Pergamon

0731-7085(94)00106-5

Journal of Pharmaceutical & Biomedical Analysis, Vol. 13, No. 2, pp. 89–97, 1995

0731-7085/95 $9.50 + 0.00

Application issues in bioanalytical method validation, sample analysis and data reporting

DARIOUSH DADGAR,* PAMELA E. BURNETT, M. GERRY CHOC,† KEITH GALLICANO‡ and JOHN W. HOOPER

Analytical Department, Phoenix International Life Sciences, St Laurent (Montreal), Quebec, Canada
† Bioanalytical Department, Sandoz Pharmaceuticals Corporation, East Hanover, NJ, USA
‡ Drugs Directorate, Health Protection Branch, Bureau of Drug Research, Health Canada, Ottawa, Canada

Abstract: Although some degree of consensus has been reached concerning the requirements for acceptable method validation, the procedures used to establish them vary significantly between laboratories. Also, issues arising from application of these requirements during validation and subsequent sample analysis need to be addressed. The purpose of this paper is to discuss application issues concerning prerequisites to method validation, and all validation criteria for evaluation of method reliability and overall performance. Other poorly addressed issues such as re-validation, cross-validation, partial sample volume, multicomponent analysis and reporting will also be discussed. Although many issues discussed are of a general nature, the scope of this presentation is primarily to address issues arising from the validation and routine application of chromatographic methods.

Keywords: *Bioanalytical; chromatography; validation; stability; biofluids.*

Introduction

The validation of bioanalytical methods has been the subject of discussion in recent conferences and papers [1–8]. These have intended to provide guiding principles for validation of analytical methods used in bioavailability, bioequivalence and pharmacokinetic studies in humans and animals. It has been generally agreed that the key criteria for evaluation of method reliability and overall performance are:

(i) analyte stability;
(ii) method selectivity;
(iii) limit of quantitation;
(iv) accuracy;
(v) precision;
(vi) relationship between response and concentration (e.g. linearity);
(vii) recovery; and
(viii) ruggedness.

Although some degree of consensus has been reached concerning the requirements for acceptable method validation, the procedures used to establish them vary significantly among laboratories. Also, issues arising from application of these requirements during validation and in subsequent sample analysis, need to be addressed. A round table discussion at the recent AAPS meeting (15–19 November 1992 San Antonio, Texas) partly addressed some of these issues.

The purpose of this paper is to summarize broadly based views on application issues in bioanalytical method validation, sample analysis and data reporting from scientists engaged in bioanalytical analysis and its regulation. This paper will discuss application problems and potential solutions to these problems. Although many issues discussed are of a general nature, the scope of this paper is primarily to address issues arising from the validation and routine application of chromatographic methods.

Prerequisites to Method Validation

The first prerequisite for method validation is a developed analytical method. The purpose of method validation is to establish that an accurate, precise and rugged method has been

* Author to whom correspondence should be addressed.
Note: The contents of this paper represent the personal views of the authors and do not necessarily reflect the policies of the institutions with which they are affiliated.

developed. Criteria for evaluation of validation data should be established *a priori*.

However, often validation is begun on a "promising" method. If validation criteria are not met, further development work ensues and the method is then revalidated. In order to break the vicious circle of development work during validation, it is advisable to establish that the following parameters are acceptable prior to initiating the aspects of validation which require analysis of large numbers of samples (for example, intra-batch and inter-batch precision and accuracy).

Selectivity

It is recommended that the selectivity of the method should be established with respect to endogenous substances, metabolite(s) and known degradation products. Interferences which are likely to be present in small quantities may adversely affect the quantitation of unknown samples at concentrations approaching the limit of quantitation (LOQ).

Selectivity with respect to interferences from endogenous substances in biological fluids can usually be established by processing a minimum of six independent sources of the same blank matrix. Careful examination of chromatograms across the time windows of peaks of interest is required to evaluate selectivity. Here it should be emphasized that it is not appropriate to test only one source of blank [7]. Although it would be preferable that *all tested blanks*, if obtained under controlled conditions, be free from interference, factors such as subject food and beverage intake, use of vitamin supplements, use of over-the-counter and prescription drugs other than the one(s) being assayed, and cigarette smoking can affect selectivity. Therefore, even blanks obtained under "controlled conditions" may not ensure inter- and intra-subject uniformity both during selectivity testing and during the study. Due to factors discussed above, and/or method artifacts, zero per cent interference may not always be achievable, and in such cases, minor interferences may be allowable provided that the pre-defined precision and accuracy criteria for quantification at the LOQ are respected.

