Journal of Pharmaceutical & Biomedical Analysis
Vol. 8, Nos 8–12, pp. 613–618, 1990
Printed in Great Britain

0731–7085/90 $3.00 + 0.00

A practical approach to method validation in pharmaceutical analysis*

G.P. CARR†‡ and J.C. WAHLICH§

‡*Medgenix Group, B-6220 Fleurus, Belgium*
§*Glaxo Group Research, Ware, UK*

Abstract: Guidelines issued by Regulatory Authorities make it clear that validation of analytical methodology is now widely required in support of registration dossiers. Although some attempts are made at defining terms and some vague indications are sometimes provided within these guidelines, no clear advice is provided on how validations should be conducted and what results should be expected. In this paper it is attempted to suggest some practical approaches to conducting validation and in particular to the determination of accuracy, linearity and limit of detection/quantitation.

Keywords: *Analytical validation; accuracy; linearity; limit of detection; limit of quantitation.*

Introduction

Analytical method validation has been defined as a procedure used to prove that a test method consistently yields what it is expected to do with adequate accuracy and precision. Over recent years the Regulatory Authorities have become increasingly more aware of the necessity of ensuring that the data provided to them in applications for marketing (and perhaps clinical) authorizations have been acquired using validated analytical methodology. This has resulted in the publication of a series of requirements and guidelines by various authorities [1–4].

It is the opinion of the present authors that the contents of these documents are often very vague, sometimes quite inaccurate and misleading and rarely provide the development analyst with guidance on what should really be required of a validation exercise. In addition, a survey of recent literature on analytical development reveals very variable standards employed for validation.

Performance parameters which should be addressed in a validation exercise include: accuracy; precision; repeatability and reproducibility; limit of detection and quantitation; linearity; selectivity; ruggedness/robustness; and stability of analytical solutions.

The parameters that require validation and the approach adopted for each particular case are dependent on: (a) the purpose of the method, and (b) the sample matrix. With respect to the purpose of the method, one should consider for example whether it is to be applied to a determination of a major or minor component, or for evaluation of performance characteristics (e.g. dissolution test for a solid dosage form) and whether the test will then be used to support a release specification or a stability study.

In this paper it is intended to review and then demonstrate practical approaches by which accuracy, linearity and limit of detection/quantitation may be evaluated for various analytical methods.

Validation Parameters

Accuracy

A reasonable definition is provided in USP XXII [2]: "The accuracy of an analytical method is the closeness of test results obtained by that method to the true value".

FDA Guideline [1] requires that data be provided ". . . over the range of interest (ca. 80–120% of label claim)". Some details of how accuracy should be determined are provided in USP XXII and CPMP Guidelines [2, 4], which both contain some useful information, but there are still a number of questions left outstanding, as noted below.

*Presented at the "Second International Symposium on Pharmaceutical and Biomedical Analysis", April 1990, York, UK.

†Author to whom correspondence should be addressed.

Table 1
Recommended validation ranges for linearity studies

Purpose of analysis	Typical range (%)	Recommended validation range (%)
Release specification assay	95 to 105	80 to 120
Check specification assay	90 to 110	80 to 120
Content uniformity test	75 to 125	70 to 130
Assay of a preservative in a stability study	50 to 110	40 to 120
Determination of a degradant in a stability study	0 to 10	0 to 20

What tests need to be validated for accuracy? If one accepts the USP XXII definition, it is arguable that virtually every physico-chemical test method will require some form of validation for accuracy.

Design of accuracy determination for different types of analytical method.
(a) *Identity tests.* It should be demonstrated that a positive result is obtained if the analyte is present and a negative result if it is absent.

(b) *Physico-chemical characteristics (e.g. MP, optical rotation).* It should be demonstrated that the result is not influenced by other components which may be present in the sample matrix.

(c) *TLC tests for related substances.* It should be demonstrated that the impurity of interest in the presence of the principal component provides a response to the method of detection which is comparable to that of a standard application of the same quantity of the impurity. In addition, it is a popular approach in such TLC tests to compare the responses of impurity zones in sample applications with those of the principal zones in standard applications which are prepared by dilution of sample solutions. This approach is widely adopted in pharmacopoeial monographs and should be justified by demonstrating equivalence of response between the principal component and its main impurities.

