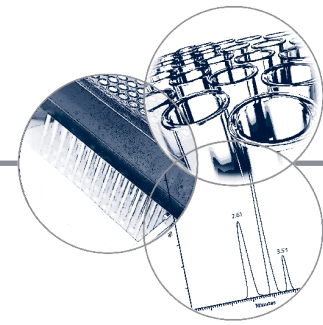
White Paper





# 2013 White Paper on recent issues in bioanalysis: '*hybrid*' – the best of LBA and LCMS

The 2013 7th Workshop on Recent Issues in Bioanalysis was held in Long Beach, California, USA, where close to 500 professionals from pharmaceutical and biopharmaceutical companies, CROs and regulatory agencies convened to discuss current topics of interest in bioanalysis. These 'hot' topics, which covered both small and large molecules, were the starting point for fruitful exchanges of knowledge, and sharing of ideas among speakers, panelists and attendees. The discussions led to specific recommendations pertinent to bioanalytical science. Such as the previous editions, this 2013 White Paper addresses important bioanalytical issues and provides practical answers to the topics presented, discussed and agreed upon by the global bioanalytical community attending the 7th Workshop on Recent Issues in Bioanalysis.

The 7th Workshop on Recent Issues in Bioanalysis (WRIB) took place in Long Beach, California, USA on 9–10 April 2013 with close to 500 professionals representing over 200 companies including multiple regulatory agencies.

The contributing chairs included Binodh DeSilva (Bristol-Myers Squibb, USA), Lakshmi Amaravadi (Biogen Idec Inc., USA), Margarete Brudny-Kloeppel (Bayer Pharma, Germany), Adrien Musuku, (Pharmascience, Canada), Lauren Stevenson (Biogen Idec Inc., USA), Mario Rocci (ICON Development Solutions, USA), and Fabio Garofolo (Algorithme Pharma Inc., Canada).

The contributing regulatory agency representatives included Brian Booth (US FDA), Sam Haidar (FDA), Jan Welink (Dutch Medicines Evaluation Board [MEB] and European Medicines Agency [EMA]), João Tavares Neto (Brazil Agência Nacional de Vigilância Sanitária [ANVISA]), Eric Ormsby (Health Canada Therapeutic Products Directorate [TPD]), Craig Simon (Health Canada TPD), Noriko Katori (Japan Ministry of Health, Labour, and Welfare - National Institute of Health Sciences [MHLW-NIHS]), Emma Whale (UK Medicines and Healthcare products Regulatory Agency [MHRA]), and Jason Wakelin-Smith (UK MHRA).

As is the case every year [1–5], a great number of topics were addressed in this year's edition of the WRIB. This White Paper focuses on the discussions, consensus and resulting recommendations on 16 recent issues ('hot' topics) in bioanalysis. From these 16 topics, eight of them pertain to issues related to LCMS methods, five were more specific to LBA, and three were related to hybrid LBA and LCMS applications.

## LCMS discussion topics

1. **Incurred sample stability** (ISS): should ISS become a regulatory requirement such as incurred sample reanalysis (ISR)? Is ISS defined as 'good science'? When is ISS needed? How do we calculate 'Time 0' in ISS?

2. Use of incurred samples for metabolite testing and specificity during method development: is there a way to avoid the use of incurred samples for metabolite/specificity testing during method development when the reference standard material is not available? What is the best approach to evaluate the impact of multiple co-administered drugs for oncology studies? Can predose samples be used? What are the pros/cons of performing small pilot studies for method development purposes: can these pilot studies be approved by the Institutional Review Board (IRB)? What are the preclinical versus clinical approaches for using incurred samples during method development to improve method quality? What is the industry standard to prove method specificity?

3. '**Fit-for-purpose**' method validation is typically used for biomarkers, tissue analysis,

**Lauren Stevenson[1], Fabio Garofolo*[2], Binodh DeSilva[3], Isabelle Dumont[2], Suzanne Martinez[2], Mario Rocci[4], Lakshmi Amaravadi[1], Margarete Brudny-Kloeppel[5], Adrien Musuku[6], Brian Booth[7], Catherine Dicaire[2], Laura Wright[8], Laurence Mayrand-Provencher[2], Mike Losauro[8], Dominique Gouty[8], Mark Arnold[3], Surendra Bansal[9], Sherri Dudal[10], Dawn Dufield[11], Jeff Duggan[12], Christopher Evans[13], Eric Fluhler[14], Stephanie Fraser[15], Boris Gorovits[11], Sam Haidar[7], Roger Hayes[16],**

Author list and affiliations continued overleaf…

FUTURE SCIENCE part of fsg

**Stacy Ho[17], Richard Houghton[18], Rafiqul Islam[19], Rand Jenkins[20], Noriko Katori[21], Surinder Kaur[22], Marian Kelley[23], Magnus Knutsson[24], Jean Lee[25], Hanlan Liu[17], Steve Lowes[26], Mark Ma[25], Alvydas Mikulskis[1], Heather Myler[3], Bob Nicholson[20], Timothy Olah[3], Eric Ormsby[27], Shefali Patel[28], Vincenzo Pucci[29], Chad Ray[30], Gary Schultz[26], Judy Shih[25], Ronald Shoup[31], Craig Simon[27], An Song[22], João Tavares Neto[32], Valerie Theobald[33], Theingi Thway[25], Jason Wakelin-Smith[34], Jian Wang[3], Laixin Wang[35], Jan Welink[36], Emma Whale[34], Eric Woolf[37] & Raymond Xu[38]**

