-\hat{b})^2+2N\bar{X}(a-\hat{a})(b-\hat{b})]/(2S^2)\sim F_{2,(N-2),\alpha}$$

where $a=\bar{Y}-b\bar{X}$; $b=\Sigma(Y-\bar{Y})(X-\bar{X})/\Sigma(X-\bar{X})^2$;

$S^2=\Sigma(Y-a-bX)^2/(N-1)$

N:number of paired data.

If the test statistic F_o is smaller than $F_{2(N-2),\alpha}$, then the point $\hat{a}=0,b=1$ is contained in the (1-α) confidence ellipse, suggesting a 100% recovery of the assay method. Hence, there is no apparent reason to doubt the absence of the constant and proportional biases based on the (1-α) confidence region.

Howerer, it is not uncommon to observe incomplete recovery in trace measurements and in the assay of biological specimen. In such cases, the acceptable criteria proposed by FDA are 80-100% recovery for the concentration above 100 ppb and 60-110% recovery for the concentration below 100 ppb[12], with satisfactory reproducibility.

II. *Precision*

Observations relatively close in magnitude are considered to be precise as reflected by a small standard deviation. Relative standard deviation (RSD) or coefficient of variation (CV) is used to describe the precision. RSD or CV can be estimated by the following equation.

$$\text{RSD (or CV)}\% = \frac{100}{\bar{X}}[\Sigma(X-\bar{X})^2/(N-1)]^{1/2}$$

where N: number of replications, preferably fewer than 5.

The concept of precision involves parallel assay, repeatability and reproducibility of an analytical method. The RSD of parallel assay is the variability of the determination of a sample solution in repeated measurements, which reflects the injection and detection errors. Repeatability is assessed by replicate determinations with the same sample and reprocessing, using the same instrument--etc; whereas reproducibility refers to the stability of replicate determinations with different samples, analysts, instrument or interlaboratory trials. The RSD should be estimated at least for the high, medium and low concentrations within the dynamic linearity range for intraday and interday variability in terms of repeatability (using the same sample and reprocessing) and reproducibility (resampling tests). Tables 5a and 5b depict useful forms for the determination of intraday and interday variability. The 95% confidence intervals of the repeatability and reproducibility for the expected assay concentration are useful

Table 5a. Intraday and interday variability

Day	Replication	Concentration					
		Low___mcg/ml		Medium___mcg/ml		High___mcg/ml	
		Repeat.	Reprod.	Repeat.	Reprod.	Repeat.	Reprod.
	1						
	•						
1	•						
	•						
	5						
Mean±SD							
Intraday CV%							
	1						
	•						
2	•						
	•						
	5						
Mean±SD							
Intraday CV%							
	1						
	•						
3	•						
	•						
	5						
Mean±SD							
Intraday CV%							

* Repeatability: use the same sample but repeated preparation of sample solution; Reproducibility: resampling and re-processing the sample preparation.

Table 5b. Low, medium, high conc.____mcg/ml

Day	1	2	3
Replication	Repeatability/Reproducibility		
1	X_{11}	X_{21}	X_{31}
2	X_{12}	X_{22}	X_{32}
3	X_{13}	X_{23}	X_{33}
4	X_{14}	X_{24}	X_{34}
5	X_{15}	X_{25}	X_{35}
$T = T_1 + T_2 + T_3$			

Analysis of Variance (ANOVA)

SV	SS	df	MS
Interday	$S_1=[T_1^2+T_2^2+T_3^2]/5\text{-}T^2/15$	2	$V_1=S_1/2$
Intraday	$S_2=\Sigma(X^2_{11}+\cdots\cdots+X^2_{35})$ $\text{-}[T^2_1+T^2_2+T^2_3]/5$	12	$V_2=S_2/12$

information of the precision. In general, the RSD should not be greater than 1.5% for the parallel assay test, nor greater than 2% for repeatability. However, in cases involving biological specimen, around 5% RSD is not uncommon. Especially when trace measurements are required, 5 to 15% RSD or CV may be tolerable.