(d) *HPLC/GC tests for related substances.* The approach described above of spiking the sample with the impurity at around the level of interest would be applicable here also. In this case, however, test results will probably be calculated by electronic integration of detector responses and based on peak heights or areas. The validation should then be evaluated by comparing the integrator values obtained in the same way for spiked sample with those of the equivalent standards. Normally for impurities for which the levels of interest are around 0.1–1.0%, a variance of not greater than ±5% may be considered acceptable.

(e) *Bulk drug assays.* The approach adopted depends on the assay method to be adopted. For titrations the expected equivalence point may be calculated on theoretical grounds, taking into account the number of titratable functions in the analyte and the MW, but it is recommended that this should then be verified by carrying out the intended titration procedure on a well characterized reference standard.

For light absorption (UV) assays based on specific absorbance values, i.e. A(1%, 1 cm), validation would require that this value be well selected. It should be noted here that these values for closely related compounds may vary considerably, so it would also be appropriate to demonstrate that this is not significantly influenced by the presence of likely impurities at their intended maximum limits.

For chromatographic assays, the mass balance approach of Kirschbaum *et al.* is recommended [5]. This is based on the comparison of total peak response with and without the chromatographic column in place; a variance of not more than ±2% should normally be achieved.

Design of accuracy determination for different sample matrices. Here it is intended to discuss additional complications which arise when the procedure includes extraction of the analyte from a sample matrix prior to carrying out a measurement. So it is now necessary to demonstrate extraction efficiency in addition to the above discussed parameters. This type of determination is generally referred to as "recovery" and is the aspect of accuracy validation which has been best recognized by the Regulatory Authorities, some of whom refer to the "technique of analysis of spiked, active-free

samples". There are still a number of practical issues to address, as noted below:

(a) *Spiking range*. This will be dependent on the intended purpose of the method, but for a procedure intended for a specification assay, a range of 80–120% is normally considered appropriate. This may also be suitable for an assay to support a stability study, but if degradation below 80% is anticipated this should be taken into consideration. A suitable approach is to spike at five levels within this range and for each level the variance should normally be no greater than ±2% of the theoretical value.

(b) *Method of spiking*. To provide an appropriate model, consideration should be given to how the active will be introduced into the mixture of excipients. For example, in the case of a tablet, manufacture includes a compression stage and this should be taken into account.

Possibilities which have been used include addition of drug into excipient mixture as a solution in a volatile solvent and then drying under vacuum to produce more intimate mixing of drug with excipient.

Another possibility is a "recovery efficiency experiment"; this is most appropriate for HPLC and GC methods and requires an internal standard. The sample is extracted with solvent as intended in the final procedure, but with the addition of an internal standard. After centrifuging or filtering, about 75% of the supernatant solution is taken, subjected to the remaining procedure and the ratio of analyte response/internal standard response noted. The sample residue (including about 25% of supernatant from the first extract) is then re-extracted with a further volume of solvent without internal standard, centrifuged or filtered, subjected to the remaining procedure and the ratio of analyte response/internal standard response is again noted. The variance between the two ratios should not be greater than ±2%. This approach has the advantage that it can be conducted on "real samples" and does not require specially prepared "validation samples". However, it should be noted that this technique will not detect inefficiency of extraction due to some irreversible binding of drug to excipient.

Another approach would be to use a radiolabelled analogue of the drug, if this were available, and then to examine for the presence of activity in the excipient residue after extraction.

Linearity

A reasonable definition is provided in CPMP Guidelines [4]: "The linearity of a test procedure is its ability (within a given range) to produce results which are directly proportional to the concentration of analyte in the sample".

All guidelines recognize that linearity is a parameter which should be determined, but a number of issues are not addressed, including the following.

What tests need to be validated for linearity? This is only required for quantitative methods but not necessarily in all cases. It is not normally necessary to demonstrate linearity for a titration based on a well established equivalence factor.

Linearity is usually conducted to justify single point standardization, i.e. the analyst is assuming a linear response with zero intercept. In some methods, e.g. fluorescence, or atomic spectroscopy, this assumption is not made and calibration curves are used. Then it is not necessary to demonstrate linearity, but it should be shown that the calibration curve has a satisfactory slope sensitivity over the range of interest.

