Accred Qual Assur (1998) 3:189–193

PRACTITIONER'S REPORT

Reports and notes on experiences with quality assurance, validation and accreditation

Robert J. Wells

Validation requirements for chemical methods in quantitative analysis – horses for courses?

Received: 22 September 1997
Accepted: 28 November 1997

R.J. Wells (✉)
Australian Government Analytical Laboratories, 1 Suakin Street, Pymble, NSW 2073, Australia

Abstract Although the validation process necessary to ensure that an analytical method is fit for purpose is universal, the emphasis placed on different aspects of that process will vary according to the end use for which the analytical procedure is designed. It therefore becomes difficult to produce a standard method validation protocol which will be totally applicable to all analytical methods. It is probable that far more than 30% of the methods in routine laboratory use have not been validated to an appropriate level to suit the problem at hand. This situation needs to change and a practical assessment of the degree to which a method requires to be validated is the first step to a reliable and cost effective analytical industry.

Key words Validation · Verification · Veterinary drugs · Anabolic steroids

Introduction and background

All methods used in analytical chemistry are subject to error. Therefore it is vital that each method should be evaluated and tested to ensure that it produces results which make it suitable for the intended purpose. Method validation and verification is the implementation of this evaluation process [1–5]. However, the extent and rigor with which a particular method is evaluated is dependent on the intended use and past experience with the method.

Method validation is the process in which every stage of a new analytical method is subjected to rigorous series of tests to ensure that the method will be able to deliver all the outcomes required of it. The confidence that the method can deliver these outcomes is expressed in terms of statistical probability over the whole analyte concentration range established during the validation process. Verification of a method is a simplified form of the validation process. It involves the testing of a series of method parameters to ensure that a previously validated analytical procedure performs as reported when it is introduced into a new environment where, at the very least, equipment may not be identical to that employed in the initial validation.

Established methods must, as a minimum requirement, be verified when introduced into a laboratory for the first time. Verification, strictly speaking, is also necessary if the method is modified or applied to a new situation, for example a different sample matrix. A new method must be subject to a much more searching series of validation procedures, each one of which adds further confidence in the analytical results obtained. While method validation is mandatory for assurance of analytical quality, the cost to a laboratory is significant. It is therefore important for the financial well-being of a laboratory that validation should adopt no more than those procedures necessary to ensure the analytical quality demanded by a client.

In general, validation processes are universal, but the rigor with which these processes are applied will vary with the intended use of the method. This paper

will outline validation principles and discuss the degree to which these principles must be implemented to achieve the necessary confidence in the analytical result obtained.

Types of method, screening vs confirmation

Methods can be classified in a number of ways, but in the present instance an important distinction should be made between screening and confirmatory methods. Confirmatory methods should include some or all of the parameters shown in Table 1. Screening methods require a limited sub-set of all of the parameters used for method validation, but should include those parameters indicated in the first column of Table 1.

Stages of validation

A distinction can be made between establishing the performance characteristics of a method and complete method validation. After the performance of a method has been assessed, complete validation requires evaluation during its application to the intended purpose, using reference methods for comparison, certified reference materials if available, and interlaboratory comparisons in order to obtain a realistic estimate of the uncertainty of routine results. It therefore follows that, where possible, laboratories should not work alone but collaborate in interlaboratory studies.

Important protocols for method validation in the literature have been derived, amongst others, from the Current Good Manufacturing Practice, Code of Federal Regulations, Food and Drug Administration, National Drug Administration, the United States Pharmacopoeia Convention, the American Public Health Association and the International Conference on Harmonization.

The scheme shown in Table 2 is the *ideal* validation procedure for most methods. However, this is a daunting list, which many laboratories would wish to abbreviate without compromising the required analytical quality.