$$\bar{X}=\frac{\Sigma X_{ij}}{15}$$

$$\text{Interday CV, \%}=\frac{\sqrt{V_1}}{\bar{X}}\times 100$$

$$\text{Interday CV, \%}=\frac{\sqrt{V_2}}{\bar{X}}\times 100$$

95% Confidence interval of Interday

$$=\bar{X}\pm t_{2,0.05}\sqrt{\frac{V_1}{3}}$$

95% Confidence interval of Intraday

$$=\bar{X}\pm t_{4,0.05}\sqrt{\frac{V_2}{5}}$$

VI. *Sensitivity*

Sensitivity of an analytical method refers to the limit of detection (LOD) as well as the limit of quantitation (LOQ) in the USP. The limit of detection is typically determined at low concentrations and in the middle of the normal range of testing. To determine sensitivity at low concentrations, the concentrations are prepared beginning with a control of zero dose, and the concentrations are increased by appropriate increments. If a concentration shows a statistically significant response that is different from the control, then that difference in concentration is considered detectable. When determination of sensitivity at higher concentrations is performed, a sample of concentration near the middle of the normal range of testing replaces the zero control, and concentrations of about this value which vary by incremental amounts are prepared and tested. The significant difference is determined by Dunnett's t statistic for comparing a set of means with a control. The test statistic t_o is:

$$t_o=\frac{\bar{Y}_i\text{-}\bar{Y}_o}{\sqrt{2S^2/n}}$$

where $\bar{Y}_o$:mean for the control.

$\bar{Y}_i$:mean for the i-th concentration.

n:numbers of response at each concentration.

S^2:pooled estimate of the veriation within the concentration.

$$S^2=\sum_{i=1}^{k}\sum_{j=1}^{ni}(Y_{ij}-\bar{Y}_i)^2/[\sum_{i=1}^{k}n_i-k]$$

where Y_{ij}:j-th value for the i-th concentration.

K:numbers of concentration including the control.

n_i:numbers of response for the i-th concentration.

The test statistic t_o is compared with a critical value obtained from Dunnett's table. According to IUPAC[37], S_B (standard deviation of noise bands) is estimated with one-fifth of the sum of the peak height of maximum positive and negative noise bands. The LOD is expressed as $3 \cdot S_B/b$ and LOQ is $10 \cdot S_B/b$, where b is the

Table 6. Design of experiment in assessing the ruggedness of an analytical method

Subtle Condition	Treatment Combination 1	2	3	4	5	6	7	8
X_1	+	+	+	+	−	−	−	−
X_2	+	+	−	−	+	+	−	−
X_3	+	−	+	−	+	−	+	−
X_4	+	+	−	−	−	−	+	+
X_5	+	−	+	−	−	+	−	+
X_6	+	−	−	+	+	−	−	+
X_7	+	−	−	+	−	+	+	−
Assay Value	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8

slope of the standard calibration line. These definitions are also adopted by the USP XXII. Another general rule of thumb in estimating the limit of quantitation is the use of a signal-to-noise ratio of not less than 5:1, and the limit of detection of not less than 3:1. Nevertheless, the physical meaning of the limit of quantitation can only be interpreted under a specific level of accuracy and precision.

VII. *Ruggedness*

Considerable ruggedness against a small perturbation of experimental variations may be required for an assay method, especially when the method will be used in multiple laboratories where the sources of solvents, reagents, analytical columns and the models of apparatus may be different. The test of ruggedness of an analytical assay method can be carried out by using the experimental at model developed by Box et al.[38].

The model provides seven subtle conditions ($X_1 \cdots X_i \cdots X_7$) for eight determinations. Each subtle condition or factor takes two levels (+/−). These subtle factors represent different brands of reagents, columns, solvents, apparatus, small change in flow rate, pH value, amounts of bulffers and different analysts ar laboratories···etc. The proposed experimental at design is depicted in Table 6.

The size of the difference due to i-th factor, determines the degree of effectiveness of the factor. Namely, the larger the magnitude of the difference, the greater is the influence of the factor on the assay value. An estimate of the variance (S^2) is obtained by:

$$S^2 = [\frac{1}{7} \sum_{i=1}^{7} (\bar{X}_{i(+)} - \bar{X}_{i(-)}]$$

$$where\ \bar{X}_{i(+)} = \frac{1}{4}[\Sigma Y_{i(+)}]$$

$$\bar{X}_{i(-)} = \frac{1}{4}[\Sigma Y_{i(-)}]$$

Any difference due to i-th factor that is larger than $(t_{7,\alpha/2})\ S/\sqrt{8}$ may be regarded as a factor significant for the result of the assay ($p<\alpha$), and should be carefully specified and controlled.

