Validation of Immunoassays for Biomarker Detection at Sequani.

Brenda A Finney, PhD – Principal Scientist, Biological Sciences, Sequani Limited

This paper is intended to review the literature on immunoassays used for biomarker analysis and detail the route we take at Sequani for validation. It is intended as a guide where analysis kits are purchased from commercial suppliers, when their intended use is to determine the concentration of endogenous analytes and to inform as to toxicological profile or mechanism of action.

Introduction

Simply put, a biomarker is a substance that can be measured in body fluids or tissues which gives information about biological processes. This encompasses a wide range of analyses, including the standard clinical pathology assays, but the specific term “biomarker” has come to be associated with advanced analysis of protein based indicators of disease, toxicity or efficacy.

Biomarkers can be an important addition to investigations of toxicology. By using validated techniques, biomarkers can indicate toxicity with quicker and more sensitive endpoints than traditional toxicological methods. Their use can aid in the selection of candidate items as well as reduce the cost and time required for pre-clinical safety testing. Furthermore, they can assist in smoothing translation to clinical trials.

The main format for the detection of biomarkers at Sequani is the immunoassay. They are extremely useful for the detection of immunogenic responses, hormones and proteins (endogenous or pathological) in various matrices like serum, plasma, urine or digested tissue. There are different formats for immunoassays, but the basic principle of using an antibody to detect the protein of interest and then determine the concentration using additional antibodies and/or colour substrates is universal (See Figure 1).

In order to ensure that an assay is accurate and reproducible it must be validated. Method validation is the assessment of the assay for its performance characteristics where the assay routinely gives reproducible, accurate results and should not be confused with biomarker qualification. Qualification is the process by which a potential biomarker is correlated to a specific biological process or clinical endpoint. Additionally, it should be noted that biomarker method validation is separate from ligand-binding assay validation intended for pharmacokinetic (PK) or toxicokinetic (TK) analysis which must adhere to the guidance provided by the relevant regulatory authorities. It is also separate from routine laboratory validation for samples being analysed clinically for diagnosis or treatment.
Validation of Immunoassays for Biomarker Detection at Sequani.

Using our standard operating procedures for sample collection and storage, along with our controlled storage systems we ensure the highest quality samples for analysis.

Therefore, unless the analysis is intended to produce clinical information or a PK/TK profile for a biologic drug or large molecule, there are no specific guidelines on the necessary validation parameters for an assay. This paper is intended to review the literature on immunoassays used for biomarker analysis and detail the route we take at Sequani for validation. It is intended as a guide where analysis kits are purchased from commercial suppliers, when their intended use is to determine the concentration of endogenous analytes and to inform as to toxicological profile or mechanism of action.

Guidance

Our validation strategy is loosely based on bioanalysis guidelines from several agencies including the FDA\(^3\) and EMA\(^5\), while also taking into account recommendations present in the references cited, and information discussed at industry conferences and working groups. We advocate a strategy where the assay validation fits the type of study where samples will be generated, and takes into account the relative importance of, and intended use of the biomarker data in that study.

Initial Considerations

As we are focusing on commercially developed assays, there are some pre-analytical parameters which are not assessed during validation of an assay. We review published data on the analyte as well as consider the manufacturer’s recommendations for sample collection and the matrix to be tested. Multiple types of anticoagulants will only be tested where it is requested, or if there is no information available as to the best matrix for the assay (or where the client requires multiple matrices).

Through the use of study plans and our site-wide SOPs, sample collection is standardised and integrity is maintained throughout collection, storage and analysis.

For the majority of assays, we use the manufacturer’s recommended calibration curve model selection and do not investigate alternative curve fitting algorithms.

Additionally, we do not reassess the reagent stability where expiry information is provided by the manufacturer. All components are used within their stated expiry limits.

Reference material to be used as quality control samples or as a spike for the various aspects of validation will vary depending on the assay. In few cases will native material be available. Therefore it is important to review the potential reference materials available, evaluate them for their suitability with the assay and their relationship with the endogenous analyte of interest.