[1]Biogen Idec Inc.,Cambridge, MA, USA
[2]Algorithme Pharma Inc., 575 Armand-Frappier Blvd., Laval (Montreal), QC, H7V 4B3, Canada
[3]Bristol-Myers Squibb, Princeton, NJ, USA
[4]ICON Development Solutions, Whitesboro, NY, USA
[5]Bayer Pharma AG, Berlin, Germany
[6]Pharmascience, Montreal, QC, Canada
[7]US FDA, CDER, Silver Spring, MD, USA
[8]Intertek, San Diego, CA, USA
[9]Hoffmann-La Roche, Nutley, NJ, USA
[10]Hoffmann-La Roche, Basel, Switzerland
[11]Pfizer, Andover, MA, USA
[12]Boehringer-Ingelheim, Ridgefield, CT, USA
[13]GlaxoSmithKline, King of Prussia, PA, USA
[14]Pfizer, Pearl River, NY, USA
[15]Pfizer, Groton, CT, USA
[16]MPI Research, Mattawan, MI, USA
[17]Sanofi, Waltham, MA, USA
[18]Quotient Bioresearch, Fordham, UK
[19]Celerion, Lincoln, NE, USA
[20]PPD, Richmond, VA, USA
[21]Japan MHLW-NIHS, Tokyo, Japan
[22]Genentech, South San Francisco, CA, USA
[23]MKelley Consulting LLC, West Chester, PA, USA
[24]Ferring Pharmaceuticals, Copenhagen, Denmark
[25]Amgen Inc., Thousand Oaks, CA, USA
[26]Quintiles Bioanalytical and ADME Labs, Ithaca, NY, USA

Author list and affiliations continued overleaf…

and discovery bioanalysis. Is fit-for-purpose used in regulated bioanalysis? How? Where? What is considered a fit-for-purpose method validation? How can discovery bioanalysis fit-for-purpose acceptance criteria help regulated bioanalysis?

4. DBS sampling: is there an industry consensus on the major recommendations from the International Consortium for Innovation and Quality in Pharmaceutical Development (IQ Consortium) Microsampling Working Group? How close is the industry on refining this technique to meet regulatory requirements for having this methodology accepted for submission? What would constitute sufficient evidence for regulatory agencies to accept this technology in regulated bioanalysis?

5. Issues regarding metabolites in safety testing (MIST): what is the industry interpretation of the tiered approach commonly used in bioanalysis field for MIST? What are the tiers in this so-called tiered approach to address bioanalysis for MIST? Should individual or pooled samples be used for relative metabolite exposure analysis? Should *N*-glucuronide metabolites be included for MIST? Or should they be excluded given that they are Phase II metabolites and are not acyl-glucuronides?

6. Evaluation of whole blood stability: what are the industry standards based on the recent Global CRO Council for Bioanalysis (GCC) recommendations [6]? What is the criteria for 'fresh blood'? What is the best approach for the evaluation of blood stability at the collection stage: freshly spiked versus freshly extracted? Do tests need to be performed in single or multiple donors? Do special populations need to be tested?

7. Overcoming nonspecific binding: what are the parameters allowing the identification/observation of nonspecific binding in low protein matrices, such as urine? How best to detect adsorption to the container wall? What are the recommendations on the best practice to handle nonspecific binding issues? What is the most critical practice to prevent nonspecific binding in peptide analysis? What specific materials should be avoided to reduce nonspecific binding?

8. For the hyperlipidemic matrix test performed as part of method validation, what type of matrix should be used to ensure a scientifically meaningful test? Are there some cases where this test may be unnecessary?

## LBA discussion topics

1. Importance of parallelism in LBA: when is the use of parallelism evaluation recommended for PK assays (e.g., to verify analyte stability, examine for biotransformation, examine patient specific matrix effects – complex association/dissociation)? When is the use of parallelism evaluation not recommended for PK assays (e.g., determined unnecessary via risk mitigated assessment, well-characterized pharmacology and stability)? Should parallelism routinely be included in PK and/or biomarker assay validation? When is it appropriate to use parallelism to assess selectivity for biomarker assays? Is a biomarker assay selective if spike recovery fails but parallelism passes? What are appropriate acceptance criteria for parallelism assessments for PK assays? Biomarker assays? Can the hook effect always be controlled? What is the best practice to investigate in-study hook effect?

2. Immunogenicity and effect on PK assays: what is the best strategy of implementing antidrug antibody (ADA) testing in preclinical or clinical studies? Why is determination of the free-drug concentration in the PK assay critical in large molecule drug development? When 'unique PK' is observed, what approach should be taken to identify the root causes?

3. Immunogenicity and neutralizing antibody (NAb) assays: do we always need to develop cell-based and noncell-based assays for immunogenicity? Do we need to develop cell-based NAb assays only for high-risk proteins? Is this the best industry practice? What to do when a fully developed cell-based NAb assay does not meet the purpose of evaluating neutralizing potential of antidrug antibodies? (e.g., when cell-based assay is not sensitive enough due to inherent challenges related to signaling pathway involved). What to do when there is a large difference in sensitivity between binding antibody assays and NAb assays?

4. Emerging technologies in LBA: which technologies should be considered established and which should be classified as emerging? Which applications are in greatest need for ultrasensitive technologies? What are the criteria that should be considered in the evaluation

and implementation of emerging technologies in order to improve the science and technology and provide enhanced capabilities for R&D in supporting pharmaceutical R&D? Recommendations on how to develop a culture and infrastructure that facilitates the ongoing development and implementation of emerging technologies?

5. ISS for biotherapeutics: is it necessary?

## Hybrid LBA & LCMS discussion topics

1. Antibody–drug conjugates (ADCs) – 'the best of LBA and LCMS approaches': how to validate assays for heterogeneous, dynamically changing analytes and how many assays are needed for PK (ADC conjugate [antibody-conjugated drug or conjugated antibody], total antibody [ADC, partially deconjugated or fully deconjugated], total drug [antibody-conjugated drug + unconjugated drug] unconjugated drug [D], antitherapeutic antibody)? What analytes should be measured? What material should be used as standards for dynamically changing analytes? What are the best strategies for heterogeneous ADCs? What is recommended for ADC immunogenicity assessment? How to assess drug/antibody ratio changes *in vivo*?

2. Biomarkers validation: what is the industry standard for exploratory versus regulatory decision-making biomarkers? Focus on precision versus accuracy – have the industry and regulators reached an agreement? How best to address the critical issue of biomarker sample stability? What are the strategies for biomarker evaluation in the absence of reference standards? For selectivity assessments, what value is added by spiking recombinant/purified protein into biomarker samples containing endogenous analyte? What are the best strategies to validate LLOQ for biomarkers with high endogenous levels? What are the recommendations to overcome method development challenges with commercial kits for PK studies?