VIII. *Validation of a Dual-System*

The dual-system validation can be applied to analytical comparisons such as (1) performance of two analysts; (2) validity of two different assay methods; (3) suitability of two models of instrument and (4) comparison of the assay results between two laboratories···etc. In most cases, the requirement for validating the dual assay system is to test the validity of a nonofficial method in comparison with the compendial me-

thod. In a practical sense, most of the official assay methods for drugs compiled in the USP and other Pharmacopeia are very good in ruggedness, but are often unsatisfactory in terms of the specificity and sensitivity for stability and pharmacokinetic sutdies. Usually, a more sensitive and specific analytical assay method than the compendial method is needed to carry out stability and pharmacokinetic studies. Although the more sophisticated alternative method can be validated by the single-system validation procedure presented in this report, it is often necessary to show the scientific rationale between two methods. This can be done by statistical evaluation of the means and the variances obtained from the two methods. If the assay results from a series of standard solutions within the dynamic linearity range of the standard method are statistically equivalent for the two assay methods, there is no apparent reason to doubt the validity of the alternative method. Table 7 depicts the design of experiment for method validation in the dualsystem.

A series of standard solutions with different concentrations are assayed by the compendial and the alternative methods respectively. The assay data are fitted to a linear model of $Y=a+bX+e$, $e=N(0,\sigma_e^2)$. The validity of the alternative assay method is tested for the joint null hypothesis of $a=0$ and $b=1$ as the method described in the test of accuracy. This test will suggest an equal mean.

Equal variance between the two assay results can be suggested by the Pitman-Morgan F test[39]. The test statistic F_o for equal variance is:

$$F_o=[\frac{1}{n-2}(\frac{S_x^2}{S_y^2}-1)^2]/[4(1-r^2)\frac{S_x^2}{S_y^2}] \sim F_{1,(n-2),\alpha}$$

where S_x^2 and S_y^2 are the variances of data X and Y respectively; n is the numbers of paired data and r^2 is the coefficient of determination of X and Y.

If the F_o value is smaller than $F_{1,(n-2),\alpha}$ value, there is no apparent reason to suggest the two variances are not equal.

Since there is no statistical difference between the means and between the variances of the two assay results, the validity of the alternative assay method in comparison with the compendial method can be established.

IX. *Qualification of Analysts*

Any validated analytical method will not function properly if the system is unstable and/or the analyst is unqualified. Table 8 depicts the major error sources and their acceptable criteria for an HPLC assay. The total RSD or CV value estimated from this table can serve as a guide to validate the system and the analyst. The total RSD for the assay involving biological specimen (BIO) can be estimated:

$$RSD_{TOTAL}^{BIO}=[0.2^2+0.5^2+5^2+0.2^2+1^2+1^2]^{1/2}=5.23\%$$

System repeatability is usually evaluated by

Table 7. Design of experiment for dual system validation

Theoretical Concentration	Found Concentration					
	Compendial (X)			Alternative (Y)		
A	X_{11},	X_{12},	X_{13}	Y_{11},	Y_{12},	Y_{13}
B	X_{21},	X_{22},	X_{23}	Y_{21},	Y_{22},	Y_{23}
C	X_{31},	X_{32},	X_{33}	Y_{31},	Y_{32},	Y_{33}
D	X_{41},	X_{42},	X_{43}	Y_{41},	Y_{42},	Y_{43}
E	X_{51},	X_{52},	X_{53}	Y_{51},	Y_{52},	Y_{53}

Table 8. Error sources and acceptable criteria in the HPLC assay operation[40]

Error Source	Acceptable Limit (RSD)
Sample weighing	<0.2%
Sample processing	<0.5%
Sample cleanup	<5.0%
Sample injection	<0.2%
Total imprecision due to glassware and pipet	<1.0%
Detection	<1.0%

parallel assay, using six replicate determinations of a standard solution. The RSD of the six chromatographic responses should not be greater than 1.5%. The total RSD for the assay of a pharmaceutical product or a simple chemical sample (SCS) can be estimated:

$$RSD^{SCS}_{TOTAL} = [0.2^2 + 0.5^2 + 0.2^2 + 1^2 + 1^2]^{1/2} = 1.53\%$$

Therefore, any analyst whose RSD^{SCS}_{TOTAL} is larger than 2% in five replicate determinations may be regarded as unqualified. In general, 2% RSD criteria is applied to the intraday repeatability of the assay of the relatively simple chemical sample at its optimum concentration, and not to the reproducibility of the assay of complicated sample or trace measurements.

