

Towards a consensus on validation of bioanalysis methods for large molecules

A recent international conference in London considered progress towards regulatory guidance for validation of analytical methods for the quantitation of macromolecules in biological fluids, for assessment of antibodies and for validation of biomarker assays. **Howard Hill** reports

Regulatory guidance exists for validation of bioanalytical methods for small molecules. It includes the US Food and Drug Administration's "Guidance for industry bioanalytical methods validation", published in May 2001. But agreement has yet to be reached for large molecules. In March 2000, the FDA and the American Association of Pharmaceutical Scientists held a joint meeting on bioanalytical methods for macromolecules, and since then the AAPS has led attempts to provide a consensus for the validation of ligand-based assays, cell based assays, antidrug antibodies and biomarkers. Bioval 2004 formed part of this consensus process.

Quantitation of macromolecules

The first day of Bioval 2004 was dedicated to the quantitation of macromolecules in biological fluids and the assessment of antidrug antibodies. The first speaker, Russ Weiner (Bristol-Myers Squibb, US) compared and contrasted the needs for small and large molecular guidance. The differences and degrees of increased complexity associated with ligand binding assays include the complexity of the standard curve, the variability and complexity of the standard itself, reagent variability, and matrix effects further confounded by the difficulty in sample clean-up.

Because ligand binding assays are indirect assays dependent upon binding interactions, factors such as lipaemic and haemolysed samples, binding proteins and anticoagulants that interfere with this process will destabilise the assay, said Dr Weiner.

These issues were further developed by subsequent speakers. Binodh DeSilva (Amgen, US), discussing ligand binding assays for macromolecules in support of pharmacokinetic evaluation, said that hot topics include differences between the criteria for acceptability for a valid method and those for accepting runs or batches when in routine use. In addition, she identified the impact of a range of variables involved in the assay, such as reagents, antibody specificity,

analyte stability, plate coating, matrix variables, and sample collection conditions.

Dr DeSilva also highlighted two important issues as problems associated with choosing the correct calibration model, and the application of total error to batch acceptability. These were then reviewed in detail by the following two speakers.

Bruno Boulanger (Lilly, Belgium) discussed the statistics around the total error concept, that is, the definition of accuracy being an additive function of bias and the random error component, defined as precision. This leads to the proposed "4:6:30" rule, where four out of six QC samples are expected to fall within 30 per cent of nominal for a batch to be acceptable, said Dr Boulanger, and it assumes that the combination of random error (precision) and systematic error (bias) for the assay is 30 per cent or less.

John Little (CentraLabS, UK) discussed the challenge to obtain the right standard curve for ligand binding assays. It was generally accepted that logistic regression methods are the most robust for ligand assays.

Dr Little recommended curve fitting using four-parameter logistics plots for limited reagent immunoassays such as enzyme immunoassays and radioimmunoassays. He considered that five-parameter logistics plots are more appropriate for excess reagent immunoassays, such as immunoenzymic assays, enzyme-linked immunosorbent assays (ELISA) and immuno-radiometric assays. In addition, because the assay response error changes in a non-uniform manner, it is essential to weight the curve to compensate for this non-uniformity.

Dr Little finished on the contentious issue of "editing" points from the curve that fall outside a set limit, such as 15 per cent. He showed how "choice" of editing of the same standard curve could result in different accuracy and precision of derived data. He posed the question, "What rational criteria can be derived for logical unbiased editing of standard curves?" The discussion failed to resolve the issue.

Immune system

Mark Wing (CentraLabS, UK) set the scene for the afternoon session by describing the complexity of the immune system, an important feature of which is its ability to distinguish itself from foreign material. In characterising the immune response to

drugs, antidrug antibodies are usually measured. Immune responses may be described as binding, enhancing or neutralising, having no effect, increasing or decreasing bioavailability, respectively. A further consequence of neutralising antibodies, said Dr Wing, is the potential to see the endogenous product as "foreign", with potentially serious consequences. These issues formed the basis for the rest of the afternoon's discussions.

Arno Kromminga (Institute for Immunology, Pathology and Molecular Biology, Hamburg, Germany) discussed the development of assays for detecting autoantibodies in autoimmune disease. A major concern is that the assay developed is too sensitive and may recognise autoantibodies of low affinity, with no clinical significance.

