White Paper

Theme: Best Practices for Bioanalytical Methods: Recommendations from the Global Bioanalysis Consortium

Guest Editors: Binodh DeSilva and Philip Timmerman

Large Molecule Specific Assay Operation: Recommendation for Best Practices and Harmonization from the Global Bioanalysis Consortium Harmonization Team

Lauren Stevenson,^{1,10} Marian Kelley,² Boris Gorovits,³ Clare Kingsley,⁴ Heather Myler,⁵ Karolina Österlund,⁶ Arumugam Muruganandam,⁷ Yoshiyuki Minamide,⁸ and Mario Dominguez⁹

Received 14 May 2013; accepted 15 October 2013; published online 16 November 2013

Abstract. The L2 Global Harmonization Team on large molecule specific assay operation for protein bioanalysis in support of pharmacokinetics focused on the following topics: setting up a balanced validation design, specificity testing, selectivity testing, dilutional linearity, hook effect, parallelism, and testing of robustness and ruggedness. The team additionally considered the impact of lipemia, hemolysis, and the presence of endogenous analyte on selectivity assessments as well as the occurrence of hook effect in study samples when no hook effect had been observed during pre-study validation.

KEY WORDS: hemolyzed samples; hook effect; lipemic samples; parallelism; selectivity.

INTRODUCTION

The Global Bioanalysis Consortium is a worldwide organization consisting of representatives from the pharmaceutical industry, contract research organizations, and academia with the mission to come forward with scientifically driven recommendations to health authorities and regulatory bodies worldwide on globally agreed best practices for bioanalytical method validation (BMV). To this end, 20 teams, addressing different aspects of BMV have been formed, with 6 teams focusing on topics which apply to ligand binding assays for the measurement of large molecules. The L2 team was charged with addressing large molecule specific assay operation for validation of ligand binding assays in support of pharmacokinetics. Team L2 membership included representatives worldwide, with participants from North America, South America, Europe, and Asia Pacific.

SCOPE

The scope of team L2 included the following:

- Setting up balanced validation design
- Specificity testing
- Selectivity testing, including consideration of the impact of lipemia, hemolysis, and presence of endogenous analyte
- Dilutional linearity
- Hook effect
- Parallelism
- Robustness and ruggedness

Over the course of discussions, the team engaged their local colleagues as well as other regional experts in order to bring a comprehensive view to each of the topics. In addition, regular engagement with the bioanalytical community and regulatory agency representatives was achieved *via* presentations at multiple forums including, but not limited to the following: 2nd Japan Bioanalysis Forum, Tokyo, Japan (March 2012); the 6th Workshop on Recent Issues in Bioanalysis, San Antonio, TX (March 2012); 1st Latin American Meeting on Bioanalysis, São Paulo, Brazil (May 2012); AAPS National Biotechnology Conference, San Diego, CA (May 2012); and European Bioanalysis Forum, 3rd Focus Meeting, Brussels, Belgium (June 2012).

In general, the team's recommendations are well aligned with previously published white papers and regulatory guidance (1,2). However, some recommendations do differ in the particulars, albeit not in spirit, from recently issued guidance. Specifically, the team does not recommend routine testing of lipemic and hemolyzed samples in pre-study validation selectivity

¹ Biogen Idec, 14 Cambridge Center, Cambridge, Massachusetts 02142, USA.

² MKelley Consulting, West Chester, Pennsylvania, USA.

³ Pfizer, Andover, Massachusetts, USA.

⁴ Quotient BioAnalytical Sciences, Fordham, UK.

⁵ Bristol-Myers Squibb, Princeton, New Jersey, USA.

⁶ Gyros AB, Uppsala, Sweden.

⁷ Affigenix Biosolutions Pvt Ltd, Bangalore, India.

⁸ Shimadzu Techno-Research Inc., Kyoto, Japan.

⁹ DominguezLab, Parana, Entre Rios, Argentina.

¹⁰ To whom correspondence should be addressed. (e-mail: Lauren.Stevenson@biogenidec.com)

assessments, although such evaluations should be considered when the characteristics of the therapeutic molecule, its target, the disease indication, or the assay format provide a scientific rationale for doing so. Similarly, routine parallelism assessments are not recommended, but rather should be considered based upon the characteristics of the therapeutic molecule, consideration of the effects of potential binding partners *in vivo* as well as the characteristics of specific critical reagents employed in the assay. In order to provide additional context and clarity, these topics are addressed in greater detail.