ACKNOWLEDGEMENT AND NOTES

This work has been supported by research grant No.80009 from Mr. Cheng's Pharmaceutical Academic Foundation(鄭氏藥學文教基金會), Taipei, Taiwan, R.O.C. Computer software (PSJ) based on the model described has been written by Hsing-Chu Hsu and Hai-Lin Lu for IBM PC and its compatibles, and is available for use at the Foundation.

REFERENCES

1. U.S. Code of Federal Requlation. 1971. Jan. 25.
2. U.S. Code of Federal Requlation. 1978. Sept. 29.
3. Youden, W.J. 1964. Statistical Methods for Chemists. John Wiley & Sons Inc. N.Y.
4. Youden, W.J. and Steiner E.H. 1975. Statistical Manual of the Association of Official Analytical Chemists. A.O.A.C. Washington, D.C.
5. Mandel, J. 1964. The Statistical Analysis of Experimental Data. Interscience. N.Y.
6. Haynes, J.D., Pauls J. and Platt, R. 1977. The Greenbrier Procedure. Pharm. Mfg. Assoc./QC Section, Washington, D.C.
7. Bohidar, N.R. 1983. Proc. of American Statistical Association Annual Joint Meeting Biopharmaceutical Section. pp. 57-62.
8. Diding, N., Feiden, K., Fischer, P., Bentejac, R., Chissell, J. F., Furtwangler R. and Wiesenthal, K. 1980. Guidelines for Good Validation Practice (GVP). Pharmazeutische Industrie 42(10) : 982.
9. Kawamura, T. 1983. GMP Technical Reports Vol. 1. pp.67-70. Yakugio Jiho Shia, Tokyo.
10. The United States Pharmacopeial Convention, Inc. 1985. The Pharmacopeia of United States of America. XXI Revision. p. 1174.
11. Debesis, E., Boehlert, F. P., Givand, T. E. and Sheridan, J. C. 1982. Submitting HPLC Methods for Compendia and Regulatory Agencies. Pharm. Technol. 6(9) : 120-137.
12. Guerra, J. 1986. Techniques and Validation Procedures comply with FDA and GMP Regulations. Pharm. Technol. Japan. 2(6) : 555-559.
13. Finkelson, M. J. 1986. Requirements for Validation Methods. Pharm. Technol. Japan. 2 (7) : 693-698.
14. Sheinin, E. B. 1987. Laboratory Evaluation of Proposed NDA Analytical Methodology (Part 1). Pharm. Technol. Japan. 3(1) : 33-40.
15. Sheinin, E. B. 1987. Laboratory Evaluation of Proposed NDA Analytical Methodology