Dr Kromminga said that administration of macromolecular therapeutic agents leads to the production of neutralising antibodies but the route of administration has a significant effect on the levels. For example, more antibodies are produced by giving interferon-beta by the subcutaneous route than by the intramuscular route.

Geoff Hale (Oxford University and BioAnalab) described attempts to circumvent the production of antidrug antibodies by "hiding" macromolecule drugs from the immune system, using agents such as PEG. Another approach is to engineer the drug — the example was the antibody Campath (alemtuzumab) — so that the body tolerates its presence. This stealth approach may allow for repeat, chronic administration. Professor Hale reviewed the development of antidrug antibody assays with particular emphasis on the problems of obtaining positive controls, determining, and then establishing, a quantitation limit.

Validation of biomarker assays

The topic for the morning session of the second day was biomarker assays. David Perrett (Barts and The London Hospital) described how clinical biochemists had struggled for years with problems of assaying biomarkers as diagnostic aids in the clinical environment. Now that the pharmaceutical industry had discovered their use in drug development, consideration was being given to appropriate assay acceptance criteria that would satisfy regulatory authorities, Professor Perrett said.

John Allinson (Bioanalytical Systems, UK) discussed the issues of assays that are

Details Bioval 2004 was organised by the **Joint Pharmaceutical Analysis Group** (a joint venture of the Royal Pharmaceutical Society and the Royal Society of Chemistry), and was held at the Society's London headquarters on 12 and 13 February. Dr Hill is chairman of the meeting's organising committee

deemed acceptable for diagnostic purposes. Many assay kits are validated to the US FDA 510K guidelines and are run in laboratories regulated by the Clinical Laboratory Improvement Amendments and which have College of American Pathologists accreditation. Proficiency tests are run under guidelines defined by the US National Committee for Clinical Laboratory Standards. Research-based assay kits are not used for diagnostic purposes but can be validated to some extent for use in research biomarker assays.

Ian James (Pfizer, UK) spoke about the role of biomarkers in the early clinical phase of drug development. He illustrated their application, from the use of a broad range of biomarkers at the early stage in drug development process, to a limited panel of validated biomarker assays in the later stages.

Dr James stressed the need to ensure that the biomarkers are clinically relevant because the cost of validating and running these assays can add significantly to the development costs.

Jean Lee (MDS Pharma Services, US) described how the AAPS ligand binding assay bioanalytical focus group is working towards a consensus approach on biomarker assay validation in biological samples for drug development. She summarised the outcomes of the focus group workshops held in Salt Lake City in October 2004.

Currently, said Dr Lee, the seminal reference is the paper on "Validation of bioanalytical assays for novel biomarkers: practical recommendations for clinical investigation of new drug entities" (in Bloom JC, Dean RA, eds. Biomarkers in clinical drug development. New York: Marcel Dekker; 2003. pp119–48).

Some of the problems arise from a lack of "official" primary or secondary standards, and the inability to obtain an "analyte free" biological matrix for the preparation of standards. The assay dynamic range is often different from the concentrations found in the test species.

The aim of most clinical studies involving biomarkers is to compare dosed and control populations. Relative, rather than absolute, changes in biomarker values may be all that is required to meet the study objective. Therefore it can be argued that assay precision and relative quantitation, rather than absolute quantitation, would suffice. The approach used in the FDA-BMV guidance is recommended for the validation of biomarker assays.

Some of the validation issues were discussed by Ron Bowsher (Linco Diagnostics, US) in his talk on challenges in validating test kits for quantification of biomarkers. Dr Bowsher described how the lack of regulatory guidance has led to confusion among bioanalytical scientists about what procedures are necessary and appropriate for validation of biomarker assays, especially kit assays. He classified assay types using analytical performance

characteristics that differentiate the assay types.

Definitive quantitative assays can be validated to the FDA-BMV guidance and use well characterised reference standards for calibration and give a response proportional to concentration that is continuous over the calibration range. Relative quantitative assays are the same as definitive quantitative assays in all respects except that the assays use reference standards that are uncharacterised. "Quasi-quantitative" assays have no definitive standards and the standard curve units may be based on activity units.