Design of linearity determination. The FDA Guideline [1] recommends a concentration range equivalent to 80–120% of the theoretical content of active. In the literature it is often seen that a range of perhaps 0–200% is examined. In practice the study should be designed to be appropriate for the intended analytical method, as shown in Table 1. The range selected for validation should not be unrealistically wide, as this may lead to rejection of a method which is really quite suitable for the intended purpose.

Evaluation of results. Very variable standards are observed in the literature. Authors may sometimes simply claim that the procedure was linear, with the additional mention of a "near zero intercept". More often, data are processed by linear least-squares regression and authors quote values obtained for the regression coefficients "a" and "b" of linear equation $y = ax + b$, together with a corre-

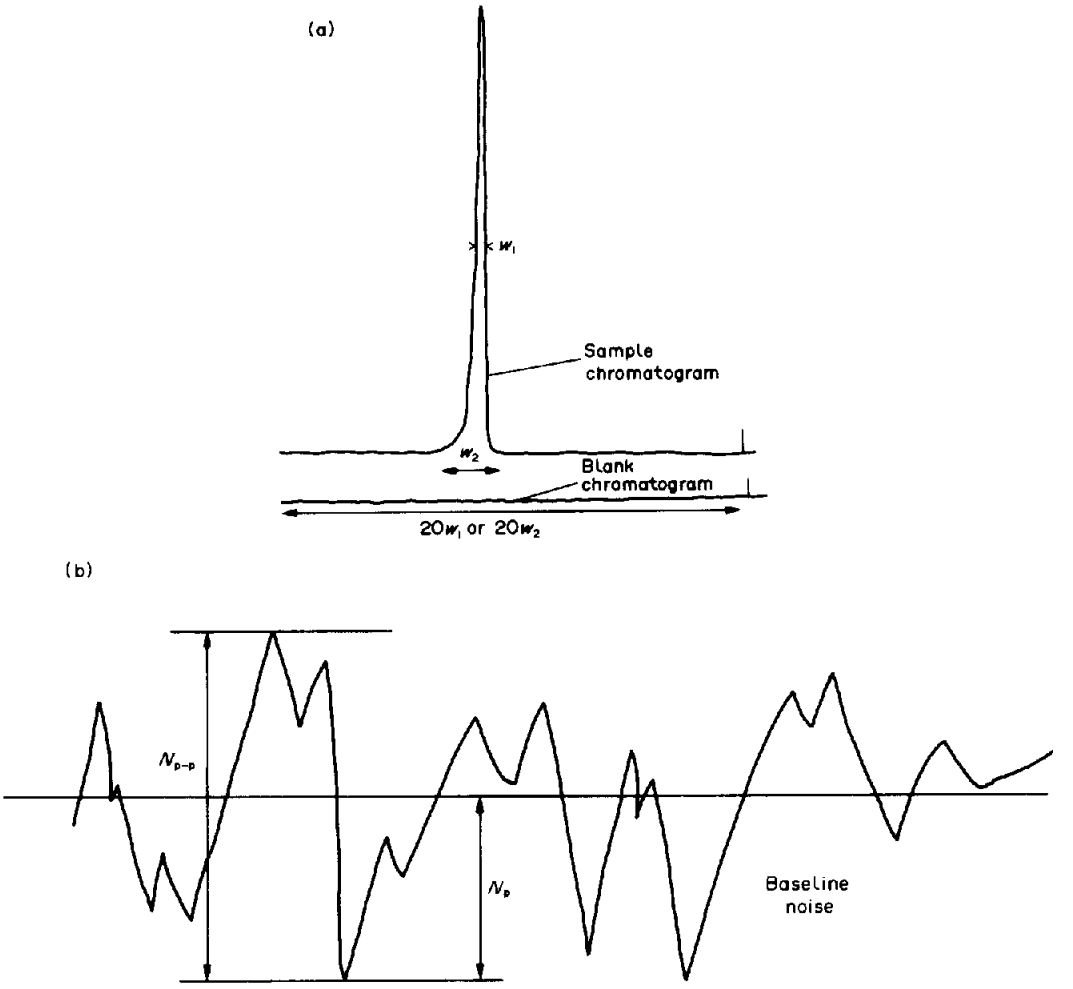

Figure 1
Estimation of chromatographic baseline noise. (a) Noise is measured over the region of a blank chromatogram corresponding to 20 times the width of the analyte peak and situated around the region where the analyte peak would be located in a sample chromatogram. (b) Noise magnitude is determined from the largest peak-to-peak fluctuation (N_{p-p}) or from the largest positive or negative deviation from the mean (N_p).

lation coefficient. This is inappropriate because:

(a) The value of a does not provide any information on the linearity of the procedure or on the goodness of fit to a regression line.