Acceptance Criteria and Method Production

The acceptance criteria for any assay validation will be stated in the validation study plan. However, each assay is evaluated on a case-by-case basis and acceptance criteria set based on what is known of the assay’s variability and the objective of the study from which samples are being generated\(^6\). This means that on occasion, the acceptance criteria for an assay will be updated during the course of the validation to take into account characteristics unique to that assay. This change will be conducted in a fashion appropriate for the stage of validation and will be controlled, particularly where the study is being conducted to GLP standards. Where these changes need to follow through to study sample analysis, the information will also be present in the method produced for that assay.

In general, our acceptance criteria for quality control samples follows the 4-6-X rule, where 4 out of 6 QC samples must be within the acceptance criteria of X for that plate. Our current standards are detailed in our SOPs.

Each assay validated will be associated with a method which will have its own unique identifier like, BMK001. This is then followed by a designation which will indicate the method format such as EL, for ELISA, IM for Immulite Immunoassay (Figure 2, Siemens assay system) or FC when the immunoassay is conducted using the Flow Cytometer (BD Biosciences FACSVers).
Validation of Immunoassays for Biomarker Detection at Sequani.

The rigor with which the assay is validated is dependent upon the importance of the biomarker data and should increase where the data are to be used for critical clinical or business decisions 2.

A copy of the method for each assay used in a study will be present within the study data, so that it may be referred to as needed. The method generated is a controlled document such that changes result in a new version of the document and are detailed in the revision history. Each method produced by our staff is electronically signed by the scientist responsible, then reviewed and electronically signed by a second scientist to ensure accuracy and clarity for use.

Stages of Validation

We consider that the majority of assays which undergo validation are “relative quantitative” assays. Relative quantitative assays depend “upon a response-concentration calibration function”, where reference standards are “not fully characterised, or fully representative of an endogenous biomarker” 2. The rigor with which the assay is validated is dependent upon the importance of the biomarker data and should increase where the data are to be used for critical clinical or business decisions 2. The sections below describe our step-wise approach to validation.

Feasibility (F)

At the beginning of the validation process it is important to have an idea of whether a kit is suitable and functions in the expected manner. In order to do this, a feasibility investigation is performed using kits from at least two different suppliers (where available). This enables a comparison between the assays in terms of basic function, range of detection, ease of use and length of assay. It also enables initial analysis of potential quality control (QC) material (which could be reference material, recombinant or purified analyte) as this can vary from assay to assay and analyte to analyte. In some instances, the limits of the assay will not be tested during the feasibility stage, but the manufacturer’s stated limits will be used 3. This is generally the case where there is a need for testing multiple sources of quality control materials.

Fit for Purpose (FFP)

This is a validation which is based on a minimum of three separate analytical runs. It does not generally include an assessment of analyte stability or Quality Assurance (QA) oversight. Stability will be set at that recommended by the kit manufacturer and other published data, where available. Although, if agreed at the outset, limited stability can be tested. This validation will result in the production of the assay method and a summary report.

Comprehensive Validation (ComV)

A comprehensive validation is an extended assessment of the assay with additional assay runs as well as robustness tests and assessment of the stability of the analyte in the analysis matrix for up to one month. Our normal practise is for this level of validation to be subject to QA oversight and review so that GLP compliance can be claimed. This level of validation is recommended for several reasons which include: limited information available on the assay or analyte, feasibility testing has not been conducted to determine the best assay, stability information on the analyte is limited, the most suitable matrix for the assay needs to be determined, selectivity investigations need to be conducted, or GLP compliance is required/desired. This validation is conducted over at least five separate analytical runs. It will result in production of the assay method and a full report.
Validation of Immunoassays for Biomarker Detection at Sequani.