Real problems arise when analyte metabolites or known degradation products are not available. Under these circumstances, if possible, custom synthesis of metabolites and known degradation products should be undertaken, for use in demonstrating selectivity. In the absence of reference samples for metabolites or degradation products, either or both of the following experiments may sufficiently demonstrate selectivity.

(i) Use of biological samples from dosed subjects may be the best solution (e.g. from a pilot study). These samples should be analysed using the usual chromatographic conditions, and under "stretched" conditions i.e. changing chromatographic conditions to resolve as many potentially merged peaks as possible. Examination of chromatograms from subject samples collected at various times following the dose can reveal related peaks due to substances that are absent in the pre-dose sample and increase in response following the dose, and subsequently decrease.

(ii) If concentrations are high enough, and the UV spectra of potential metabolites/degradation products are different from that of the parent drug, diode array or other multiple wavelength detectors may be used to ensure peak purity.

(iii) In the absence of subject samples from a pilot study, just before the start of analysis of study samples, a few samples at the estimated t_{max} s for parent drug, metabolites and degradation products can be analysed, using only a portion of the available sample volume, as described above for pilot study samples. This will ensure that study samples will not be wasted should interference from metabolites/degradation products be found.

(iv) If analysis of samples from dosed subjects is not feasible, a number of the potential metabolites could be produced by *in vitro* incubation with liver homogenate, and resulting metabolites analysed using the method's chromatographic conditions.

(v) To account for potential degradation products produced by acid hydrolysis in the stomach, *in vitro* incubations in acidic media may be performed.

(vi) Anticipating interferences from concomitant medications and their metabolites may not always be feasible. Method modifications are usually required if such problems arise during analysis of the study samples, and may necessitate revalidation, the extent of which is described later in this paper.

(vii) Last, and by no means least, is to use a more specific detector such as the mass spectrometer. The new generation of atmospheric pressure HPLC–MS, MS–MS and

sensitive GC–MS instruments have greatly enhanced the power of analytical chemists to produce highly sensitive and specific methods.

Stability

To ensure that compound integrity is maintained throughout the workup process, certain stability tests should be performed during the final stages of method development.

Analyte stability. Published literature should be investigated, or laboratory tests should be conducted to determine whether pure analyte and/or solutions of the analyte (e.g. drug, metabolite(s) and internal standard) are stable under normal laboratory conditions of heat, humidity, light and air exposure.

In process stability. It must be demonstrated that drugs remain intact if left for several hours at room temperature in the biological matrix. Certain analytes, e.g. captopril, acetyl salicylic acid, etc. undergo immediate changes/degradation in the biological matrix. In such cases, appropriate additives, e.g. enzyme inhibitors, anti-oxidants and/or derivatizing agent may be added. Reduction of the collecting device temperature might improve stability. It may be necessary to extract samples immediately after collection, and store the dried frozen residue for later reconstitution and analysis. If this procedure is used, the calibration standards and QC samples must be treated in the same manner.

It should also be noted that some matrices encountered at intermediate stages of sample workup such as buffered biological matrix and the back-extraction medium may affect the stability. Again, if instability of analytes is observed at any of these stages, appropriate additives should be considered, or processing temperature reduced.

Processed sample stability. Reconstituted extracted samples must remain stable in the reconstitution solvent at the temperature of the auto-injector until they are injected, as well as at other lower temperature(s), e.g. 4°C, to allow for storage of prepared samples that cannot be immediately chromatographed owing to instrument problems. Processed sample stability should be examined in replicate at each QC concentration. After reconstitution, replicate samples of each QC concentration should be pooled, mixed and aliquoted for injection at time 0, intermediate times and after the expected maximum delay before completion of a batch of samples, e.g. 24, 36, or 48 h. By pooling the extracted QCs at each concentration, the starting concentration for all aliquots is identical, and the comparison of determined concentrations following storage to the time 0 concentration is independent of the extraction reproducibility.