(b) The value of b is very important as it is the value of the intercept of the regression line. Unfortunately this is often presented as the number computed in the regression calculation and will not have any meaning to the reader without some further information, e.g. about the absorbance range of a detector. This is easily remedied by expressing b as a percentage of the value of analytical response at the 100% analyte level. It is then suggested that this should normally fall within the range −2.0 to +2.0% for the validation to be considered satisfactory.

(c) The value of the correlation coefficient as a criterion of linearity has been criticized in the literature [6, 7], but nonetheless authors of analytical development papers continue to adopt it, or its squared product.

The correlation coefficient (r) was developed to demonstrate whether or not any relationship exists between two sets of data. If none, $r = 0$; if a linear relationship exists, $r = +1$ or -1 depending on the slope.

For validation the question is not whether a relationship exists; it is normally expected to be linear. Another approach which has been suggested is based on curve fitting [7, 8]. Response versus concentration data are fitted to equations of the type: $y = ax^n + b$; $y = ax^z + bx + c$; $y = a(\exp)bx$; or $y = a + bx + c/x$, to establish the best fit.

These approaches may be of interest if the objective were to establish calibration curves, but in practice it is intended to use these data to justify the linear model: $y = ax$. This would then allow the conduct of analyses using single-point standardization. For this to apply, a satisfactory intercept value must be demonstrated as described above.

In addition, goodness of fit of data to the regression line may be evaluated by a procedure based on the residual sum of squares. Taking the regression line as the mean, a relative standard deviation (RSD) is calculated for the data; normally this value should not be greater than 2.0%, but when evaluating this determination, the results of precision determinations should also be taken into account.

Limit of detection (LOD)/quantitation (LOQ)

Limit of detection (LOD) is a parameter of limit tests and may be defined as the smallest quantity of analyte which may be expected to produce a response which is significantly different from that of a blank.

Limit of quantitation (LOQ) is a parameter of determination tests for minor components and may be defined as the smallest quantity of analyte which can be determined with acceptable accuracy and precision.

What tests need to be validated for LOD/LOQ? These are essentially the parameters of "trace analysis methods", i.e. procedures that require methods to be operating at or close to maximum attainable sensitivity. It may require judgment from the development analyst whether this validation is appropriate, but in general it may be assumed that it will be required for tests for impurities including: TLC and HPLC tests for impurities and related substances; GC tests for residual solvents; atomic spectroscopy tests for metals, e.g. catalysts.

However this validation may also be required to support some assay methods such as content uniformity tests for dosage forms containing very small amounts of active per unit dose, or assays of dissolution test samples.

Determinations for non-instrumental methods. This would probably most commonly apply to TLC tests for related substances. In practice LOD determination may be achieved by chromatographing samples of analyte which have been spiked with decreasing quantities of the impurity of interest to establish the smallest quantity which can be reliably detected. The following points should be noted.

(a) The weight of sample and volumes of sample solutions applied to the plate should remain constant and be the same as intended in the analytical procedure.

(b) After development, plates should be examined by all intended detection methods, e.g. UV radiation at 254 nm, UV radiation at 366 nm and any spray reagents. LOD values should then be separately reported for each method of detection.

(c) Results are best expressed as a weight and also as the equivalent percentage (or relative scale of units) in relation to the drug of interest e.g. x (μg) and equivalent to y (%) or z (ppm).

USP XXII [2] also includes a proposal for the determination of LOQ for non-instrumental methods, but in the opinion of the authors this is not appropriate.

Determination of LOD and LOQ for instrumental methods. According to IUPAC [9] for spectrochemical methods these values may be determined for the smallest concentration (c_1) or amount (q_1) from: c_1 (q_1) = ks_B/S, where k is a constant, s_B is the standard deviation of analytical blank signal, and S is the slope of response versus concentration curve (slope sensitivity). Recommended values for k are: for LOD, $k = 3$; and for LOQ, $k = 10$.

Justification of these values has been provided by Long and Winefordner [10] and by the Analytical Methods Committee [11].

For this to be used in practice then, requires determination of the value of s_B, which is related to baseline noise.