(Part 2). Pharm. Technol. Japan. 3(2) : 151-155.
16. Sheinin, E. B. 1987. Laboratory Evaluation of Proposed NDA Analytical Methodology (Part 3). Pharm. Technol. Japan. 3(3) : 257-260.
17. U.S.Department of Health and Human Services, Food and Drug Administration, Maryland, U.S.A. Feb. 1987. Guidelines for Submitting Samples and Analytical Data for Methods Validation.
18. Committee for Proprietary Medicinal Products Guidance Note on Analytical Validation, Final Draft, Document. 1989. No. 111/844/87. EN.
19. The United States Pharmacopeial Convention, Inc. 1990. The Pharmacopeia of United States of America. XXII Revision. pp. 1711-1712. Washington, D.C.
20. Carr, G. P. and Wahlich, J. C. 1990. A Practical Approach to Method Validation in Pharmaceutical Analysis. J. Pharm. Biomed. Anal. 8: 613-618.
21. Duarte, J. E. and Vest, D. K. 1979. Validating Alternate Laboratory Assay Methods. Pharm. Technol. 3 : 60-66.
22. Carey, R. N., Wold, S. and Westgard, J. O. 1975. Principal Component Analysis: An alternative to "Referee" methods in method comparison studies. Anal. Chem. 47 : 1824-1829.
23. Lawton, W. H., Sylvestre, E. A. and Young-Ferraro, B. J. 1979. Statistical Comparison of Multiple Analytic Proccdures : Application to clinical chemistry. Technometrics. 21 : 397-409.
24. Koopmans, L. H., Owen, D. B. and Rosenblatt, J. I. 1964. Confidence Intervals for the Coefficient of Variation for the Normal and Log Normal Distributions. Biometrika. 51 : 25-32.
25. Westgard, J. O. and Hunt, M. R. 1973. Use and Interpretation of Common Statistical Tests in Method Comparison Studies. Clin. Chem. 19 : 49-51.
26. Snyder, L. R., Glajch, J. L. and Kirkland, J. J. 1988. Practical HPLC Method Development. Chapter 2. New York : John Wiley & Sons.
27. Snyder, L. R. and Kirkland, J. J. 1979. An Introduction to Modern Liquid Chromatography. 2nd edition. Chapter 29. Wiley-Interscience. N. Y.
28. Martin, R. F. 1980. More on Quantitative Evaluation of Linearity. Clin. Chem. 26 : 1509-1510.
29. Mitchell, D. G., Mills, W. N. and Garden, J. S. 1977. Multiple-Curve Procedure for Improving Precision with Calibration Curve-Based Analysis. Anal. Chem. 49 : 1655-1660.
30. Bolton, S. 1984. Pharmaceutical Statistics. p. 209. Marcel Dekker Inc. U.S.A.
31. Wilson, T. D. 1990. Liquid Chromatographic Methods Validation for Pharmaceutical Products. J. Pharm. & Biomed. Anal. 8 : 389-400.
32. Cardone, M. J., Willavize, S. A. and Lacy, M. E. 1990. Method Validation Revisited: A Chemometric Approach. Pharm. Res. 7 : 154-160.
33. Montgomery, D. C. 1984. Design and Analysis of Experiments. 2nd editon. pp. 416-419. New York : John Wiley & Sons Inc.
34. NCCLS. 1979. PSEP-4. Instrument Evaluation Subcommittee, Evaluation Protocols Area. Villanova, PA. 1905.
35. Burnett, R. W. 1980. Quantitative Evaluation of Linearity. Clin. Chem. 26 : 644-646.
36. Bohidar, N. R. and Peace, K. E. 1988. In Biopharmaceutical Statistics for Durg Development. pp. 199-201. Peace, K. E.(ed). New York : Marcel Dekker Inc.
37. Spetrochim Acta. 1978. Nomenclature, Symbols, Units and their Usage in Spectrochemical Analysis II, Spectrochim Acta B33 : 242.
38. Box, G. E. P., Hunter, W. G. and Hunter, J. S. 1979. Statistics for Experiments. John Wiley & Sons Inc. N. Y.
39. Maloney, C. J. and Rastogi, S. C. 1971. Significance Tests for Grubbs' Estimators. Biometrika. 26 : 671-676.

40. Dolan, J. W. and Snyder, L. R. 1989. Troubleshooting LC System. p. 465. Humana Press Inc. U.S.A.

分析方法的驗證：高效液相層析定量法的簡易驗證模式

許興智　*簡俊生

嘉南藥專藥學科　*行政院衛生署藥物食品檢驗局

摘　要

本報告綜評近代定量分析法驗證之演化，優良藥品製造規範(GMP)及優良驗證規範(GVP)的法令要求。針對高效液相層析定量法(HPLC)整理出簡易可行之驗證模式。

對分析系，分析方法及操作人員之具體驗證步驟以及評價定量分析之正確性，精確度，專一性，靈敏度，堅牢性，非法定分析法及法定分析法的比較驗證所須用之實驗設計和統計解析皆適宜地闡述。

本驗證模式雖不能因應科學上嚴謹的各項驗證要件，但確可在一般實驗室所能容忍的時間及經費範圍內明確地彰顯分析系及分析方法的優缺點。診斷並提示分析系，分析方法及操作上之缺陷及問題之所在，以供改善分析條件，並確認改善之效果。