Over the past decade, the drug development landscape has shifted to one where much of the early phase drug discovery and development work is conducted at small/virtual/mid-size biotech companies. Many of these companies have turned to CROs for bioanalytical analysis. The use of CROs is beneficial for balancing budgets with timelines, modulating the risk of drug candidate failure, and meeting regulatory agency expectations. There is a clear expectation that during late stage drug development supporting GLP Toxicology or human clinical studies require full assay validation of the bioanalytical methods used, as per FDA guidelines. It is less clear for the level of rigor that should precede this stage of drug development. While assay validation is not required for early stage drug development, situations such as go/no-go decisions or when bioanalytical data are not a critical endpoint often leave sponsors wondering what level of data scrutiny will meet the needs of the study. An open conversation between the sponsor and CRO/contract lab is crucial to ensure the bioanalytical method rigor is suitable for the intended use of the data.

This article assists in determining the appropriate level of method performance and evaluation required for small molecule and immunogenicity bioanalysis for each phase of development and anticipated data use.

**FIT-FOR-PURPOSE**

**Assay Development in Bioanalysis**

The success of bioanalytical studies relies on the selection of the most suitable analytical method but the timeline of method development and types of analyses involved vary greatly.

Over the past decade, the drug development landscape has shifted to one where much of the early phase drug discovery and development work is conducted at small/virtual/mid-size biotech companies. Many of these companies have turned to CROs for bioanalytical analysis. The use of CROs is beneficial for balancing budgets with timelines, modulating the risk of drug candidate failure, and meeting regulatory agency expectations. There is a clear expectation that during late stage drug development supporting GLP Toxicology or human clinical studies require full assay validation of the bioanalytical methods used, as per FDA guidelines. It is less clear for the level of rigor that should precede this stage of drug development. While assay validation is not required for early stage drug development, situations such as go/no-go decisions or when bioanalytical data are not a critical endpoint often leave sponsors wondering what level of data scrutiny will meet the needs of the study. An open conversation between the sponsor and CRO/contract lab is crucial to ensure the bioanalytical method rigor is suitable for the intended use of the data.

This article assists in determining the appropriate level of method performance and evaluation required for small molecule and immunogenicity bioanalysis for each phase of development and anticipated data use.

**SMALL MOLECULE (LC-MS/MS) PERSPECTIVE**

The characterization criteria and ruggedness required of bioanalytical assays has incrementally increased over the past 25 years. In 1990, the first annual American Association for Pharmaceutical Scientists (AAPS) and US Food and Drug Administration (FDA) Crystal City Meeting ushered in the first wave of standardized
characterization. The subsequent 2000 Crystal City II Meeting sought to harmonize bioanalytical method validation in late stage clinical and/or bioequivalence studies where pharmacokinetics (PK) is the critical endpoint and served as the basis for the 2001 US FDA Guidance on Bioanalytical Method Validation. The guidance and subsequent FDA white papers were critical in outlining the determination of robust and reliable PK data and therapeutic endpoints. Due to the onerous time and cost for increased level of characterization as well as for less critical decisions, the 2006 Crystal City III meeting attendees and subsequent white papers briefly discussed a “fit-for-purpose” approach.

Subsequently, this “fit-for-purpose” approach to bioanalysis has been discussed in various forums, such as AAPS Journal, 2014 (S. Lowes, et al), the Workshop on Recent Issues in Bioanalysis (WRIB), Global Bioanalysis Consortium, and Delaware Valley Drug Metabolism Discussion Group (January 2015). It is a common topic of conversation in the industry and a source of confusion for many biotech and pharma companies due to the lack of direct guidance. These discussions led to the emergence of four levels of method performance and evaluation, from most to least stringent: 1) Validated, 2) Qualified, 3) Research, and 4) Screening. Selection of the appropriate level, in conjunction with anticipated use of the data generated from the method in question, can often be far from straightforward. Teasing out key information from projects, such as stage and criticality of the decisions to be made using the data, can also be difficult due to the complexity of drug development projects.