Throughout the course of deliberations and discussion, the team was mindful to ensure that the recommendations put forth were done so with clear scientific rationale. Therefore, it is acknowledged that while many assessments should be considered, not all will be found to be necessary, nor appropriate, for every molecule. The application of rigorous scientific principles in determining the appropriate assessments to be performed should ensure the reliability of the assay and the data generated therefrom.

BALANCED VALIDATION DESIGN

A key element of balanced validation design is to ensure that the same number of observations is made in each precision and accuracy (P&A) run. Specifically, the number of sets of control samples in each run should be the same. This approach aligns with team L1's recommendation for precision and accuracy batches which states that, in each batch, three independent sets each of five QC levels (LLOQ, LQC, MQC, HQC, ULOQ) should be assessed in ≥ 6 P&A runs (3).

When considering the number of analysts performing validation as a variable for balanced validation design, it is recommended that the number of analysts used reflects the subsequent sample testing practices. It is widely acknowledged that one analyst often performs the bulk of the validation experiments, with inclusion of second analyst only for a subset of P&A runs. This practice is deemed acceptable since testing laboratories later qualify new analysts on a given assay, generally by the acceptable performance of at least two consecutive P&A runs. Although the inclusion of at least two analysts during validation is widely accepted as best practice, the team recognizes that it is possible to justify using only one analyst during validation if study sample analysis will also only use one analyst, which is potentially the case when testing small studies. However, if multiple analysts will be employed to test samples, then the use of at least two analysts in validation is strongly recommended. This practice not only allows the opportunity to identify systematic analyst bias, but also contributes to the understanding of assay ruggedness.

SPECIFICITY

From a ligand binding assay point of view, specificity is defined as the ability of a binding reagent (*e.g.*, an antibody) to bind to the analyte of interest (*e.g.*, antigen). The team agrees with the definition for specificity as described in De Silva *et al.* (4) and EMA guidelines (2), which indicate that the specificity of a binding reagent (antibody) refers to its ability to bind to the analyte (antigen) of interest.

To assess assay and reagent specificity, potentially cross-reacting molecules (variant forms of the analyte, physico-

chemically similar compounds) as well as other potential interferents, such as concomitantly administered drugs, are evaluated to determine whether they affect the performance of the ligand binding assay. Possible cross reactants are chosen based on the structure of the drug molecule. This assessment is more relevant for types of large molecule drugs other than monoclonal antibodies, since the selectivity assessment is already performed in the presence of high levels (milligrams per milliliter) of endogenous immunoglobulins. For concomitant medications, those specified in the protocol to be co-administered should be tested. There is usually no need to test for over the counter medications or co-administered small molecules with no structural similarity to the analyte of interest. In addition, since anti-drug antibodies (ADA) including pre-existing specific, auto or heterophilic antibodies and circulating soluble target may influence ligand binding kinetics and assay performance, testing the effects of the presence of these potential interferents should be considered.

The approach employed to test for specificity is to spike pooled blank matrix, LLOQ, and HQC validation samples with increasing amounts of possible interferents up to the maximum anticipated physiological concentration in a given study population.

In order to claim specificity, or lack of interference, acceptance criteria are based on recovery of drug molecule in the spiked samples. Recovery should be within 20% of nominal concentration for HQC and 25% for LLOQ. Unspiked samples (*i.e.*, blank matrix) should measure $< \text{LLOQ}$. However, in cases where interference is expected to occur, for example, in the presence of ADA, this evaluation may be used to define the levels of interferent tolerated by the assay as defined by the levels at which the above acceptance criteria are met.

SELECTIVITY

Generally speaking, the team accepted the standard definition of selectivity as explained by DeSilva *et al.* (4), *i.e.*, the ability of an assay to measure the analyte of interest in the presence of other constituents in the sample. Either an inhibition or enhancement of the assay signal may be observed due to factors such as heterophilic antibodies, rheumatoid factor, enzymes, or structural proteins in the matrix. These factors will vary between individuals and may depend upon diet, ethnicity, disease-state, age, gender, fasting state, and more.