3. What are the present industry standards in the analysis of large molecules by LCMS? What are the recommendations on cross-validation of LBA with LCMS for regulatory submissions (LBA orthogonal method)? What are the recommendations on the use of LCMS for immunogenicity (LBA orthogonal method)? What is the best strategy to employ when the results obtained from both methods are not comparable?

## LCMS discussions, consensus & conclusions

### ■ 1 ISS

The topic of ISS, first introduced in the 2012 6th WRIB White Paper [5], was again discussed in the small molecule session, since some participants still have concerns regarding ISS and its applicability in regulated bioanalysis. The relevance and value of conducting stability assessments with study samples beyond what is inherent to the well-established ISR experiment was extensively debated. The consensus of the audience was that ISS should not be included as a regulatory requirement, since the vast experience of industry with respect to bioanalytical method performance supports the standard use of stability QCs to satisfactorily demonstrate stability of an analyte. However, there are examples where it can be scientifically postulated that stability of analyte(s) may be influenced by other molecular entities present in patient-generated samples or as an unintended consequence of the bioanalytical measurement itself. Metabolite instability is the primary cause of variance to spiked-matrix QC sample stability (i.e., the typical stability assessment). ISS evaluation may be indicated when previous drug metabolism or preclinical studies (*in vitro* or *in vivo*) have been conducted and the results are available. In such circumstances, it may be appropriate to take a proactive stance to avoid subsequent sample analysis inaccuracies by employing an appropriate ISS evaluation.

Once the decision to conduct an ISS evaluation is made, the issue of determining the 'Time = 0' concentration surfaces. ISS is a relative assessment as it is practically impossible to obtain a true Time 0 value [5]. Consequently, it has been agreed that the best approximation is to consider the first analysis of sample as the Time 0 value and subsequent determinations are thus an evaluation of the relative stability.

If an assessment of ISS confirms a potential stability issue, a more extensive investigation should be considered to evaluate any impact upon sample analysis accuracy and establish options for corrective actions. Any resolution derived from such an investigation should be appropriately tested to ensure repeatability and applicability to the bioanalytical method.

Although there are no industry standards for ISS and there are practical challenges in accurately performing this test, it is still recommended to

[27]Health Canada TPD, Ottawa, ON, Canada
[28]Janssen R&D, Springhouse, PA, USA
[29]Merck Research Laboratories, Boston, MA, USA
[30]Pfizer, San Diego, CA, USA
[31]AIT Bioscience, Indianapolis, IN, USA
[32]ANVISA, Brasília, Brazil
[33]Sanofi, Framingham, MA, USA
[34]UK Medicines and Healthcare products Regulatory Agency (MHRA), Welwyn Garden City, UK
[35]Tandem Labs, Salt Lake City, UT, USA
[36]Dutch Medicines Evaluation Board, Utrecht, The Netherlands
[37]Merck Research Laboratories, West Point, PA, USA
[38]AbbVie, North Chicago, IL, USA
*Author for correspondence:
Tel.: +1 450 973 6077
Fax: +1 450 973 2446
E-mail: fgarofolo@algopharm.com

**Key Terms**

**Incurred sample stability:** Reanalysis of a portion of incurred samples over a given period of time to determine whether the analyte is stable and concentrations are reproducible.

**Fit-for-purpose:** Validation of a bioanalytical method with scientific rigor for the intended purpose where all applicable parameters may not be evaluated as per regulatory guidance.

**Key Term**

**Blood microsampling:** Sampling of blood volumes small enough (less than 1% of blood volume/24 h period) to enable collection of the desired number of samples from a single subject without any measurable negative effects.

include the evaluation of potential instability of study samples in the design of a bioanalytical program in selected cases as described above.

## ■ 2 Use of incurred samples for metabolite testing & specificity during method development

Incurred study samples are used beyond the initial analysis to evaluate analytical repeats, dilution repeats and ISR. However, using incurred study samples for other purposes such as metabolite testing and/or the evaluation of specificity during method development has generated some interesting views from the LCMS bioanalytical community. A survey was recently conducted among the members of the GCC, which confirmed that using incurred samples is a common practice for certain specific applications where spiked QC samples cannot be used, such as establishing the analyte concentration range for the assay, verification of metabolites and preliminary stability. Moreover, the use of predose samples to determine potential interferences due to the presence of concomitantly administered medications was considered useful. Finally, the use of incurred study samples to assess levels of endogenous biomarkers is common practice in establishing such bioanalytical assays. GCC is planning to publish the results of this survey following the positive comments received at the 7th WRIB.

There are obvious concerns associated with any replicate analysis of study samples that may be construed as opening up potential for misuse. Any analysis of study samples in bioanalytical strategy should only contribute to ensuring accuracy and confidence of the resulting data. Approaches that may predetermine values or replicate previously obtained and accepted analyte concentrations in study samples are of particular concern. Pooling samples, an auditable process for tracking samples/identity and SOP driven procedures around study sample usage are all critical in the scientific justification and use of study samples. For clinical studies, it is also important to ensure that the study patients/volunteers agree with any further use of the study samples if this goes beyond just obtaining accurate and precise analyte concentrations in the biological matrix samples collected.

Since preclinical samples are more readily available than clinical samples, using these samples judiciously could help alleviate some problems that might be encountered in the analysis of clinical samples. Similar to clinical study samples, in this case for GLP animal studies, all intended uses of the samples are required to be included in the protocol. Although it would be most desirable to obtain clinical samples that better represent actual samples, most often these samples are not available. One possibility of obtaining clinical samples would be to conduct small pilot studies. It is believed that these small pilot studies may be approved by the IRB based on the benefit-to-risk approach. Another approach to obtain clinical samples for method development would be to pool samples from multiple studies. Alternatively, obtaining samples from volunteers or patients that are already on the relevant medication could also be explored.