Table 1

Requirements for both screening and confirmatory methods	Additional requirements for confirmatory method
Specificity (selectivity)	Recovery
Sensitivity	Limit of quantitation (LOQ)
Calibration and linearity	Quality control
Accuracy (bias)	Repeatability (between analysts)
Repeatability of method	Reproducibility (between laboratories)
Range	Ruggedness
Limit of detection (LOD)	System suitability test

Table 2

Stage	Process involved
Preliminary stages	
1. Identify requirements	Needs of client
2. Method selection	Literature search, recommendation of colleagues Novel approach, regulatory requirements Staff expertise and requirements for staff training
3. Development candidate method	Preliminary testing
Method validation	
4. Calibration and linearity	Goodness of fit
5. Accuracy (bias)	Comparison with reference method, spiked samples Certified reference materials, collaborative tests
6. Repeatability	Within day, within lab
7. Reproducibility	Day to day, between labs
8. Sensitivity, LOD and LOQ	Instrumental specifications
9. Specificity	Interference studies
10. Recovery	Samples of known concentration, preparation efficiency
11. Quality control	Spiked samples, blanks, critical control points
12. System suitability test	Routine acceptance criteria, control charting
Validity of method to meet required analytical outputs, including calculation of uncertainty for all relevant stages, is now established	
13. Produce method documentation	Write up experimental/validation work
14. Write user instructions	Prepare standard operating procedure (SOP)
15. Management authorisation	
16. Use method	
17. Method review	Updating SOPSs

Metrological requirements

A cornerstone of the development of a method is that all results are traceable. This has been discussed in detail in an earlier paper in this journal by Price [6]. Current requirements of QA and GLP are that all laboratory equipment should have maintenance and calibration records kept in a log book. Such calibration procedures must be traceable to a reference or primary standard. This is true of both general and specialized

laboratory equipment, but traceability of some specialized laboratory equipment (e.g. mass spectrometers) may be problematical.

A more serious problem arises when the traceability of the whole analytical process is considered. Complete traceability through to the mole is an excellent ideal but nowhere near achievable at present. Comparability between chemical analyses by use of certified reference materials (CRMs) may be regarded as the initial step in the traceability chain to the mole. Even this relatively small first step is constrained by a number of factors. Some of these are:

- A CRM is usually certified using a specific method. Therefore the same method should be used unchanged for all analyses deemed to be traceable to that CRM.
- There are only a limited number of CRMs available.
- CRMs are often only available for a particular matrix at a particular analyte concentration. Any deviation from either sample matrix or a specific concentration would invalidate traceability.

Added to these constraints are the issues raised by the certainty of a result in terms of both analyte identification and quantitation. The uncertainty of a result is dependent on the analyte concentration. In trace analysis it might be argued that the traceability of a method in identifying a substance would have a traceability chain completely different from that of a method which quantitates the same substance. The traceability chain for a method which both quantitates and identifies a substance could be different again. This differentiation is important in many regulatory analyses in which a zero tolerance for a substance has been set. Under these circumstances, only the presence of the substance has to be established. This problem is discussed in more detail later in this paper.

Practical applications of the validation process

Thus far, this paper has simply summarized information covered in standard texts. Two different practical situations are now described in which it is necessary to emphasize certain parts of the validation process to ensure confidence in the results, while minimizing or even ignoring other processes which have no impact on the analytical outcome. In discussing these separate examples, we hope to demonstrate that a pragmatic rather than a strictly rigid approach to validation must be adopted to economically accommodate various needs.

Establishment and regulation of maximum residue limits (MRLs) for veterinary drugs

One of the two Joint WHO/FAO Expert Committees on Food Additives (JECFA) has the task of recommending appropriate MRLs for veterinary drug residues in edible animal tissues based on toxicological assessments of a particular drug coupled with an accurate knowledge of the depletion of the drug and/or its metabolites from various animal tissues over time. These recommendations then go forward to the Codex Alimentarius for ratification and adoption for regulatory purposes. However, the validation requirements for an acceptable method to generate residue depletion data may be distinctly different from the requirements for a satisfactory regulatory method.