Selectivity should be assessed during assay development since identification of issues at this stage offers the opportunity to reconsider assay format, antibody reagents, blocking buffers, and MRD. Formal assessment should then be included in the method validation as follows:

Ten or more individual lots should be tested unspiked and spiked at the LLOQ. Testing at a higher analyte concentration is also recommended and is typically performed at the HQC level. Samples may be tested when freshly prepared (taking care to mix thoroughly to allow for applicable molecular interactions with native matrix components) or, for flexibility, after one freeze/thaw cycle, assuming that the appropriate freeze/thaw stability has been established.

For unspiked samples, $\geq 80\%$ (8/10) should measure $< \text{LLOQ}$. For spiked samples, $\geq 80\%$ (8/10) should measure within $\pm 20\%$ of the nominal concentration for higher level (HQC) spike and within $\pm 25\%$ for LLOQ spike. Furthermore, the same 80% of samples should meet these criteria at both spike levels.

The evaluation of selectivity should be based on clinically relevant considerations, and therefore interpretation of the guidance maybe required. For example, validations to support clinical trials in patient populations should typically include the use of disease-state matrix for selectivity testing, since matrix factors that bias results could differ between these individuals and healthy volunteers. Matrix from autoimmune conditions such as rheumatoid arthritis or inflammatory conditions such as sepsis and viral infection may be more likely to affect LBA readouts (5).

In some cases, it is not possible to source disease-state matrix, or ten individuals may not be available. Acknowledging this, the team recommends that as many individuals as feasible are included within the ten samples used for selectivity testing. If none are available at the time of validation, in-study selectivity assessments should be included once pre-dose study samples become available.

Special Cases

Endogenous Protein

When the therapeutic molecule has an endogenous counterpart, measuring only the exogenous drug can be challenging. Occasionally, it may be possible to generate critical reagents that only recognize the exogenous form of the molecule, but typically other approaches are necessary. In general, it will be important, when possible, to create standards and QCs in matrix with low endogenous levels of the analyte. This will require screening of multiple individual lots to generate a “low endogenous” pool. Several approaches for dealing with the presence of endogenous protein should then be considered. Sacrificing the level of the LLOQ such that the signal generated by the therapeutic molecule is sufficiently above the “noise” of the endogenous molecule is commonly employed. Alternatively, when the endogenous concentrations are relatively high (measuring within assay range), subtracting the endogenous concentration present in the unspiked matrix sample from the measured result in the spiked sample has, in some cases, proved useful. However, subsequently, it is imperative that a pre-dose sample is available for all subjects tested during sample analysis, and it should also be understood that pre-dose levels of endogenous analyte may not be representative of post-dose levels. Lastly, if the endogenous concentration is generally very high across individuals, consider whether the range of the assay is appropriate.

Lipemic Samples

If the medical condition being treated is associated with high lipemia, such as certain metabolic conditions like type 2 diabetes mellitus, then this should be taken into consideration in the selectivity assessment. Typically, standard selectivity testing using disease-state individuals during validation will

address this concern and any additional testing in artificially created lipemic samples is not deemed necessary. However, since food intake or infusion of lipids can lead to varying triglyceride concentrations in serum or plasma, regardless of disease indication, it is prudent to perform a case by case assessment to determine whether sufficient scientific rationale exists to warrant additional selectivity assessments in lipemic samples. Only a very small number of examples within the team could be identified where lipemia caused a selectivity concern. This may be unsurprising, given that protein drugs are generally hydrophilic and therefore the binding reactions in an aqueous environment are not likely to be affected by lipids. Cases where lipemia has caused a selectivity concern, though, may be related to the type of drug being developed (6) or the assay format. For example, homogeneous assays appear to be more susceptible to interferences of this type. Overall, the team does not recommend routine selectivity testing of lipemic samples for the majority of method validations.

Hemolyzed Samples

Hemolysis of samples is typically associated with poor collection procedures. For small molecule bioanalysis, inclusion in method validation is routine and required (7,8); however, for large molecules, this requirement is less clear. Immunoassays are mostly unaffected by hemolysis unlike other analytes measured by spectral or chemical means (9). Accordingly, the team was unable to identify multiple examples where hemolyzed samples yielded out of trend results, again suggesting that this is a rare selectivity issue for ligand binding assays.