Parameters Tested in Validation

Table 1 describes the parameters which can be tested in an assay validation, and the validation stage where they are generally performed. Information regarding our general acceptance criteria are available in our outline protocols. At the start of each study, a study plan will be produced detailing the parameters to be tested, with methods and acceptance criteria. The acceptance criteria for a specific assay may be different than those listed in the outline protocol and will be set as is relevant for the assay being tested.

Table 1. Validation Parameters.

<table>
<thead>
<tr>
<th>Validation Stage</th>
<th>Parameter</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>F (dependent upon assay) FFP ComV</td>
<td>Assay Range</td>
<td>This will constitute the reportable range of the assay and spans from the Lower Limit of Quantitation (LLOQ) to the upper limit of quantitation (ULOQ). Within this range, the analyte is measurable with acceptable levels of accuracy, precision and error.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>i. Limit of Blank (LOB) will be determined where possible from samples with no analyte present. Lower Limit of Detection (LLOD) will also be determined from samples with no analyte or low concentrations (spiked or endogenous) of the analyte.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. LLOQ will be determined using the samples for LLOD where these samples meet bias and imprecision criteria (typically CV ≤ 20%).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii. The calculations used to determine these values have been taken from Christenson and Duh 2012 and Armbruster and Pry 2008.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv. The ULOQ will be set using the highest calibration standard (see below).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>v. These investigations are not an exhaustive investigation of the absolute limits of the assay, rather a confirmation of the manufacturer’s information and establishment of our in-house reporting limits.</td>
</tr>
<tr>
<td>F FFP ComV</td>
<td>Calibration Curve</td>
<td>Calibration curves are run in duplicate on each plate of the assay. They are evaluated on each run using criteria detailed in our SOP to ensure that they have acceptable accuracy and precision. It is important that curves contain sufficient points to accurately plot the correct non-linear fit, therefore kits with 6-8 non-zero concentrations present for the calibration curve will be used preferentially for analysis. The calibration curve fit recommended by the manufacturer should be evaluated over at least the initial 3 runs to confirm that it is suitable. Where calibration curves are inconsistent across plates, this may indicate a wider issue with the assay and care should be taken when deciding to progress with the use of the assay.</td>
</tr>
</tbody>
</table>
### Validation Stage Parameter Information

<table>
<thead>
<tr>
<th>Validation Stage</th>
<th>Parameter</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Accuracy</td>
<td>Recovery assessment is used as an indication of the relative accuracy of the assay(^2). This is conducted by adding the analyte (QC - as described above) to the intended matrix (spiking) in multiple concentrations which span the range of the assay and includes assessment of the endogenous concentrations in the intended matrix which are then subtracted from the spiked sample results.</td>
</tr>
<tr>
<td>F</td>
<td>Precision</td>
<td>Precision assessment will provide information on how the assay functions over replicates, and over multiple plates/assay runs. If an assay does not function in our hands with a reasonable amount of precision, the data obtained on study samples may be nonsensical. Within run precision (replicates) - F, FFP and ComV Between run precision (multiple plates/runs) – FFP (2 runs) and ComV (at least 3 runs)</td>
</tr>
<tr>
<td>F</td>
<td>Quality Control</td>
<td>Reference material used for quality control samples should reflect the endogenous analyte. However this is not always possible. Therefore, multiple options may be tested in the feasibility stage of validation before choosing the best one which will progress for use on the rest of the validation. Ranges will then be established as needed throughout the remainder of the validation process.</td>
</tr>
<tr>
<td>FFP (dependent on assay)</td>
<td>Dilutional Linearity</td>
<td>There has been much discussion about the utility of testing dilutional linearity. Dilutional linearity can be problematic as the affinity or binding of the antibody to the analyte can vary in different matrices(^6). It is generally thought that parallelism is a better indicator of assay function(^2,6) and(^9). However, where the reference material used for the validation is reflective of the endogenous biomarker, dilutional linearity can be a further indicator of accuracy and demonstrate if there is a high dose hook effect and support parallelism investigations(^5). Once the reference material/quality control has been assessed for its utility with the assay, dilutional linearity may be assessed by preparing a sample which is 5 times above the ULOQ and then diluted into the range of the assay. Alternatively, where samples are heavily diluted prior to use in the assay (10-fold or more), and expected endogenous concentrations allow, a single additional dilution step may be tested for precision.</td>
</tr>
<tr>
<td>ComV (dependent on assay)</td>
<td>Parallelism</td>
<td>Assessed if native samples available for validation, however since this is rarely the case, this investigation is generally performed during study analysis.</td>
</tr>
</tbody>
</table>
### Validation Stage Parameter Information