The benefit of using incurred study samples as part of method development is ultimately to improve the quality of data obtained from the analytical methods used in the eventual analysis of subjects. However, there is the need to address both ethical issues, as well as any regulatory concerns in this area. Consensus from this workshop was that quality of bioanalytical methods and drug-development timelines could all benefit from appropriate flexibility associated with use of incurred study samples in method development.

## ■ 3 Fit-for-purpose validations

The term 'fit-for-purpose' for bioanalytical method validation (BMV) has been a topic of extensive discussion in recent years. The definition was clarified as part of previous meetings and publications [7–9], including previous WRIB editions [3,4]. A fit-for-purpose approach is applied when the assay does not fully comply with all current regulatory guidance requirements, but still has scientific and technical validity. Such an approach is typically employed in situations where the type of assay presents inherent difficulties and limitations, such as biomarker assays, tissue analysis, and early-stage discovery studies. Assay optimization could progress using a tiered approach and flexible methodology depending on the development stage of the assay, with increasing compliance to a full validation as the drug transitions from early discovery to late development. It has been confirmed that fit-for-purpose BMV is rarely or not applied to traditional LCMS small-molecule regulated bioanalysis. The main challenge encountered when applying a fit-for-purpose BMV resides in whether the data generated for a given study will be accepted by regulatory agencies, although

prospectively established acceptance criteria were applied based on scientific rationale. In the context of regulated bioanalysis, an assay (independent of its technical difficulty) should normally be fully validated if the data will be used for decision-making purposes such as a product label claim or supporting clinical safety assessments. On the other hand, complete validations may not be mandatory for exploratory goals or when providing supportive evidence. As a general principle, the extent of the validation of an assay should be in line with the end use of the bioanalytical data generated, and should be adequate to support the decision based upon this data. Whether it is a full validation or a fit-for-purpose validation approach, the purpose of an assay is to demonstrate that quantitative measurements generated under specific assay conditions will yield accurate and precise determinations.

It has been confirmed that fit-for-purpose approach is well-established for biomarker assay. In this specific field, it should be noted that, in quantitative measurements using a fit-for-purpose approach, the use of QC samples may not always be necessary, as it depends on the development stage at which the assay is applied. For instance, QC samples may not be needed as part of a biomarker screening assay. However, for safety and efficacy assessments where the biomarker is the end point, QC samples are necessary to confirm assay accuracy and reliability. When used, QC samples should meet the following two requirements:

- Be of known concentration (either by spiking with known amounts of reference standard or by performing repeated measurements);
- Be representative of the incurred sample matrix as closely as possible. In relation to the latter requirement, parallelism constitutes a critical parameter to assess in order to allow the use of a different matrix for calibrators and QC samples in endogenous biomarker assays.

When current regulatory method validation performance specifications cannot be met, acceptance criteria in fit-for-purpose validations should account for sufficient accuracy and precision, and should take into consideration the dynamic response range of the biomarker being measured. In other words, the extent of the biomarker change anticipated in the study will help define the validation acceptance criteria. In the absence of a reference standard, a statistical approach may be employed to establish suitable assay acceptance criteria. In most instances, biomarker assays are developed and validated for their applications in mechanism of action/efficacy studies. As previously mentioned, a complete validation is normally needed for safety assessment, depending on the business decision made with the results. Furthermore, biomarker assays applied for efficacy assessments starting from Phase IIb, should be as close to a full validation as possible. Although challenging in nature, it is possible to validate assays for endogenous analytes in accordance with current regulatory guidance. That being said, a well-implemented fit-for-purpose approach based on the intended use of the assay is expected to be positively received by regulatory agencies when the limitations to the validation are scientifically justified.

## 4 DBS

The topic of DBS has been thoroughly discussed in the last few years [3–5]. Recent advances and a better understanding around the underlying fundamentals of the DBS technology indicate that the present regulatory challenges [10,11] will be overcome in time. While there are still many hurdles, confidence was expressed that innovative solutions will be found. Also, **blood microsampling** technology still presents important benefits, which include improved PK/PD data by enabling a complete sampling profile to be collected from the same animal thus significantly reducing the number of animals (rodents) [11], smaller sample volume that makes the technique more favorable for pediatric study support, simplified sample handling and storage, increased safety by means of deactivation of bacteria and viruses due to coating materials, and the possibility of self-sampling.

IQ Consortium Microsampling Working Group, sponsored by the industry, has dedicated considerable efforts towards elucidating a better understanding of the critical factors that can lead to potential issues during DBS bioanalysis. An industry consensus has been reached on the major recommendations from the IQ Consortium Microsampling Working Group, but there are still some intricate differences between analytical laboratories.

There seems to be cautious optimism from regulators around the technology, where the majority of the concerns focus around hematocrit, differential recovery from DBS, and concordance between wet and dry samples. Therefore, in a clinical setting, dual sampling is needed to meet regulatory requirements (concordance must be shown between wet and dry matrices). Different

approaches are being used to monitor and determine concordance; typically these should be completed in the intended population and be time-matched samples at relevant doses. Sparse sampling paradigms have been found acceptable, however these approaches should only be considered after discussion and agreement with regulators. Moreover, the IQ Consortium Microsampling Working Group has devised several scenarios to demonstrate concordance between wet and dry samples of the same matrix, as well as sampling paradigms between plasma and blood matrices. As the technology matures, it is hoped that this dual sampling requirement will be lifted. For now, it was acknowledged that, for acceptance of this technology in regulated bioanalysis on human samples, there must first be a critical mass of evidence supporting its use. At present, DBS can be implemented in a 'fit-for-purpose' manner where appropriate; however when used in a regulated environment, particularly in a clinical setting, close communication with health authorities is needed. There was consensus that this technology continues to mature and strengthen from a scientific and compliance point of view within the industry.

## 5 Issues regarding MIST

Several important lessons in metabolite analysis have been learned over the years in order to meet the FDA MIST guidance. The elements in metabolite analysis that were worth considering for MIST include the following: use of incurred *in vivo* samples whenever possible for method development; comparing *in vivo* metabolite profiling across species prior to method validation to assure that the method development from one species is sufficient for another; assessing the risk for potential issues in MIST early and remaining alert for *in vitro*/*in vivo* metabolite profiling differences.