Therefore, the team does not recommend routine selectivity testing of hemolytic samples for the majority of method validations. However, when the drug, target, disease indication, or assay format suggest that assay interference is likely, then this assessment should be considered. For example, hemolysis may be unacceptable for immunoassays of relatively labile analytes like insulin, glucagon, calcitonin, parathyroid hormone, ACTH, and gastrin due to the release of proteolytic enzymes from erythrocytes that degrade these analytes (10).

DILUTIONAL LINEARITY AND HOOK EFFECT

Despite advances in ligand binding assay platforms and formats, these assays frequently have relatively limited quantitative ranges (LLOQ to ULOQ), spanning only one to three orders of magnitude. Therefore, in order to analyze samples of high concentration in such assays, significant sample dilution may be required. The examination of supra-physiologically high concentrations of biotherapeutic analyte can help define the region in the dynamic assay range that may be susceptible to false negative results (11–13) or under-recovery due to hook effect (or prozone). Hook effect can be attributed to a number of factors including high molecular weight species or aggregates associated with high analyte concentrations (14,15), saturation of binding capacity of the critical reagents in the assay by an overwhelming amount of analyte, and interference by binding proteins such as ADA, target, or other antigen (12,13,16–18).

Dilutional linearity experiments are performed to demonstrate that high concentrations of the analyte of interest can be accurately measured by diluting into the assay's quantitative range and multiplying the measured concentration by the dilution factor. Hook effect is typically assessed in the same experiment by including samples spiked with very high concentrations of analyte which are tested without dilution beyond MRD. A prozone, or range in which observed concentrations may not be reflective of actual concentrations, is identified when greater or increasing dilutions result in stagnant or increased signals compared to the preceding lesser dilutions.

When possible, blank matrix used to dilute linearity and hook effect samples should be representative of matrix that will be used for sample dilution in production. The blank matrix sample should be spiked with analyte of interest at or above the maximum anticipated concentration (Cmax) levels expected in study samples. In the absence of information on anticipated Cmax, it is recommended that samples be spiked at the highest feasible concentration of analyte where the sample is composed of least 90% matrix ($\leq 10\%$ stock solution: $\geq 90\%$ matrix, v/v). The sample with the highest feasible concentration (or Cmax, as applicable) can then be diluted to carefully examine the dilution to signal relationship and analyte recovery throughout the assay range. At least one sample above the ULOQ (hook effect) and 3–5 samples within the quantitative assay range, and one sample below the LLOQ should be examined to fully characterize dilutional linearity. For an excellent graphical representation, see DeSilva *et al.* 2003 (4).

During validation, dilutional linearity samples with measured concentrations within the quantitative range of the assay should return values that are within $\pm 20\%$ of theoretical. In addition, the precision (%CV) of the cumulative back-calculated concentrations for all in-range samples should be $\leq 20\%$. Hook effect samples spiked at concentrations between the maximum anticipated concentration and ULOQ concentrations should measure $> \text{ULOQ}$. If samples spiked within this range recover at values $< \text{ULOQ}$, a hook effect is present and measures should be taken to manage the hook effect. Table I shows an example of a dilutional

linearity/hook effect experiment where no hook effect is observed and dilutional linearity is demonstrated within the dynamic assay range.

Managing Hook Effect

Ideally, the presence or absence of hook effect will be tested during method development. When a hook effect is discovered during the development phase, there is an opportunity to re-optimize the assay in order to eliminate the hook effect. This might include further titration of capture and detection antibody concentrations, increasing the MRD to break up binding complexes, and/or optimizing assay buffer ionic strength, surfactant levels, or pH to break up aggregates.

If hook effect is not uncovered until pre-study validation, and assay re-optimization is not an option, then measures will need to be put in place to manage the hook effect during sample analysis. Initially, it is advisable to verify the hook effect in the applicable study population by running a predefined subset of highest anticipated concentration samples from each dose group at multiple dilutions to fully characterize the hook effect in that study population. Should the hook effect be verified in the study population, the implementation of multiple or predefined sample dilutions during study sample testing will likely be required. However, once the pharmacokinetic properties of the therapeutic have been fully characterized and the PK is predictable, it may be possible to define and implement a single predefined dilution in place of the multiple dilution strategy (19).