<table>
<thead>
<tr>
<th>Validation Stage</th>
<th>Parameter</th>
<th>Information</th>
</tr>
</thead>
</table>
| **FFP (dependent on assay)**, ComV | Stability | Stability is assessed after storage on a subset of validation samples at a relevant storage temperature. This storage temperature, refrigerated ($2-8^\circ C$) or frozen ($\leq -15^\circ C$ or $\leq -70^\circ C$) is initially based on any published information available for the analyte of interest. Stability investigations generally cover at least 3 of the categories listed:  
  i. Short-term – storage for up to 7 days  
  ii. Freeze-thaw – stored frozen at least overnight and then thawed, this process can be repeated up to three times  
  iii. Long-term – storage for up to 1 month  
  iv. Processed – storage on the bench at room temperature for up to 24 hours, post initial analysis (may be omitted in cases were sample volumes make this scenario unlikely)  
  v. Extended – storage for more than 1 month, should be tested where extended storage of samples is likely or necessary |
| **FFP**, ComV | Carry-over | Carry-over is assessed to ensure that there is no transfer of sample/reagent from well to well on an assay plate. This is tested by placing low concentration samples after high concentration samples. Inflation of the resulting concentration in low concentration samples would indicate carry-over has occurred. Where carry-over is indicated above the acceptable limits, samples should be arranged in a manner to reduce the impact of any carry-over on sample concentrations. |
| ComV | Robustness | Robustness reflects the ability of the assay to produce consistent results in the face of small changes to non-critical points of the procedure. For example, automation (Figure 3) vs. manual, analyst vs. analyst or small changes to incubation times, or wash volumes.  
  The tests conducted will vary by assay, not all kits have suitable volumes of reagents for automation for instance. When these tests are performed the results of similar samples on different plates will be compared. |
| ComV (dependent on assay) | Specificity | Specificity is the ability of assay reagents to distinguish between the analyte and other structurally similar components present in the matrix or non-specific matrix effects $^2$, $^6$ and $^9$ and will generally have been investigated during method development.  
  Generally, if present, reduction of these components by dilution in an appropriate buffer will overcome issues of non-specificity $^4$. Be aware that where kits state a minimum required dilution, this could be an indication that the desired matrix may interfere with the specificity of the assay. It may not be useful to test specificity where the QC/reference material is not reflective of the native analyte and the need for this test should be evaluated on an assay-by-assay basis.  
  When conducted, specificity is tested by spiking different batches of sample matrix near the LLOQ and determining the accuracy of the reported result. |
When assays are calibrated with a recombinant material it is important that the recombinant material parallels the native species in terms of recognition. Antibody and ligand-binding affinities may be different to surrogate matrices or analyte. Therefore, when this type of material is used, as will be the case in most validation studies, parallelism should be tested at the first opportunity where native samples are available.

Additional Considerations

This section describes additional investigations which may be required for validation. These aspects may need to be added to the validation if indicated by results during the performance of a planned validation, or included in the validation study plan from the outset where published information indicates that they will be useful or needed.

1. Selectivity is the ability of the method to determine the analyte in the presence of other components present in the sample.
   a. Additional specificity/selectivity investigations may be required where matrix from the relevant population, either diseased/transgenic/treated, is thought to potentially contain factors which would interfere with the assay.
   b. Potential sources of the interference should be considered and if possible, assessment of the accuracy in replicates and multiple sources from relevant populations should be tested. It may not be possible to obtain these samples during the validation, but if necessary this investigation should be done as soon as samples are available.