Qualitative assays may have no accuracy or precision criteria and are based purely on ordinal data (high, medium or low nominal data) where results are either positive or negative. Currently, only methods having well characterised reference standards meet the full requirements for the FDA-BMV guidance. This can be difficult to achieve for many biomarkers.

New techniques

The meeting's final session looked to the future and considered the development of alternative techniques. Robin Thorpe (National Institute for Biological Sciences and Control, UK) discussed the role of *ex vivo* bioassays in evaluating the biological activity of "analytes" in a living system. This differentiates them from *in vitro* based immunoassay and receptor binding assays.

Dr Thorpe described a range of bioassays, many of which are cell-based systems, as opposed to whole animal *in vivo* methods. The diversity and complexity of such assays was exemplified by more than 14 observed effects of interleukin-1 on different target cells and tissues. The effects ranged from fever induction at the microgram level, through the release of collagenase from the synovium at nanogram level, to interleukin-2 receptor induction on T-cells at the femtogram level.

Other issues ranged from standardisation of conditions to standardisation of the "test-substance" since the purity — or activity — may vary, depending upon the source and tests used to characterise it.

Berend Oosterhuis, (Pharma-Bioresearch, Netherlands) described a case history of clinical application of an *ex vivo* stimulation assay as a biomarker for anti-inflammatory effects. Using lipopolysaccharide stimulation of whole blood (monocytes), he reviewed the impact of different drug inhibitors of P38 kinase, ICE and CD14 antibody on cytokine release.

Dr Oosterhuis said that by careful optimisation of the many parameters, it is possible to achieve coefficients of variation of 15–30 per cent. Such characteristics mean that differences in the inhibition of tissue necrosing factor- α can be used to differentiate different dose levels and therefore develop useful pharmacokinetic/pharmacodynamic relationships.

Brian Shenton (University of Newcastle, UK) provided an historical overview of the

development of flow cytometry, including fluorescent activated cell sorter (FACS), which sorts cells on the basis of physical parameter and fluorescent tags. The development of a wide range of flow cytometers has been accelerated by a worldwide screening programme for AIDS, the detection of which depended upon destruction of the T-helper lymphocyte population during viral infection.

Currently, flow cytometers are used in the clinical and research environment as a routine tool for biomarker assays such as cytokine analysis and lymphocyte phenotyping. It is essential to ensure comparability of results and these depend upon well controlled quality control and quality assurance (QA) procedures.

Dr Shenton highlighted the fact that many QC procedures using dyes and beads are instrument-specific, making comparability of results between laboratories difficult. Correct instrument set-up (training and qualified operators) is essential before invoking the appropriate QC checks.

Ulrich Kunz (Boehringer Ingelheim, Germany) described the advantages and limitations of Biacore methods for drug level and immune response measurements of therapeutic proteins. Limitations of the technique include matrix effects that limit the number of samples or cycles per sensor chip, long-term stability of the chip ligand and stability of samples in the autosampler.

Although pure "analyte" ELISA bridge assays are 100 times more sensitive than Biacore assays, in real life applications the Biacore system can detect the immune response as little as three weeks after dosing whereas ELISA can only detect it after eight weeks.

Dr Kunz hypothesised that the reasons could be the access to all epitopes by Biacore, and the ability to detect low affinity antibodies. Thus, Biacore has the advantage of early detection, with the disadvantage of expensive instruments and chips, and low throughput.

The final talk, by Geoff Hale, focused on the application and validation of flow cytometry to the measurement of the pharmacokinetics of Campath-1, a humanised IgG1 monoclonal antibody, against the CD52 antigen of human lymphocytes. Because the humanised drug closely resembles endogenous human IgG (which is present at concentrations 100,000 times that of the drug), measurement is complex.

After evaluating a wide range of ligand binding assays and bioassays it was decided to use flow cytometry. Using this technique they were able to validate an assay from 0.5mg/ml to 50mg/ml. When the critical variables are identified and controlled, meaningful data can be generated.

Professor Hale illustrated the application of this assay and showed a correlation between trough concentrations of Campath in leukaemia patients in early therapy and their ultimate clinical response.