Notably, hook effect may not be observed during pre-study validation, but become apparent during in-study sample analysis. The cause of such hook effects may be due to accumulation of drug to higher concentrations than anticipated, an increase in endogenous binding partners after drug treatment, or the generation of ADA. In-study hook effect is typically noted as anomalous PK during scientific review of the data. In these instances, it is important to examine all of the relevant data (ADA status for example) and to design an investigation that characterizes the hook effect. Once

Table I. Dilutional Linearity—No Hook Effect Observed

Sample ID	Target concentration (ng/mL)	Dilution (Fold)	Measured concentration (ng/mL)	Final concentration (ng/mL)	Recovery (%)
DL1	250000	1	$> \text{ULOQ}$	$> \text{ULOQ}$	NA
DL2	2500	1	$> \text{ULOQ}$	$> \text{ULOQ}$	NA
DL3	25	1	$> \text{ULOQ}$	$> \text{ULOQ}$	NA
DL4	9.0	55556	9.127	507060	101.4
DL5	5.00	100000	5.186	518600	103.7
DL6	2.50	200000	2.607	521400	104.3
DL7	1.00	500000	0.940	470000	94.0
DL8	0.250	2000000	0.274	547000	109.4
DL9	0.050	10000000	$< \text{LLOQ}$	$< \text{LLOQ}$	NA
			Mean	527030	105.4
			N	5	5
			Stdev	28242	5.66
			%CV	5.36	5.37

Reference standard=500,000 ng/mL

ULOQ=10.0 ng/mL

Final conc. = observed conc. X dilution

characterized, the need to retest additional samples can be assessed, and if deemed necessary, an *a priori* testing and data reporting scheme should be put in place.

In summary, the team recommends that tests for hook effect be performed in development where possible and reassessed during pre-study validation. Even when no hook effect has been observed, careful scientific review of in-study data is required to ensure that no hook effect is present in-study. When hook effect has been identified and well characterized, several approaches for sample testing and data reporting may be valid as long as they adhere to rigorous scientific procedures and are defined prior to sample testing.

Parallelism Evaluation

The concept of parallelism is similar to dilutional linearity except that parallelism assesses incurred study samples. At present, routine parallelism assessments are not being broadly implemented industry-wide. Examples of non-parallelism experienced by the team were rare, involved non-mAb therapeutics, and were apparent during scientific review of the sample data. Therefore, the need to perform a parallelism assessment for a given biotherapeutic depends upon the characteristics of the drug, its binding partners, and assay reagents' specificity. Scientific rationale should drive the decision to perform parallelism for a given therapeutic in a given study and novel non-mAb-based modalities may require closer attention compared to typical mAb-based biologics when determining whether parallelism assessments are advisable. Examples of factors important when evaluating the need for a parallelism investigation include propensity of the drug to aggregate, drug stability *in vivo*, presence of anti-drug antibodies, endogenous binding partners, and assay specificity toward complexes formed. Partial degradation of the compound while in circulation *in vivo* may result in formation of stable fragments with various degree of specificity toward assay reagents and therefore ability to impact compound PK assay. The presence of binding or neutralizing anti-drug antibodies can impact interaction between the drug and specific assay reagents resulting in a reduced assay signal. Finally, *in vivo* circulating biotherapeutic may be present in free or bound with the target form (complexed) and binding of these forms to assay reagents may differ.

If a parallelism assessment is deemed necessary, by definition, it cannot be performed until incurred study samples become available (1,2,17,20). Generally speaking, careful review of study PK data will ensure that instances of in-study hook effect (as discussed above) and the likely presence of non-parallelism are noted and investigated appropriately. An investigation may include, for example, a correlation of PK and ADA data, and review of the assay performance and PK profiles across study subjects and dose levels. Analysis of incurred study samples at multiple dilutions (parallelism test) may then facilitate the understanding of the potential interferences of circulating matrix components and yield important information about biotherapeutic processing *in vivo*.