In the bioanalytical field, a tiered approach is often employed in order to obtain relative exposure data in animals versus humans for MIST risk assessment in early drug development. In this approach, a preliminary evaluation is first performed using an LCMS method, which is able to provide both metabolite MS fragmentation patterns for structure confirmation and the relative quantitation in animals versus humans. At this point, samples pooled by AUC as per the Hamilton approach could be used [12]. The next tier is to quantitate the metabolite exposure using a qualified method if the relative metabolite exposure in human versus animals is higher from the preliminary evaluation. A qualified method is defined as a method with an adequate level of validation to allow making scientific judgment based upon the concentrations of the metabolites [2]. Ultimately, a fully validated method will be required if the metabolite fulfils the MIST criteria for further testing.

There was a debate in the industry as to whether or not *N*-glucuronide metabolites should be included for MIST, given that they are Phase II metabolites and are not acyl-glucuronides. There was consensus at the WRIB that *N*-glucuronide metabolites should be included for MIST if they are deemed to be reactive. There are quite a few approaches for metabolite reactivity determination: covalent binding of the proteins is one approach (a protein binding experiment where noncovalent binding is eliminated) and chemical trapping (with glutathione or potassium cyanide, for instance) is another. This work should be performed in close collaboration with biotransformation scientists.

## 6 Evaluation of whole blood stability

The evaluation of whole blood stability has been a topic of considerable interest in the past [2,4,6]. Although this evaluation is performed in the industry, some unresolved method development issues regarding this stability assessment still exist. The most common way to perform this evaluation is to fortify fresh blood at 37°C with the analyte (and metabolites, if appropriate) at two concentration levels (low QC and high QC) and equilibrate. Afterwards, 'Time 0' samples are withdrawn and immediately centrifuged to harvest plasma. The remaining blood samples are maintained at the desired test conditions for an established period of time, when aliquots are withdrawn and centrifuged to harvest plasma. The plasma aliquots are analyzed simultaneously and compared to assess the stability over a desired time period (typically 2 h).

Some aspects of the evaluation of whole blood stability are still debated. One is related to the definition of 'fresh' blood. Fresh blood typically refers to nonfrozen blood collected and utilized within 1 day, but there is currently no industry consensus on the time period for utilization after collection (observed delays can range from 30 min to 5 days). Also, the most commonly used temperatures to assess stability in the industry are 4°C and room temperature, while some laboratories conduct studies at 37°C with the argument that it represents the worst case scenario. There was consensus that the evaluation

of whole blood stability should be conducted in clinical studies, but not everyone agrees that it should be performed for preclinical studies.

A consensus was also reached on these remaining issues discussed among the attendees. For small molecules, it was suggested that evaluation of whole blood stability should be performed the same day in which the whole blood samples were spiked. However, if the necessary plasma stability was demonstrated, the plasma samples generated by the spiked whole blood samples can be extracted at later times, if needed. As for the number of donors to be tested, a single donor representative of the study population is generally sufficient, but it was acknowledged that it would be best if multiple donors were used for some specific studies (evaluated on a case-by-case basis). Also, it was recognized that the blood source should match the intended population when appropriate. For studies meant for a multi-ethnicity population, blood collected from the predominant ethnicity of that population should be used. For pediatric studies, blood from children should be used as many studies have demonstrated that enzymatic activity is different for the pediatric population compared to adults. However, blood from regular donors may be used for studies in renally and hepatically impaired populations.

## 7 Overcoming nonspecific binding

Urine represents a matrix that can be difficult to work with for a number of reasons, including low levels of proteins or lipids, the presence of urinary salts that may precipitate at different temperatures, and the wide pH range (4–10). One of the most common challenges observed (and often incorrectly addressed during development of urine-based assays) is the nonspecific binding or adsorption of the analyte to the container surface [13]. Several factors may contribute to this phenomenon, such as the analyte physicochemical properties (hydrophobic compounds are more prone to binding to container walls), the type of container used and the sample collection procedures employed. One should be aware that other low protein matrices, such as cerebrospinal fluid (CSF) and bronchial lavage fluids, are also prone to this problem.

For small molecule assays, nonspecific binding in urine is typically detected via serial dilutions or sequential transfers in the same type of container where nonlinear analyte responses would then be observed. To overcome nonspecific binding for small molecule urine assays, a generally effective and widely used solution consists of adding plasma or serum albumin, such as bovine serum albumin (BSA), to the collection container to create an environment more similar to plasma and help to prevent analyte adsorption to the container by the protein blocking the surface and/or allowing the analyte to bind to the protein instead. Other anti-adsorptive agents, such as surfactants and organic solvents, can also be used to prevent adsorption.

In general, sample collection for urine (and some other fluid matrices) is often not fully controlled, and data obtained from urine assays are rarely used as the primary end-point. Consequently, most urine assays are qualified or validated using a fit-for-purpose approach. However, if conducting a rigorous urine method qualification/validation where urine results are used as primary data, it is of crucial importance to thoroughly evaluate all aspects of sample collection early in the method-development process. Appropriate anti-adsorptive agents should be chosen and proper sample-collection procedure should be shared with the clinical sites to avoid bias in measured analyte concentrations.

The analysis of peptides is also prone to nonspecific binding to a container issues, as peptides (especially those with many uncharged hydrophobic residues) have been shown to bind more readily to various materials than small molecules due to their size and physicochemical properties. Binding can occur throughout the analytical process; for instance when pure solution is used, during the extraction process, and even during chromatography [14]. As in the case of urine analysis, the addition of BSA is often an appropriate and practical solution to reduce the nonspecific binding for peptides. Care should be taken during method development in selecting appropriate options in terms of solution additives, container, and handling procedures. Although the use of additives is a common approach, one has to keep in mind that some additives may cause significant matrix effects. Among the options that could be used are the dilution of matrix with acetonitrile 1:4 v/v, the addition of BSA, the addition of acid (in combination with other options or not), and the addition of surfactants. The use of low-binding plates is highly recommended when working with peptides and proteins. Polypropylene, polystyrene and glass are most commonly used materials in collection devices. The use of polyethylene should be avoided. Another important procedural aspect to consider is avoiding the use of serum separator

tubes made with silica gel. Also, injection of a matrix-free peptide stock onto an LCMS system in some cases can result in prolonged system contamination from strongly adsorbed analyte and should be considered case-by-case.