Ruggedness

If it is discovered late in the validation process that a method is not easily transferred between systems, analysts, or analytical columns of the same type, further development and revalidation becomes necessary. To minimize the chances of this occurring, the following should be investigated during the latter stages of development.

Column effect. Analyses should be performed on at least two columns containing two different lots of the identical packing material, preferably one column which was used for method development and the other previously unused.

Mobile phase effect. For HPLC methods, the effect of small variations in the mobile phase solvent ratio (i.e. less than 2% of the amount of each component), and buffer pH were applicable, should be examined and reported. For GC methods, other parameters such as small variations in oven temperature and gas flow rates should be examined.

Assessment of the above parameters should provide an indication of the method's ability to maintain critical separations when faced with expected day to day mobile phase variations, and column to column variability.

Ruggedness acceptability. Peak shapes and resolution from other peaks in the matrix must remain visually acceptable. The limit of quantitation must demonstrate a reproducible response readily distinguishable from the noise level.

Internal standard (IS)

Many bioanalytical methods are based on the internal standard quantitation method. Internal standard is added at the earliest stage of sample preparation to minimize error by compensation. It is commonly believed that a good IS should be structurally similar to the

analyte under analysis. Although the above statement is generally true, there are occasions when it is not so. Such was the case in liquid–liquid extraction method followed by HPLC analysis of the tricyclic antidepressant doxepin, and its metabolite, nordoxepin (Dadgar, unpublished data). The structurally similar drug, imipramine did not minimize error by compensation for nordoxepin, whereas it successfully did so for doxepin in spiked samples. This is because parent and metabolite differed from each other in their extractabilities under the extraction conditions used. Under this circumstance, it was necessary to use two internal standards, imipramine as IS for doxepin, and desipramine as IS for nordoxepin.

Two other important criteria for choosing an internal standard are:

(i) it should elute close to the retention time of the analyte of interest. Therefore, for some multicomponent analyses two or more internal standards are needed;
(ii) it should normally elute after the parent drug peak (where reversed-phase chromatography is used) so that the possibility of interference with faster eluting more polar metabolites is obviated.

Method Validation

In this section, commonly encountered application issues in bioanalytical method validation are discussed under the heading of each validation criterion.

Calibration curve

The joint conference report [1] states that batch acceptance should be based on QC acceptance criteria. However, bioanalytical laboratories react differently and somewhat arbitrarily regarding inclusion or exclusion of a standard point deviating greatly from the calibration curve. Some leave it in the curve and some drop it. As it happens, there are times when, for example, the LOQ standard may deviate by, say 40%, but the next point in the calibration curve demonstrates a minor deviation from nominal. However, dropping the LOQ standard causes the next point to deviate greatly from the nominal value.

Another example of a commonly encountered problem is where removal of a given standard causes a QC to fail to meet its acceptance criteria when the QC would otherwise have been within the acceptable range. In light of the above, inclusion/exclusion of points in the standard curve should be established *a priori*, and the following provides possible guidelines for this.

Provided the calibration curve consists of at least seven non-zero single standards, up to two non-zero standrds may be removed from the calibration curve if at least one of the following valid reasons exists and a minimum of five non-zero standards remain in the curve.

(a) loss of sensivitity;
(b) poor chromatography;
(c) loss during sample processing;
(d) if, when included in the calibration curve, it clearly biases the QC results, and the back-calculated standard concentration deviates substantially from its nominal value.

In many bioanalytical laboratories, the calibration curves and QC samples are prepared simultaneously and frozen for storage, normally at the same temperature as is intended for storage of biological samples derived from clinical/toxicological samples. However, some laboratories prefer to prepare calibration curves fresh with each batch of samples analysed. The advantage of preparing and freezing calibration curves is that the effect of time and possible degradation of analyte(s) is the same for the calibration samples and the study samples.

In order to generate an accurate "analytical" calibration curve independent of possible time effect, the calibration curve may be prepared fresh with each batch of samples while QC samples are prepared and stored frozen with study samples in order to account for the "time effect". A simple approach to using fresh calibration curves is to prepare a series of working calibration standards at concentrations 10 or 20 times greater than those intended for biological standards, in a suitable dilution solution such as 1:1 methanol–water. These working calibration standards may be stored provided stability is previously demonstrated over the maximum period over which they will be stored. Then, on a daily basis, blank biological matrix is spiked with the working calibration standards in a ratio of, e.g. 1:20 working standard:biological blank. Dilution of biological matrix with working standards will be compensated for by adding an equal volume of the working standard dilution solution (free of analyte(s)) to the study samples.