**VALIDATED BIOANALYTICAL METHOD**

This level of method performance and evaluation is the simplest to match to the client’s needs. The main intent of this method is to support a regulated (GLP) study. Contract Research Organizations (CROs) evaluate at this level extensively for characteristics such as precision, accuracy, selectivity, sensitivity, and stability of the analyte from sample throughout the bioanalytical process. Three core runs, as well as many evaluations as described in FDA guidance papers, are performed. A validated method is appropriate during preclinical and clinical development phases of the drug development process. The situations when it is appropriate to use a validated bioanalytical method are as follows:

1. The data will be used in support of a GLP toxicological study.
2. The data will be used in support of a clinical study, providing pivotal exposure and/or PK data.
3. The data will be used to make pivotal or safety-related decisions for a project. This includes identified active metabolites through the Metabolism is Safety Testing (MIST) guidance.

If any of the previous statements are accurate for the study, then a validated method would be appropriate. This becomes a slippery slope when supporting data is important, but not necessarily pivotal to understanding the drug development picture. An example of this would be for a dose range finding study in a species that may not be used for your GLP studies. It would be more cost and time effective to qualify the method and validate later if the species is used. However when a safety study is planned, it is typically a best practice to then validate the method to be used. It is always important to keep an open relationship between client and CRO when making these decisions, so the assay rigor is sufficient to support the intended use of the data.

**QUALIFIED BIOANALYTICAL METHOD**

Qualified methods are not required to be as robust as validated methods. Many commenters have questioned the purpose of this level of method performance and evaluation. This method is most relevant in situations where clients want the data to make non-regulated or pivotal decisions, and have reasonable confidence in the data. For example, CROs still need to develop the method with scientific rigor to withstand the scrutiny of method validation; however, the degree of characterization and evaluation is less often used for a validated method. A qualified method can be used at any point in the drug development process, but is most commonly used for additional assessments of tissues or other matrix during preclinical development or late discovery.

Typically, a CRO will perform single method with a statistically appropriate number of quality control (QCs), typically an n of 5 at each level, that span the calibration range, similar to a validated method. This single run could be compared to a single core run of a validation method; however, only precision and accuracy of the single run are evaluated, and a method report is not typically provided.

Some key examples of situations when a qualified method may be appropriate include:

1. When a validated method exists for plasma; however, drug levels in urine or tissues will be measured to add further information to the drug development picture.
2. Non-GLP early toxicological studies, such as dose-range finding studies.
3. Metabolite/prodrug in any matrix during GLP/non-GLP toxicological studies (prior to MIST evaluation).

**RESEARCH BIOANALYTICAL METHOD**

This level of method performance and evaluation is designed to fit the middle to late discovery phase of a drug development project. This level is intended to characterize an analytical method for the analyte that may eventually move into a validated and/or qualified method. It should provide sufficient scientific rigor to provide confidence around late discovery decision-making.

CROs should develop this method with scientific rigor; however, an abbreviated approach may be appropriate. They will complete this method evaluation prior to sample analysis, showing precision and accuracy, but no acceptance criteria needs to be set. Rather, the targeted method performance for the run precision and accuracy should be 20% for the % Relative Standard Deviation (%RSD) and % Relative Error (%RE) and 30% for both on the lower limit of quantitation (LLOQ).

Here are some key situations when a research method may be appropriate:

1. During the discovery phase of the drug development project for decision-making evaluations and as the phase continues closer to preclinical development.
2. To look at additional analytes (e.g., biomarkers, possible metabolites, etc.) during non-GLP toxicology studies or discovery studies to deepen understanding of the project.
SCREEnING BIOANALYTICAL METHOD

The screening bioanalytical method level of method performance and evaluation is the most basic. In this process, a method is not specifically developed for the analyte. A generic method is run that is likely to give adequate results for the analyte to be evaluated. This type of method is ideal for early discovery to see how quickly drug is found in plasma or similar substance. This method type should be used for simple data evaluations where yes/no answers are desired and/or when comparing bioavailability across multiple lead compounds.

A calibration curve, and possibly an abbreviated set of QC samples, are included with the samples to be tested. No acceptance criteria are set and a targeted method performance of 30-40% for %RSD and %RE are reported. These methods are therefore straightforward to set up for sample analysis. The data should be used cautiously and not interpreted for anything other than large-scale changes in analyte level. The situations where a screening method would be appropriate are:

1. Early discovery studies
2. Studies where the data interpretation is intended to be simple and provide basic answers with large margins for variability.