Whenever sample matrices were modified to prevent or mitigate the potential nonspecific adsorption of the analytes of interest, one must keep in mind that other potential bioanalytical parameters should also be considered, especially the degradation of Phase II metabolites to the parent/Phase I metabolites to avoid overestimation of the analyte concentrations. The stability of metabolites themselves should also be under consideration if these metabolites need to be measured.

### ■ 8 Hyperlipidemic matrix test

The first mention of a regulatory recommendation to include a hyperlipidemic matrix lot during the validation of an analytical method in plasma or blood was from ANVISA back in its 2003 ANVISA Manual for Bioavailability and Bioequivalence Practices [15], where a hyperlipidemic matrix was to be included as part of the selectivity assessment. Later, the EMA introduced hyperlipidemic matrix as part of the matrix effect evaluation in its draft Guideline on BMV in 2009 [16], and then in the final version of the Guideline issued in 2011 and effective since February 2012, where it is stated that *"In addition to the normal matrix it is recommended to investigate matrix effects on other samples e.g., haemolysed and hyperlipidaemic plasma samples"* [17]. ANVISA also included this recommendation in its new BMV guideline issued in May 2012, Resolution RDC no. 27 as part of both selectivity and matrix effect testing [18]. The FDA has also started focusing its attention on this topic during inspections.

The need for hyperlipidemic matrix testing in BMV can be justified by the potential presence of matrix effects that could be caused by the presence of various lipids in the samples to be analyzed, considering that the levels of lipids in the study samples may significantly vary between subjects, and that some subjects may show naturally high lipid levels, even in fasted studies. Using a stable isotope-labeled (SIL) IS often compensates for this potential effect, but it may not always be the case especially at concentrations close to the LLOQ and if the ion suppression effect is significant. Thus, testing of hyperlipidemic matrix as part of chromatographic analytical methods is considered useful and a good practice in demonstrating that the presence of lipids does not affect the performance of the assay.

Although the hyperlipidemic matrix testing is included in EMA and ANVISA regulatory guidance documents, the definition of a hyperlipidemic sample in the context of method validation is not clearly established. The ANVISA guideline defines a lipidemic sample as a *"high lipids degree sample, for example, coming from post prandial collection"* [18]. As a general rule, to be scientifically meaningful in BMV, the hyperlipidemic matrix test should be representative of the samples destined to be analyzed with the method, by taking into account the type of lipidemic samples encountered in clinical or preclinical studies, as well as the expected approximate degrees of lipidemia.

In light of these considerations, the most logical and appropriate type of hyperlipidemic matrix to use in BMV would represent a naturally lipidemic matrix obtained from donors with abnormally high levels of triglycerides (either consistently high-level donors or donors following a high-fat meal).

## LBA discussions, consensus & conclusions

### ■ 1 Importance of parallelism in LBA

Parallelism assessments are performed to evaluate whether the sample dilution–response curve is parallel to the standard concentration–response curve. In this analysis, a plot of the measured concentrations of the analyte at multiple dilutions against the expected concentrations at each dilution should have a slope close to 1.0. The experimental methods used to evaluate parallelism are therefore similar to the dilutional linearity assessment performed during pre-study validation, except that parallelism is assessed with multiple dilutions of incurred study samples. While an assay may have proven dilutional linearity, some incurred samples may contain interferents/binding proteins or other factors that affect interaction of the analyte of interest with the assay critical reagents. While a sample result may not be invalidated due to non-parallelism, it should be noted that potential interferents in such samples may be affecting the relative accuracy of the result.

For PK assays, evaluation of parallelism should be considered when issues are anticipated due to the nature of the molecule, disease indication or patient population. For example, parallelism may be employed to verify analyte

stability, examine biotransformation or understand patient specific matrix effects. However, parallelism should not be routinely included in the validation of PK assays but rather the need to do so should be evaluated on a case-by-case basis and parallelism performed when necessary or relevant. The decision to perform parallelism or not should be driven by scientific rationale [19].

For biomarker assays, where the analyte of interest is an endogenous compound, it is always necessary to evaluate parallelism between the recombinant/purified standard calibrator and the endogenous analyte. Since samples containing endogenous analyte can be obtained during the assay development, stage parallelism should be assessed at that time to inform assay optimization and enable an early understanding of the level of decision making that the resulting sample data will support. The parallelism assessment results should then be confirmed during assay validation [9].

Parallelism assessments can also be leveraged to inform other parameters for biomarker assays. For example, selectivity, the ability of an assay to differentiate and quantify the analyte in the presence of other components in the sample can, and should, be assessed using a parallelism approach whenever samples are available that have adequate levels of analyte to enable testing of multiple dilutions. In cases where endogenous levels of analytes are too low to enable testing of multiple dilutions, then a spike recovery approach may be necessary. In this case, it should be noted that the spiked material will be recombinant/purified calibrator material and, therefore, not the same as the endogenous analyte that will be measured in study samples. While this assessment might provide some additional level of confidence that the recombinant/purified standard calibrator material and the endogenous analyte have similar binding to the critical reagents, it is not as informative with respect to the assay's selectivity towards the endogenous analyte as the parallelism approach. For exploratory biomarker assays, an assay would be considered selective if parallelism passes, even if a spike-recovery assessment fails. For assays supporting late-phase markers where regulatory agency decision making is involved, however, acceptable parallelism coupled with unacceptable spike-recovery would warrant some investigation to understand the spike-recovery failure, but would not mean that the assay could not be validated.

For both PK and biomarker assays, the current working method of 30% CV is acceptable, although not perfect. The industry continues to evaluate alternative criteria. In the interim, when applying the 30% criterion, data should be carefully monitored as results that pass this criterion may still reveal trends of non-parallelism.