Number of calibration curves

The question of how many calibration curves to run with each batch should be answered by consideration of whether or not study samples are run, for example, singly, in duplicate or in triplicate. If study samples are run singly, then a single calibration curve should normally be used. If replicates of study samples are analysed, then identical replication of standard curves is desirable.

In order not to mask the accuracy of the method, calibration curves should be processed identically during validation and during study sample analysis.

Validation of partial or increased sample volume

A common difficulty arises during biological sample analysis when less than the validated sample volume is available, and a partial volume must be used for analysis. This procedure needs to be validated. In addition, samples analysed and found to be above the calibration curve range require dilution for reanalysis.

A simple form of validation for partial sample volume is to use 1/4, 1/3 and 1/2 of replicate volumes of QC samples of appropriate concentrations. QC concentrations are chosen both above the calibration curve range to demonstrate accurate dilution to within the range, and at concentrations such that once diluted, they are near, but not below the LOQ. The partial sample volume is brought up to the validated volume by addition of blank matrix, and analysed. The calculated concentrations of diluted QCs, when multiplied by their respective dilution factors must fall within the defined precision and accuracy criteria for that QC.

Occasionally, it may be desirable to increase the volume of sample matrix, compared with that of the calibration standards, so that the instrument response is within the standard curve range. For example, when analysing urine, it may be necessary to use a larger volume to obtain reliable concentrations owing to a large void volume. The interpolated concentration, corrected for volume, can be below the LOQ concentration established for the assay, provided that the response is above that of the LOQ standard. Validation is necessary to show that selectivity is not compromised and the pre-defined criteria for accuracy and precision are met with the larger sample volume.

QC samples can be prepared at, e.g. 1/2 and 1/4 the LOQ for replicate extraction of double and quadruple the usual extraction volume, respectively. These LOQ QCs are analysed against the calibration curve extracted using the usual volume. The usual precision and accuracy criteria should be met. In addition, selectivity tests should be conducted using the increased blank urine volumes.

Limit of quantitation

The limit of quantitation (LOQ) must be differentiated from the limit of detection (LOD). The value of the LOD is the smallest concentration that can be distinguished from the background noise. The limit of detection can be defined in different ways. However, no matter how the limit of detection is defined, the limit of quantitation should be at least twice the response of the LOD, and within the pre-defined accuracy and precision boundaries, normally within $\pm 20\%$ of nominal with a RSD $\leq 20\%$.

It is important that the accuracy and precision of the LOQ be obtained using LOQ QC samples, i.e. independently from the calibration curve, because the LOQ standard which is used to construct the calibration curve influences the regression equation. In this manner, the precision and accuracy of the LOQ can be defined in both inter- and intra-batch tests. The practice of pushing the LOQ to the limits of the LOD in order to compete with other more sensitive methods should be avoided.

Stability

Data obtained from analysis of study samples are suspect if not accompanied by supporting data assuring the stability of the drug and metabolites in the matrix. It is important to emphasize that stability tests must be designed in such a way as to detect any degradation, over the maximum period of time that study samples will be stored prior to analysis. Often, stability data are based on duplicate or triplicate determinations of the concentrations of high and low QC samples, at multiple time points after the start of storage to allow "trends" to be detected. However, the issue is not whether there is a trend in degradation, but whether the study samples with the longest storage time are adequately

preserved at the time of analysis. For example, in order to demonstrate 12 month stability at −20°C, the stored samples should be analysed when freshly collected or prepared, at intermediate times (e.g. 3 and 6 months following preparation) and after 12 months of storage. To show processed sample stability, for example over 36 h (as discussed previously), samples should be analysed at time 0, at intermediate times such as 12 and 24 h, and at 36 h.

If stability testing is conducted on QC samples (as opposed to samples from dosed subjects) it is recommended to directly compare responses of stored and fresh QCs analysed at the same time, or to compare interpolated values provided the values are interpolated from the same calibration curve.

Two types of essential stability study have already been discussed. Here, stability of the analytes under prolonged storage conditions and freeze–thaw will be discussed.