Presently, there are differing views on best practices for performing parallelism assessments, when applicable, and no clear consensus has yet emerged. Some advocate the use of pooled samples for practical considerations such as creation

of larger sample volumes and avoiding generation of multiple datasets for the same sample. Others advocate the use of individual samples as a means of ensuring that the "correct" concentration is reported for each individual sample. Complexities of execution and subsequent data reporting arise in either scenario. In principle, however, it is agreed that the parallelism assessment is similar to evaluation conducted during dilutional linearity. Incurred samples (pooled or individual) are tested at multiple dilutions that are expected to yield concentrations that fall above the assay ULOQ (to evaluate prozone or hook effect) as well as within the assay range. Since the goals of the parallelism evaluation are to ensure that the study PK data are representing the most accurate information and to understand the potential impact of circulating matrix components, some of which could be accumulating or induced as a result of drug administration, testing samples collected at later study time points should be performed as opposed to samples collected early in the study.

Should non-parallelism be detected, a sample testing scheme will need to be put in place to mitigate the issue and an *a priori* strategy for data reporting will need to be established. In some cases, simply increasing the dilution at which samples are tested may be adequate to drive dissociation of complexes causing the non-parallelism. In others, an alternative PK assay design that utilizes different critical capture and detection reagents may be required.

It is also important to be aware that even when a parallelism assessment passes proposed acceptance criteria, trends that may have meaningful impact on the study data may still be present. An example is presented in Table II which shows multiple dilution data from samples collected from two animals at different time points after drug injection. Animals 1 and 2 were dosed with two separate drug compounds and samples were analyzed in separate PK assays. For the Animal 1, samples were tested at days 29 and 57, while for Animal 2, samples were tested at days 10 and 14 after drug administration. For all samples tested, the inter-dilution precision (%CV) is well within the commonly applied $\pm 30\%$ criterion (4). However, the data suggest a trend of increasing drug recovery with increasing dilution for the samples collected from Animal 1 on days 29 and 57 and Animal 2 on day 14. A scientific judgment should therefore be applied to determine whether an additional investigation is warranted.

In summary, the team does not recommend routine parallelism assessments, but advocates such assessments where supported by scientific rationale. In general, incidents of non-parallelism can be caught by careful review of study data and mitigated through scientifically rigorous investigation and implementation of appropriate sample testing procedures.

ROBUSTNESS AND RUGGEDNESS

Robustness/ruggedness terms are often used interchangeably and there has historically been confusion regarding the absolute definition of each term. What is more important than prescribing definitions, though, is that it is understood that both parameters are indicators of assay reproducibility under varied conditions. Therefore, any

Table II. Parallelism Trends

Sample ID	Sample dilution	Reportable result	Statistics	
Animal 1	100	>ULOQ		
Day 29	2,000	23,969		
	3,000	30,991		
	4,500	35,237		
	6,750	37,398		
	10,125	<LLOQ	SV	5,919
	15,188	<LLOQ	Mean	31,899
	22,781	<LLOQ	CV%	18.6
Animal 1	100	>ULOQ		
Day 57	1,000	19,867		
	1,500	24,059		
	2,250	32,227		
	3,375	28,331		
	5,063	32,830	SV	5,468
	7,594	33,255	Mean	28,428
	11,390	<LLOQ	CV%	19.2
Animal 2	20	563		
Day 10	30	598		
	45	566		
	68	608		
	101	652		
	152	625	SV	35.8
	228	629	Mean	613
	342	660	CV%	5.9
Animal 2	20	165		
Day 14	30	210		
	45	243		
	68	273		
	101	329		
	152	293	SV	61.0
	228	235	Mean	237
	342	152	CV%	25.7

robustness/ruggedness analysis should address the question of whether the assay will perform well under real-life changes in standard laboratory situations, with consideration for the specific circumstances anticipated during study support.

Robustness/ruggedness testing should generally be incorporated into the method development process and typically includes, but is not limited to, assessment of the following: variations in incubation times and temperatures, reagent and plate lot changes, multiple analysts, and multiple instruments. In determining the appropriate conditions to test, it is important to be mindful of the needs of the assay and the conditions under which it may be run, such as differences in ambient temperatures in different testing labs and regional differences in serum/plasma sources.

When addressed thoroughly during assay development, robustness/ruggedness is further demonstrated during validation, by virtue of the use of multiple instruments/analysts and typical variations in incubation times. Furthermore, as the development program for the molecule progresses, cross validation of the assay at multiple labs provides additional demonstration of robustness/ruggedness. When robustness/ruggedness has not been adequately addressed during method development, a formal evaluation during validation may be considered.