There was limited discussion at the 7th WRIB on the hook effect. Further discussions on the hook effect and its impact on assay development are needed and the responses in this White Paper are limited to the specific questions raised. Although questions related to the hook effect were raised in the parallelism section, there is no direct correlation of the two concepts. The main difference being that the hook effect is mainly an assay issue while lack of parallelism is primarily attributed to analyte or incurred sample composition issues.

The hook effect should be investigated during method development using a highly concentrated QC sample that is diluted into the assay range. If the higher concentration samples return final concentrations less than the actual values, a hook effect is suspected. The hook effect cannot always be controlled. However, assay performance can be optimized during assay development to reduce the risk, for example by optimizing the coating concentration. Some assay formats may be more prone to having the hook effect. For example, homogeneous assays where the reagents are in excess and there are multiple epitopes on the analyte have a higher probability of having a hook effect [20].

In rare instances, in-study samples expected to have very high concentrations of analyte may return anomalously low values, suggesting the possibility of an in-study hook effect, even when no hook effect was observed with QC samples of even higher concentrations during pre-study validation. In these cases, a scientific investigation is warranted and performance of a parallelism-type assessment on the suspect samples may prove informative since additional dilution can drive dissociation of interfering complexes that may exist in the sample.

## 2 Immunogenicity & effect on PK assays

PK data interpretation for large molecules is more complex than PK data interpretation for small molecules due to a large number of factors, including immunogenicity. Generally speaking, when trying to determine the effect of ADAs on PK assays, a data-driven approach should be implemented as not all abnormal PK is inherently due to ADA. First and foremost, the assessment of ADA with respect to PK is used to

determine the potential effects on exposure and safety of the compound. For preclinical studies, the measurement of ADA is primarily used to assess exposure: it is acceptable to run samples in a screening assay only. However, in the absence of a confirmatory analysis, this may increase the number of positive results being reported. For clinical ADA assays, where interpretation will also have safety implications, it is expected that screening, confirmatory and titration assays should be conducted.

In order to enable appropriate interpretation of study exposure data, an evaluation of the ADA interference in the PK assay during assay development can be done, just as it is important to know the drug tolerance level for the ADA assay [21]. ADA interference in the PK assay can be evaluated using the positive control to inform potential *in vivo* interference. It is also important to understand the specificity of the capture and detection antibodies employed in the PK assay when trying to interpret the effect of the ADA results on the drug concentration. For example, for monoclonal antibody (mAb)-based therapeutics, one can measure free (unbound) or total (unbound and bound) concentrations of the drug. Because the PK assay reagents will determine the nature of the analyte detected (unbound versus bound mAb), the extent of impact of ADA on the PK assay will be method-dependent. For preclinical studies, many toxicologists are more interested in the total drug concentration for safety assessment and determining exposure in first-in-human. However, understanding the free drug, which is the biologically active form, in clinical assays is critical when interpreting the effect of binding ADA [22].

When an unexpected PK profile is observed, a risk-based, data-driven approach should be applied when conducting investigations, as not all abnormal PK results will require a thorough investigation. It is advisable to consider the expected drug therapeutic window during PK assay method development to guide the appropriate amount of effort applied to examining ADA impact on assay performance.

## ■ 3 Immunogenicity & neutralizing assays

Assessment of neutralizing potential of anti-drug antibodies is a necessary and important component of the tiered approach to immunogenicity assessment in clinical studies. The use of cell-based versus noncell-based NAb assays has been a hot topic of discussion among industry scientists and regulators in recent years. The development and validation of NAb assays should be determined using a risk-based approach [23]. For high immunogenicity risk biotherapeutics, a cell-based NAb is commonly developed prior to first-in-human. However, for low immunogenicity risk biotherapeutics, while cell-based NAb may still be desirable, alternative assay formats may be explored. In addition, other criteria may impact whether a cell-based NAb assay is necessary or appropriate. For instance, the mechanism of action for the molecule should be considered. If cell signaling is involved, then cell-based assays are recommended. However, if the cell-based NAb is simply a cell-based binding assay, it is less likely to be any more informative than a noncell-based assay. Scientists and regulators should bear in mind that the purpose of the NAb assay is to assess the neutralizing potential of detected antibodies. Therefore, if a cell-based NAb assay has a large difference in sensitivity compared to the binding antibody assay, then the cell-based assay may not serve the purpose in that most of the detected binding antibodies may not be appropriately evaluated due to the inherent sensitivity differences. In these cases, it may also be erroneously assumed that the detected binding antibodies are not neutralizing. Therefore, if adequate sensitivity cannot be achieved with a cell-based approach, then adapting the assay to other platforms becomes a necessity and may be the ideal option in order to comprehensively assess the overall immunogenicity profile of a given therapeutic. It was recommended that the regulatory authorities be engaged to on a case-by-case basis to discuss the impact of switching to a noncell-based assay.

## ■ 4 Emerging technologies in LBA

There is much discussion on how to define what qualifies as an emerging technology. Is it defined in the context of the application? Has the technology been reviewed by the regulatory authorities? Is it an established technology with a novel application? Some believe that pre-commercial prototypes should be considered emerging, such as Quanterix™, a technology for miniaturized digitized ELISA that is not yet on the market. However, others are of the opinion that it should not include prototypes as most of these will not make it into widespread application.

While these questions are still up for debate, there was consensus that established technologies used for novel applications should be

considered emergent. For example, developing an assay on the Singulex® platform, but using home-grown reagents and kits might be considered novel. Using LCMS for large molecule bioanalysis may also be considered emergent, because although LCMS is a long established technology, LBAs have historically been the application of choice for quantification of large molecules. While LBAs are standard for mAb and proteins, the need for new ultrasensitive technologies has increased in recent years for certain applications. Biomarker and hormone assays, in particular, could benefit from alternative methods of detection. In addition, special applications of PK such as when drug measurements are required in special matrices (e.g., CSF), that may be volume limited or have extremely low levels of analyte, could benefit from new approaches. However, emerging technologies should be implemented where appropriate and where they provide some benefit above the standard practice, whether in cost, throughput, or assay sensitivity.