Long term stability. The stability and tendency for adsorption to the storage container should be assessed for all analytes in the biological matrix using the exact type of container (e.g. glass, polypropylene) to be used for study sample storage. Stability must be proven over at least the maximum period of storage of study samples, under the temperature conditions to be used for study samples.

Ideally, the control samples used for evaluation of long term stability, should be those obtained from dosed patients/volunteers, collected, pooled, aliquoted and stored at the same time and in the same way as study samples. This provides maximum assurance of the integrity of all analytes in study sample for a given study. However, for most drugs where metabolite reversion to the drug, or ongoing metabolism in the frozen matrix are not problems, blank matrix samples spiked at different concentrations are an acceptable alternative and are usually used.

Two procedures for the assessment of long term frozen stability are discussed below, the first traditional, and the second developed recently by some of us [8].

(1) Replicates of stored and freshly prepared QC samples are analysed and their responses compared.
(2) A batch of stablility samples is prepared and divided into two sub-batches, the first stored at the temperature intended for use in storing study samples (typically −20°C; the stability samples), and the second in liquid nitrogen (nominally at ≤−196°C), or in another suitable freezer (colder than −130°C). The latter are the reference samples. After storage for an appropriate time, replicates of both the stability and the reference samples are analysed on the same standard curve and the results compared.

At very cold temperatures (≤−130°C) stability is assured even for unstable drugs. This has been demonstrated using theoretical calculations based on the Arrhenius equation, which show that at −130°C, (the temperature in the vapour above liquid nitrogen was used to be conservative) reaction rates are a factor of approximately 10^3 lower than at −20°C for reactions with very low activation energies (5 kcal $mole^{-1}$), and factors of 10^6–10^{13} lower for activation energies in the usual range (10–20 kcal $mole^{-1}$). We have confirmed this by evaluating the degradation of acetylsalicylic acid (ASA) in human plasma at −20°C and in liquid nitrogen. ASA is known to be one of the most unstable drugs in plasma. While the drug was degraded approximately 60% over 4 months at −20°C, no degradation was detected after storage in liquid nitrogen for the same period. Further, calculations based on the ASA reaction rate constant obtained from the −20°C data, and assuming an improbably low activation energy of 5 kcal $mole^{-1}$ (worst case), showed that calculated ASA degradation at −130°C was <0.05% after 4 months. These and other evaluations of this approach will be published separately.

Our studies suggest that the "second" method of assessing long term frozen stability has superior precision and accuracy to that of the traditional method. This is probably because it eliminates errors resulting from preparation of different fresh batches at each stability time point. It appears that such errors, when combined with assay variation, can lead to erroneous conclusions concerning stability or lack thereof. However, many laboratories do not have access to liquid nitrogen equipment, and this is a limiting factor in the application of this method.

Freeze–thaw stability. The influence of three

freeze–thaw cycles at concentrations representative of the high and low analyte concentrations in the matrix should be examined in replicates. Data for freeze–thaw stability may be obtained in two ways:

(i) stability samples are frozen and thawed three times and then analysed. This method is less time consuming, however, if samples are proven unstable, then experiment design, (ii) should be performed.

(ii) Quality control samples are analysed without being frozen at first, and then after each cycle of freeze–thaw so that a trend, if it exists can be seen. If it was proven, for instance, that QC samples are only stable after two cycles of freeze–thaw, then study samples should not be frozen and thawed more than twice.

Recovery

High recovery of analyte(s) from the matrix is desirable. However, sometimes it may be necessary to intentionally sacrifice high recovery in order to achieve better selectivity, and this is acceptable provided that adequate sensitivity, precision and accuracy are achieved. Solvents such as ethyl acetate normally give rise to high recovery of analyte, however this solvent simultaneously extracts many interfering compounds, therefore, provided that an adequately sensitive detection limit is attained with good precision and accuracy, the extent of recovery should not be considered an issue in bioanalytical method development and validation.

Recovery is best tested by directly comparing responses of replicates of extracted QC samples with replicates of extracted blank matrix to which analyte has been added at the same nominal concentration. In this way, any effect of the matrix on, for example, peak shape does not complicate interpretation of recovery data. Recovery may be determined by comparison of interpolated concentrations of extracted vs unextracted QCs provided they are from the same calibration curve.