This tiered approach provides much better fit-for-purpose of the bioanalytical method for the drug development process. Following this approach retains the balance of time and cost with scientific rigor and evaluation level and improves the decision-making process. Since there can be ambiguity around the appropriate level of method performance and evaluation required for a study, it is paramount to have open discussions with a bioanalytical professional well-versed in the drug development field to ensure appropriate selection. It is always best to ask more questions and make the best-informed decision than to base a decision on haste or cost.

IMMUNOGENICITY PERSPECTIVE FOR NON-CLINICAL SPECIES

Since 2004, many papers and guidance documents have described the impact of, and methods to characterize, immunogenicity as well as describing the need to perform assessments for all therapeutics that are biologic in nature. Early publications discuss the level of method characterization to be conducted in both non-clinical and clinical species, regardless of the therapeutic’s physicochemical properties. However, as early as 2008 (Koren, et al., 2008) (Shankar, et al., 2008), there have been suggestions that analytical rigor may not need to be equal between non-clinical and clinical species.

More recently, the need for immunogenicity assessments in non-clinical species for biologics intended to treat humans has undergone some scrutiny as to the data’s relevance. Considering that much of today’s biologically-derived therapeutics are intended for humans—and in many cases are fully human proteins—it is no surprise to observe elevated levels of anti-therapeutic antibodies (ATAs) in non-clinical species. Hence, it is common to include a disclaimer, such as seen in the August 2015 draft EMA guideline on immunogenicity assessment of biotechnology-derived therapeutic proteins, which states, “the predictivity of non-clinical studies for evaluation of immunogenicity in humans is considered low” (EMA, 2015).

With the need for such disclaimers one might ask: Why assess immunogenicity in non-clinical species at all? Several reasons to assess the induction and persistence of ATAs in non-clinical species exist, and may well prove useful, even if they are not predictive. The first reason is to identify causes of unusual pharmacokinetic profiles, such as unusual half-life or decreased maximum concentration. Similarly, a toxicologist can correlate immunogenicity with toxicological findings by identifying ATAs. While this correlation may still qualify as “not a reliable predictor”, some level of understanding regarding the effect an immune response may have in the clinical phase is provided. Finally, with new delivery mechanisms and novel therapeutic scaffolds being developed daily, exploring immunogenicity can provide a “sneak peek” to the clinical immunogenicity risk profile, if compared to a well-characterized system.

The International Council of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidelines govern immunogenicity assessments. The 2011 addendum to the ICH S6(R1) guideline seems to imply that ATA and immunogenicity may not be needed in all non-clinical studies (Swanson & Bussiere, 2012) (Harmonization, 2011), but rather states that ATA should be evaluated where there is evidence of altered pharmacodynamic activity (e.g., effect of the drug is altered from expected), unexpected changes in a pharmacodynamic marker, or evidence of immune-mediated reactions. The ICH guidance specifies that while the samples may not need to be assessed, they should be collected and stored for subsequent analysis if needed. The latest draft guidance from the EMA goes further and indicates that, for biosimilar comparability, non-clinical ATA is not recommended (EMA, 2015).

What does this language in these documents mean for the bioanalytical lab? In reality, not much: There is still little room for interpretation regarding the need to develop and validate an ATA assay for Investigational New Drug (IND) enabling studies. These assays should still be developed and validated because, if the need for the data arises and samples need evaluation, there is rarely an abundance of time available to complete the assay development and validation and submit an IND to the agency. However, there is room to define levels of characterization, similar to those outlined above. For consistency with LC-MS/MS methods, these levels can be defined as; validated ATA methods, research ATA methods and screening ATA methods.

Validated ATA methods

Validated ATA methods are mandatory for regulatory toxicology or IND-enabling studies. Validation parameters are clearly defined in several industry white papers as well as FDA guidance documents ((Koren, et al., 2008) (Shankar, et al., 2008) (U.S. Department of Health and Human Services, Center for Drug Evaluation and Research, & Center for Biologics Evaluation and Research, 2009)). The specifics of how to validate these assays is too extensive for discussion here.