For spectrochemical methods this is not a problem as the analyst can conduct a series of absorbance determinations with blank solutions, using instrumental conditions intended for sample determinations. The standard deviation may then be calculated.

For chromatographic methods, the method is not so obvious, but an approach based on one proposed by Foley and Dorsey is recommended [12]. This has now been introduced into the European and British Pharmacopoeia monographs on gas chromatography and liquid chromatography as the basis of the method for determining signal-to-noise ratios [13, 14].

The value of s_B is estimated from the magnitude of noise in a blank injection over a

representative section of baseline which is 20 times the peak width of the analyte, as illustrated in Fig. 1. There is no internationally recognized approach to conducting this determination and some workers use 20 times the baseline peak width (w_2) [12], while others use 20 times the peak width at half height (w_1) [13–15].

The noise magnitude may then be determined from the largest peak-to-peak fluctuation (N_{p-p}) [12], when $s_B = N_{p-p}/5$, or from the largest deviation from the mean (either positive or negative) (Np) [13–15], when $s_B = N_p/2$.

Whichever approach is adopted it must be noted that the results depend on predictions made with analytical blank solutions. In addition, estimations of noise intensity are based on the heights of noise fluctuations, whereas the analytical method being validated may well be based on peak area measurements.

Thus, it is essential that the development analyst verifies the predicted results by examining sample solutions containing the computed concentrations of analyte to ensure that the values are indeed realistic.

Conclusions

It is clear from the various guidelines issued by Regulatory Authorities that analytical methodology should be thoroughly validated. Although the guidelines provide indications of the validation required, they do not normally include much information on acceptable approaches to conducting these studies, or on the kind of results which should be considered acceptable. This is not very surprising, as such detail can be very dependent on the nature of the sample, the type of analytical methodology and the purpose of carrying out the test.

Some indications are provided in this presentation of approaches which may be adopted for validating various types of analytical method, in particular for accuracy, linearity and LOD/LOQ. Proposals for ensuring that methods continue to maintain their valid status during routine use (system suitability tests) are presented elsewhere [16].

In considering these validation procedures, the authors have attempted to recommend approaches which will challenge the method being examined to demonstrate that it is capable of providing the desired information.

References

[1] Guideline for Submitting Samples and Analytical Data for Methods Validation. US Department of Health and Human Services, Food and Drug Administration, Maryland, USA, Feb. (1987).
[2] United States Pharmacopeia XXII, Section 1225, 1710–1712. United States Pharmacopeia Convention Inc., Rockville, MD (1990).
[3] Department of Health Medicines Act 1968, Guidance Notes on Applications of Product Licences (MAL2) A20-7 and A2E-2. HMSO (1987).
[4] Committee for Proprietary Medicinal Products Guidance Note on Analytical Validation, Final Draft, Document No. 111/844/87-EN, Aug. (1989).
[5] J. Kirschbaum, S. Perlman, J. Joseph and J. Adamovics, *J. Chromatogr. Sci.* **22**, 27–30 (1984).
[6] M.F. Delaney, *Liq. Chromatogr.* **3**, 264–268 (1985).
[7] Analytical Methods Committee, *Analyst* **113**, 1469–1471 (1988).
[8] E. Debesis, J.P. Boehlert, T.E. Givand and J.C. Sheriden, *Pharmaceut. Technol.* 120–137 (1982).
[9] Nomenclature, Symbols, Units and their Usage in Spectrochemical Analysis II, *Spectrochim Acta B* **33**, 242 (1978).
[10] G.L. Long and J.N. Winefordner, *Anal. Chem.* **55**, 712A–724A (1983).
[11] Analytical Methods Committee, *Analyst* **112**, 119–204 (1987).
[12] J.P. Foley and J.G. Dorsey, *Chromatographia* **18**, 503–511 (1984).
[13] European Pharmacopoeia (2nd edn), 11th Fascicule, V.6.20.3 and V.6.20.4, Maisonneuve SA, Sainte-Ruffine, France (1987).
[14] British Pharmacopoeia, Vol. II, A82 and A84. HMSO, London (1988).
[15] J.E. Knoll, *J. Chromatogr. Sci.* **23**, 422–425 (1985).
[16] J.C. Wahlich and G.P. Carr, *J. Pharm. Biomed. Anal.* **8**, 619–623 (1990).

[Received for review 5 April 1990; revised manuscript received 1 June 1990]