REFERENCES

1. FDA. Guidance for industry bioanalytical method validation guidance for industry bioanalytical method validation. Washington, DC; 2001; Available from: <http://www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf>
2. European Medicines Agency E. Guideline on bioanalytical method validation 2011; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf
3. Kelley M, Beaver C, Stevenson L, Bamford R, Gegwich P, Katsuhiko Y, *et al.* Consensus and recommendation of the L1 Global Harmonization team on run acceptance for large molecule bioanalysis to support pharmacokinetics. AAPS J. 2013;in press.
4. DeSilva B, Smith W, Weiner R, Kelley M, Smolec J, Lee B, *et al.* Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. Pharm Res. 2003;20(11):1885–900.
5. Lee J, Ma H. Specificity and selectivity evaluations of ligand binding assay of protein therapeutics against concomitant drugs and related endogenous proteins. AAPS J. 2007;9(2):E164–70.
6. Martinez-Subiela S, Ceron JJ. Effects of hemolysis, lipemia, hyperbilirubinemia, and anticoagulants in canine C-reactive protein, serum amyloid A, and ceruloplasmin assays. Can Vet J. 2005;46(7):625–9.
7. Berube E-R, Taillon M-P, Furtado M, Garofolo F. Impact of sample hemolysis on drug stability in regulated bioanalysis. Bioanalysis. 2011;3(18):2097–105.
8. Garofolo F, Rocci Jr ML, Dumont I, Martinez S, Lowes S, Woolf E, *et al.* White paper on recent issues in bioanalysis and regulatory findings from audits and inspections. Bioanalysis. 2011;3(18):2081–96.
9. Tate J, Ward G. Interferences in immunoassay. Clin Biochem Rev. 2004;25(2):105–20.
10. Schiettecatte J, Anckaert E, Smits J. Interferences in Immunoassays. In: Chiu NHL, editor. Adv Immunoass Technol. 2012.
11. Heidelberger M, Kendall F. Quantitative theory of the precipitin reaction: III. The reaction between crystalline egg albumin and its homologous antibody. J Exp Med. 1935;62(5):697–720.
12. Gillet P, Scheirlinck A, Stokx J, De Weggheleire A, Chauque HS, Canhanga ODJV, *et al.* Prozone in malaria rapid diagnostics tests: how many cases are missed? Malar J. 2011;10:166.
13. Dubois-Galopin F, Beauvillain C, Dubois D, Pillet A, Renier G, Jeannin P, *et al.* New markers and an old phenomenon: prozone effect disturbing detection of filaggrin (keratin) autoantibodies. Ann Rheum Dis. 2007;66(8):1121–2.
14. Lu C-H, Kalmar B, Malaspina A, Greensmith L, Petzold A. A method to solubilize protein aggregates for immunoassay quantification which overcomes the neurofilament “hook” effect. J Neurosci Methods. 2011;195(2):143–50.
15. Wang X, Das TK, Singh SK, Kumar S. Potential aggregation prone regions in biotherapeutic: a survey of commercial monoclonal antibodies. mAbs. 2009;1(3):254–67.
16. Kozel TR, MacGill RS, Percival A, Zhou Q. Biological activities of naturally occurring antibodies reactive with *Candida albicans* mannan. Infect Immun. 2004;72(1):209–18.
17. Lee JW, Kelley M, King LE, Yang J, Salimi-Moosavi H, Tang MT, *et al.* Bioanalytical approaches to quantify “total” and “free” therapeutic antibodies and their targets: technical challenges and PK/PD applications over the course of drug development. AAPS J. 2011;13(1):99–110.
18. Talamo G, Castellani W, Dolloff NG. Prozone effect of serum IgE levels in a case of plasma cell leukemia. J Hematol Oncol. 2010;3:32.
19. Myler HA, Phillips KR, Dong H, Tabler E, Shaikh M, Coats V, *et al.* Validation and life-cycle management of a quantitative ligand-binding assay for the measurement of Nulojix, a CTLA-4-Fc fusion protein, in renal and liver transplant patients. Bioanalysis. 2012;4(10):1215–26.
20. Kelley M, DeSilva B. Key elements of bioanalytical method validation for macromolecules. AAPS J. 2007;9(2):E156–63.