For most non-clinical ATA methods, the primary purpose is to describe the PK and pharmacodynamics (PD) data. While the ATA data itself may not be a significant factor in determining if a
therapeutic goes forward into the clinical phase, it can still provide invaluable information as to how potential adverse events could be a result of the immune reaction (such as vasculitis). In some cases, additional characterization may be needed; in particular, if there were unusual or unexpected findings in pre-IND non-clinical studies. These additional considerations could be assessment for cross reactivity with endogenous counterparts, domain specificity, or even immune complex determination.

Research ATA methods
A research method is a good option when evaluating higher order species (non-human primates) that may have proteins with high homology to humans. Screening samples would help identify those that are potentially positive for immunogenicity as well as a means of confirming the immunogenic positive result is accurate and specific to the therapeutic. Delving deeper into the characterization of the ATA response could certainly provide more information; however, the return on investment for this is questionable, as a molecule’s clinical destiny is still somewhat uncertain. Additionally, for some companies, there are still multiple versions of the drug in evaluation at this point in a program’s lifecycle. In such cases, developing a method capable of detecting antibodies against all the different constructs of a molecule and characterizing the extent of that response accurately and reproducibly can prove to be very challenging.

Screening ATA methods
An ATA screening method, the least common type, is very basic and only used when deemed critical to understanding the immune response against a dosed therapeutic at a very early stage in the drug discovery/development life cycle. One example of this would be to assess pre-existing reactivity of a molecule in order to reduce the number of potential candidates from moving forward. Level of characterization considerations noted above also apply to these methods but with the additional complication that the drug is likely in short supply. Given this limitation, these methods could be generic; that is, detecting anti-human antibodies. Due to the likely need to detect all anti-human antibodies for human protein therapeutics, it is best to reserve these kinds of methods to rodent models as a relevant non-human primate positive control would need to be constructed or generated. That said, if this is a type of method that is going to be used over a large number of development programs, the cost and time required to generate an appropriate positive control could prove beneficial.

Table 1: Level of rigor for the bioanalytical methods and the stage of drug development for which they are most appropriate.

<table>
<thead>
<tr>
<th>Method</th>
<th>Level of rigor</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validated</td>
<td>****</td>
<td>Pre-clinical GLP and clinical development</td>
</tr>
<tr>
<td>Qualified</td>
<td>***</td>
<td>Any but most commonly during pre-clinical development or late drug discovery</td>
</tr>
<tr>
<td>Research</td>
<td>**</td>
<td>Mid- to late drug discovery</td>
</tr>
<tr>
<td>Screening</td>
<td>*</td>
<td>Early drug discovery</td>
</tr>
</tbody>
</table>

When determining the level of assay characterization needed for ATA assessment, consider these issues. First, is the therapeutic a protein replacement therapy (e.g. insulin or erythropoietin)? What is the homology between the human and the non-clinical species under evaluation? These are important questions as an immune response to a replacement protein therapeutic that generates cross-reactive antibodies with the endogenous protein can have devastating effects. One classic example is the PEG-rHuMGDF clinical trial (Li, et al., 2001), which resulted in neutralizing antibodies that cross-reacted to endogenous thrombopoietin.

A second consideration for determining if an immunogenicity assay is merited for early discovery is the novelty of the therapeutic and/or the delivery mechanism. If a molecule is a fully human monoclonal antibody, there is a growing body of work that can describe what to expect in non-clinical species. If the compound is a fusion protein that has several potential novel epitopes, or is delivered in a unique way, then it may be worthwhile to evaluate an assay to minimally detect if ATAs are present.

A final consideration, is the proposed administration route of the therapeutic. The 2014 FDA guidance for industry (Immunogenicity Assessment for Therapeutic Protein Products) notes that, in general, intradermal, subcutaneous, and inhalation routes are at increased risk of developing ATAs. It is especially important to understand the desired route of administration if the required level of method characterization is unclear.

Note: For a complete list of references, please visit the online version of this article at ContractPharma.com.