Revalidation

When it is necessary to make changes to chromatographic conditions, or the sample processing procedure of an analytical method, revalidation may be necessary. The decision regarding which parameters require revalidation should be based on logical consideration of the specific validation parameters likely to be affected by the change.

For example, changes to extraction or back-extraction media may be expected to affect selectivity, recovery, precision and accuracy, without affecting freeze/thaw stability in the biological matrix. A change to the analytical column or mobile phase may be expected to affect linearity and selectivity without affecting recovery. A guideline to revalidation is presented in Table 1.

Cross-validation

Linearity, intra-batch precision and accuracy

Cross validation refers to applying a validated method in a given biological matrix to the same type of matrix from another species, or to a similar matrix (e.g. plasma and serum) from the same species, or to the same matrix with a change in anticoagulant.

Cross validation may be carried out as follows: a calibration curve is prepared in the validated matrix; replicates at all QC concen-

Table 1
Guidelines for revalidation

Method parameter changed	Parameters to revalidate
Extraction solvent, buffer, back extraction matrix or injection solvent	Linearity, recovery, selectivity, LOQ, intra-batch precision and accuracy, in-process stability. Additionally, if injection solvent is changed, processed sample stability, but no recovery or in-process stability.
Chromatographic column, mobile phase composition (e.g. significant change in retention times), detector type or wavelength	Linearity, selectivity, intra-batch precision and accuracy for the LOQ and other QCs
Extending the upper end, or reducing the lower end of the calibration curve range	Linearity, LOQ (if reduced), intra-batch precision and accuracy at revised upper and lower levels
Internal standard	Selectivity, intra-batch precision and accuracy, and recovery

trations including the LOQ are prepared in both the matrix from the validated method, and in the matrix to be validated; all QC samples are back-calculated from the same calibration curve in the validated matrix; the method is considered cross validated if the determined concentration of QCs in the matrix to be validated satisfy the acceptance criteria.

Stability

All types of stability studies previously discussed (long term, freeze–thaw, in-process, processed sample) should be evaluated and reported under cross-validation.

Selectivity

Selectivity should be evaluated in the revised matrix as interferences may differ between matrices.

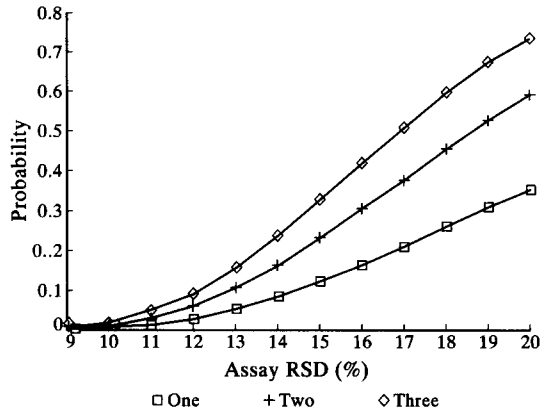

Figure 1
Probability of rejecting an analytical batch when measuring one, two and three analytes.

Analysis of Study Samples and Reporting of Study Data

Multianalyte methods

A recent conference report [1] recommends batch acceptance criteria defined as follows: duplicate QCs at three concentration levels (low, intermediate, high) with four of six calculated to be within ±20% of their respective nominal values (no two at the same concentration level may be outside ±20%). This approach is now generally accepted. In practice, using an analytical method characterized by acceptable inter-batch precision (≤15% RSD) results in minimal rejection of analytical runs on the basis of chance alone. However, rejection of a substantial proportion of runs is possible when using some multianalyte procedures.

In a multianalyte procedure where all analytes are measured in the same processed sample chromatogram, failure to meet QC acceptance criteria for any *one* analyte triggers a repeat analysis of the subject samples in order to generate reportable data for that analyte. The repeat analysis also potentially generates additional reportable data for the other analyte(s). This "compounding" of the acceptance criteria increases the probability of overall run rejection. Treating each control value independently, the probability of run rejection can be calculated for single and multianalyte methods. Figure 1 shows the resulting probability of run rejection vs assay %RSD for one, two and three analyte methods. For a three analyte assay having a RSD of 15% for each analyte, greater than 30% of the runs will have to be repeated on the basis of chance alone. This analysis assumes that variability is constant across analyte concentration levels; any increase in variability at the concentration of the low control would further increase the probability of run rejection.