A residue depletion method for a veterinary drug

This is usually developed and used in-house (under GLP) by a company for the specific purpose of measuring residues of the drug in tissues. Seldom is the method tested in another laboratory. Since the drug is administered under controlled conditions, interferences due to the presence of related drugs is never a problem. Indeed, related drugs are often used as surrogates or internal standards in such methods. Therefore only matrix interferences are important in the validation process. Analyte recovery and accuracy of quantification are of prime importance, but since no standard reference materials are available, these parameters must be assessed from spiked samples. The LOQ must be as low as possible and the CV of the method must also be low (ideally $< 10\%$ at the LOQ) and certainly lower than the animal-to-animal variability. The complexity (and therefore the cost) of the method is not a vital factor, nor is the suitability of the method for routine application of great importance. A final correction for recovery is necessary in order to establish the actual residue concentrations in tissue. Often a liquid chromatography (LC) method using pre- or post-column derivatization is the method of choice to meet the analytical objective.

A regulatory method for a veterinary drug

This often has distinctly different requirements. Since residue testing is a factor in world trade, it is vital that there is comparability of results between national laboratories undertaking drug residue monitoring. Thus, an interlaboratory trial of the method is very desirable. Furthermore, it is most unlikely that all regulatory laboratories will be identically equipped, and a regulato-

ry method must be robust and flexible enough to cope with such variation. Another major requirement of the regulator, which is of little concern to the drug manufacturer, is the strong preference for an economical multi-residue method suitable for the quantitation of all drugs of a particular class. An optimum method would combine high efficiency isolation and quantitation with the confirmation of a wide range of drugs of different classes from the same matrix. The selectivity of a method is very important to achieve this aim. The LOQs of each drug included in this ideal method would need to be only 4–10 times lower than the MRL, and the recovery and%CV at the LOQ need not be as stringent as those required in residue depletion studies. The use of related drugs of regulatory interest as internal standards or surrogates would be undesirable in this method, and a gas chromatography-mass spectrometry (GC-MS) or LC-MS method using deuterium-labeled internal standards would probably be the most cost effective way of achieving most or all of the analytical goals.

Clearly satisfactory validation of a residue depletion method would be significantly different in many respects from that demanded by a regulatory method. But who should develop the 'regulatory method' and how is it to be decided if a method is suitable for regulatory purposes?

Codex have requested that, during the evaluation process for recommendation of MRLs, JECFA also ensure that the sponsor of the drug provide a "validated method suitable for regulatory purposes". The drug manufacturer might argue that this is not their responsibility. Codex might reply, with some justification, that if a company seeks registration of a drug then the provision of a method to detect and regulate the usage patterns should be mandatory. The development of a new analytical method, particularly a multi-residue method suitable for all edible tissues, is not an inexpensive enterprise, and, not surprisingly, many companies are reluctant to develop a regulatory method which fully meets the desires of the regulators. This impasse has still to be resolved to the acceptance of each party, with JECFA caught in the middle as both arbiter and judge.

Detection of the use of anabolic steroids in sport

The detection of the illegal use of drugs in sport will be a major priority for the Australian Government Analytical Laboratories (AGAL) for the next 3 years. The analyses carried out by the Australian Sports Drug Testing Laboratory in AGAL (NSW) illustrate a further set of analytical problems which test, to the full, the implementation of a rigid validation protocol. The detection of anabolic steroids in urine will be used as an example.

International Olympic Committee requirements for drug testing are simple. If any banned non-endogenous anabolic agent is unambiguously detected then the sample is positive. Since the protocol employed requires any positive sample to be analysed three times and since the criteria used to declare a sample positive are conservative, false positives should be zero. False negatives are not as easy to control since a decision is made from the initial screen. On top of what is essentially a very sensitive qualitative analysis is a requirement to measure the ratio of testosterone to epitestosterone as a method to detect the illegal use of exogenous testosterone. Again this is not quantitative, but accurate measurement of ratios of analytes at very low levels requires extensive validation information.