2. Certain assays may be more sensitive to the order of sample placement on the plate.
   a. Edge effects can be seen when there are inconsistencies between the wells on the edge of the plate as opposed to wells in the centre. This can generally be overcome by ensuring that temperatures are consistent across the plate and plate sealers are used during incubation steps.
   b. However, investigation into the best placement of samples may be required if edge effects are indicated despite these control measures.

Continued Validation

As indicated in the table, there are some investigations which may not be reasonable to complete during a validation and are therefore required on incurred samples during routine analysis. When these investigations have been completed during routine analysis, the method will be updated to include the results.

1. When assays are calibrated with a recombinant material it is important that the recombinant material parallels the native species in terms of recognition. Antibody and ligand-binding affinities may be different to surrogate matrices or analyte than to native matrices or analyte. Therefore, parallelism is the conditions in which dilution of test samples do not result in a biased measurement of the analyte concentration. Therefore, by diluting several samples in analyte-free medium and determining the analyte concentrations; parallelism is demonstrated if the dilution curves of these samples are parallel to the standard curve, with no apparent trend for increasing or decreasing estimates of analyte concentrations.
   a. Parallelism is the conditions in which dilution of test samples do not result in a biased measurement of the analyte concentration. Therefore, by diluting several samples in analyte-free medium and determining the analyte concentrations; parallelism is demonstrated if the dilution curves of these samples are parallel to the standard curve, with no apparent trend for increasing or decreasing estimates of analyte concentrations.
   b. The precision of the diluted samples is typically acceptable if the coefficient of variation of the corrected means is ≤30%.
   c. Parallelism should be tested in a subset of samples where those samples show sufficient initial concentrations of the analyte for dilution.
   d. Where parallelism is not achieved, the use of the calibration standards provided by the kit is invalidated, and therefore only quasi-quantitative interpretation of the results is possible. It is not anticipated that this would be a common result when using commercially available kits, however some degree of non-parallelism can be tolerated.
   i. If this occurs, study plan amendment may be required, as will updates to the method on the way in which the data are reported. For example, data could be reported as a measure of absorbance change from the Control group rather than the units indicated by the calibration curve.

2. The use of historic control data can be useful, not only as a baseline for determining effects on study animals, but also as an indication that the assay is functioning properly. If possible, samples from Control animals of the correct strain, sex and age for your intended study should be assayed during the validation to give an indication of the concentrations which can be expected. This also allows for data import procedures to be tailored to the assay using the most relevant data available.
   a. Understanding the concentrations normally obtained by our lab, in animals from our site of similar age, strain and sex can provide additional information on the assay function which takes into account the physiological/biological variation within the population of interest.
   b. Physiological variation could potentially obscure differential results between groups if the assay sensitivity is not appropriate. Therefore, the use of the control database to produce intra- and inter-subject variation, or subject variation across serial samples (if available) may be helpful and give an insight into the level of variation which can be expected in normal subjects.
   c. Data from Control groups on studies may be added to the control database where those plates have passed within stated acceptance criteria. Where these data are produced they will be held centrally within our historical control database and referenced as needed for reporting.
Conclusion

We recommend a step-wise approach to validating assays for your biomarker of interest. This allows you to save time and money while achieving confidence in the assay generating your results. We will also support you through a full-scale validation with continued in-study validation where it is needed for regulatory submissions. Our experts are happy to discuss your needs at any stage of your programme, but in order to get the best assay and results we recommend we start talking as soon as a biomarker is thought to be relevant for your programme.

References


(4) ICH harmonised tripartite guideline. Validation of analytical procedures: Text and methodology Q2(R1); November 2005.

(5) EMA. Guideline on bioanalytical method validation. February 2012.


(9) Cummings J, Raynaud F, Jones L, Sugar R and Dive C. Fit-for-purpose biomarker method validation for application in clinical trials of anticancer drugs. Br J Cancer 2010 103: 1313-1317