In multicomponent methods, it is recommended that each component be treated independently with regard to acceptance or rejection, and reanalysis. Data for previously accepted components that is generated simultaneously with a repeat analysis for a failed component should not be evaluated or reported.

Acceptance of a truncated calibration curve

Sometimes, during analysis of study samples, the full concentration range of the calibration curve may not be available. If the low end of the curve is not available, an "elevated" LOQ of concentration equal to the lowest acceptable calibration standard or QC may be used. All study samples with analyte concentrations falling below the elevated LOQ must be reanalysed.

When the upper portion of the calibration curve is not available, a truncated range should be used, by setting the revised upper end of the calibration range equal to the highest acceptable standard or QC sample. All study samples with concentrations determined to be above the revised range must be reanalysed.

Reassay of samples

In most studies, some samples will require reassay. Criteria for identifying these samples, their analysis and reporting should be established *a priori*. Reasons for reanalysing bioanalytical samples can be summarized as follows:

Results are analytically unacceptable. These refer to samples with unreportable concentrations due to:

(i) equipment failure;
(ii) poor chromatography;
(iii) loss during sample processing;
(iv) samples with concentrations below an elevated LOQ;
(v) samples with concentrations above the accepted range of the calibration curve;
(vi) pre-dose samples with observed concentrations needing confirmation;
(vii) processing errors, such as incorrect addition of internal standard or other reagents;
(viii) rejected batches;
(ix) low injection volume (and thus low response).

Samples in these categories should be reanalysed singly, and provided that the batch meets acceptance criteria, the value from the reanalysis in each case is reported.

Pharmacokinetic outliers. It is debatable whether a sample with a concentration incongruous with the pharmacokinetic profile should be reanalysed. However, if a decision is made to repeat the analysis, the reanalysis should be conducted in duplicate.

Reporting a single result from multiple determinations requires pre-defined criteria for data selection. Such criteria are normally based on the percentage difference of the mean of repeat values from the original value. Based on the criteria, the decision is made to report either the original sample value, or a choice between, for example, the mean of duplicate reanalyses, and the median value of the original and reanalyses. It is helpful to use a flow chart to define the decision making process used to report results from multiple determinations, given different scenarios of original and repeat sample concentrations found.

Conclusion

It appears that bioanalytical validation procedures and acceptance criteria will continue to evolve. It is likely, and desirable, that future changes will be based more on statistical considerations and the calculated impact of various criteria on the end use of the data, rather than on "consensus" of leaders in the field. The authors of this paper are conducting research in this area, to better link the impact of validation requirements and acceptance criteria on the determination of bioequivalence and other pharmacokinetic evaluations.

References

[1] V.P. Shah, K.K. Midha, S. Dighe, I.J. McGilveray, J.P. Skelly, A. Yacobi, T. Layloff, C.T. Viswanathan, C.E. Cook, R.D. McDowall, K.A. Pittman and S. Spector, *J. Pharm. Sci.* **81**, 309–312 (1992).
[2] D. Dadgar and M.R. Smith, *Trend Anal. Chem.* **5**, 115–117 (1986).
[3] V.P. Shah, *Clin. Res. Pract. Drug. Reg. Affairs* **5**, 51–60 (1987).
[4] A.R. Buick, M.V. Doig, S.C. Jeal, G.S. Land and R.D. McDowall, *J. Pharm. Biomed. Anal.* **8**, 629–637 (1990).
[5] A.G. Causey, H.M. Hill and L.J. Phillips, *J. Pharm. Biomed. Anal.* **8**, 625–628 (1990).
[6] G.P. Carr and J.C. Wahlich, *J. Pharm. Biomed. Anal.* **8**, 613–618 (1990).
[7] H.T. Karnes, G. Shiu and V.P. Shah, *Pharm. Res.* **8**, 421–426 (1991).
[8] J.W. Hooper, D. Lessard, K. Selinger and D. Dadgar, A new, more reliable method for the determination of the long term frozen stability of drugs in biological fluids. Eighth International AAPS Conference, Orlando, Florida (1993).

[Received for review 28 January 1994; revised manuscript received 20 May 1994]