Steroids are excreted as the glucuronides and sulfates of the steroid and its metabolites. Enzymic hydrolysis yields the free steroids, which are derivatised and determined by gas chromatography-mass spectrometry (GC-MS). Positive identification of two or three different metabolites is needed to declare a positive. However, only a very few deuterated anabolic steroids or their metabolites are available as pure standards. Indeed, few non-isotopically labeled steroid metabolites are commercially available. Therefore a urine prepared by mixing incurred samples of all the known anabolic agents is used with every batch analysed to monitor the complete analytical procedure. Deuterated testosterone and methyltestosterone are added as surrogate and internal standard respectively. Only the concentration of added substances is known. When the day-to-day or even sample-to-sample variability of GC-MS is recognized together with the fact that all urines are of different complexity, the validation process becomes increasingly difficult.

Deuterated testosterone and DES (diethylstilbestrol) glucuronide allow estimation of the success of the glucuronide deconjugation and the recovery of free steroids. Use of the composite urine ensures that derivative separation and sensitivity are within experimental protocols, and deuterated methyltestosterone tests instrumental performance. Although the LOD of added substances can be measured, those of all the metabolites cannot be accurately obtained. Moreover, in real samples, the LOD for each group of metabolites will be different according to the interfering substances which are present. This is why the criteria used to date to declare a sample positive are so conservative - the athlete always gets the benefit of any doubt. The "traditional" validation of the whole process can only be based on the composite urine used as external standard and is centered around the limit of detection. However, acceptance-rejection criteria based on obtaining satisfactory GC-MS spectra for all designated ions of all target

metabolites in the GC-MS run must also be considered as part of the validation process.

The introduction of GC-high resolution mass spectrometry has added another dimension to the problem because much lower signal-to-noise spectra can be obtained and many interferences screened out. This beacons a new era where detection limits will fall and the number of confirmed positives could well rise. Also, will new problems arise in deciding if the method has been adequately validated? That is, is it demonstrably fit for the purpose for which it was designed?

Method validation vs performance criteria validation

The debate about comparability of methods based on performance criteria is still current. The AOAC stand is that results can only be validly compared between laboratories if the same method and preferably the same equipment is used. Although many good arguments can be advanced for this approach, the time and resources required to set up and evaluate interlaboratory tests are often extreme. Also, this approach leads to inertia and reluctance to commit to change or improvement. The obvious argument against performance criteria based validation is the question of what is the bench mark against which the criteria are measured. If results are compared with those obtained for an SRM, then that SRM was probably determined by a standard method. The ideal situation arises when two independent methods give comparable results. This certainly happens – sometimes.

Conclusions

This debate reinforces the need for a rigorous system for traceability in analytical chemistry. The time must come when clients can be confident in the fitness of their results for the purpose for which they were obtained. It has been estimated that at least 30% of all chemical analysis is not fit for purpose. In Australia alone, this probably represents $A 100 000 000 per annum spent on analytical work which is worthless.

However, it is probable that far more than 30% of the methods in routine laboratory use have not been validated to an appropriate level to suit the problem at hand. This situation needs to change and a practical assessment of the degree to which a method requires to be validated is the first step to a reliable and cost-effective analytical industry.

References

1. ISO/IEC Guide 25 (draft 5 1996] General Requirements for the Competence of Testing and Calibration Laboratories. International Organisation for Standards, Geneva
2. Taylor JK (1993) Anal Chem 55:600A–608A
3. Green JM (1996) Anal Chem 68:305A–309A
4. Thomson M, Wood R (1995) Pure Appl Chem 87:649–666
5. Guideline for collaborative study of procedure to validate characteristics of a method of analysis (1989) J Assoc Off Anal Chem 72:694–704
6. Price G (1996) Accred Qual Assur 1:57–66