In order to develop a culture conducive to promoting new technologies, it may be desirable to dedicate a group to this purpose. This, of course, is only possible if the resources can be made available without taking a cut elsewhere. Another way this may be achieved is by sharing data between groups when possible and where appropriate.

### ■ 5 ISS for large molecules

ISS should not be required as a standard routine test. It should serve to bridge a possible gap between spiked and incurred samples, when deemed necessary based on the physicochemical and/or biotransformation properties of the analyte.

## Hybrid LBA & LCMS discussions, consensus & conclusions

### ■ 1 ADCs

ADCs combine the specificity of a mAb with the potency of a chemical drug (or payload) producing a highly specific therapy. Both LBA and LCMS are used for the bioanalysis of ADCs. The molecular structure of ADCs is composed of three components: the mAb, the linker and the low molecular weight cytotoxic drug. One characteristic of these molecules is that they are highly heterogeneous and the heterogeneity can change *in vivo*, which therefore necessitates three key assays for bioanalysis. These assays include: ADC conjugate (antibody-conjugated drug or conjugated antibody), total antibody (conjugated, partially deconjugated and fully deconjugated mAb) and unconjugated drug (the free cytotoxic drug) [24,25]. The number/type of assays required for these compounds will depend on the study goals. It is possible that some analytes may relate to efficacy while others may relate to safety.

During the initial evaluation, more assays may be required to characterize the ADC PKs. However, the nature of the analytes to track during later clinical studies will be determined by the exposure type (e.g., ADC conjugate, unconjugated drug) that provides the signals that relate to the clinically relevant readout. A qualified assay will be sufficient to support discovery efforts, but the relevant assays are expected to be fully validated before they are utilized in support of regulated nonclinical and clinical studies.

When developing an assay, the reference standards used may vary. For total antibody, conjugated antibody and conjugated drug assays, the ADC should be used as the reference standard. For the unconjugated drug assay, the reference standards for the drug should be used.

### ■ 2 Biomarkers validation

Biomarkers are the measures of biological, pathological or pharmacologic processes. PD markers are a class of biomarkers that include proteins, nucleic acids or metabolites that are expressed in a target population and can provide the evidence that a drug hits its target to exert functional change. Biomarker data may be collected for information use only, efficacy or safety purposes. The clinical application determines whether the assay will need to be fully validated. For exploratory biomarkers, typically conducted in early Phase I with limited human exposure and no diagnostic intent, a fit-for-purpose approach is recommended. However, a full assay validation will usually need to be performed for biomarker assays employed in late-phase clinical trials to inform regulatory decision making and patient stratifications [26].

When developing/validating these biomarker assays with an LCMS, LBA or a hybrid LBA/LCMS approach, there are a number of issues that need to be addressed such as choice of precision versus accuracy, sample stability, determination of LLOQ for endogenous analytes, and imperfect commercially available kits.

Regarding precision versus accuracy, precision is the recommended choice. Accuracy of

**Key Term**

**Proteotypic peptides:** Peptides sequences that are found in only a single known protein and therefore serve to identify that protein.

endogenous molecules is very difficult to assess given the molecular differences between the synthetic calibrator and the analyte of interest. While absolute accuracy is not practically attainable, relative accuracy and thorough characterization of what is being measured should be evaluated. However, although there was a strong alignment on this point among representatives from both industry and regulatory agencies at the 7th WRIB, no official agreement has been established.

The evaluation of sample stability for a biomarker is dependent upon the availability of the calibrator standard and/or pooled samples that have relevant (high and/or low) levels of the endogenous analyte. The preference is to use incurred samples, although in some cases these samples may need to be fortified with recombinant or purified standard material to create higher concentrations of analyte in the samples. In the absence of a true calibrator, the practical strategy is to use what is available. Since the vast majority of biomarker assays are exploratory, commercially available calibrator material can be used.

In order to determine the assay LLOQ as it relates to the endogenous analyte (as opposed to the recombinant or purified calibrator standard material), samples with adequately high concentrations of endogenous analyte are required. When these samples are available, parallelism data can be used to identify the most conservative dilution that demonstrates parallelism across multiple samples. The LLOQ is set at the lowest concentration accurately measured at that dilution. Alternatively, the lowest concentration for each individual sample that demonstrates parallelism can be identified and the LLOQ set by the sample with the highest measured concentration. When samples with sufficiently high concentrations to enable a parallelism assessment are not available, then the provisional LLOQ will need to be set using the standard calibrator material.

Commercially available kits can be validated. If the kit is imperfect, changes can be made to enable validation. For example, the kit may be used only as a source of critical reagents with the assay being fully optimized independently. The data provided in the kit insert should not be used as a surrogate for the assay validation.

## ■ 3 Large molecule bioanalysis by LCMS

The quantitative analysis of large molecules by LCMS in a regulatory environment is a relatively new application that generates intense discussions within the bioanalytical community. The definition of a large molecule itself cannot be simply correlated to a molecular weight cutoff. From a mass spectrometric point of view, molecules with a molecular weight of approximately 1000 Da or above can generally be considered large molecules based on their propensity to generate multiple charged ions (i.e., charge states) in electrospray. Indeed, multiple charged ions can significantly impact sensitivity, selectivity and method-development strategies to generate reliable bioanalytical methods. Industry standards in quantitative large molecule bioanalysis by LCMS are still evolving as they depend on choice and technical details of the method. Assay strategies should be determined in a fit-for-purpose manner, depending on the pharmaceutical development stage and the information needed. Despite a flow of ideas, some clear agreements were established regarding the industry standards.

The acceptance criteria for the quantitation of a large molecule may resemble those applied to LBA, especially in the case where hybrid LBA/LCMS approaches are used, such as immunocapture, or if sample preparation is highly complicated, which increases the variability of the assay. However, if supported by the demonstrated analytical method performance during validation, the acceptance criteria can be more closely aligned to the LCMS criteria. Due to recent improvements of the latest generation of high resolution (HR) MS instruments, instruments like the quadrupole TOF and Orbitrap™ are gaining more and more interest for intact protein analysis, especially in the drug discovery setting. Regarding multiply charged ion states that are characteristic to protein/peptide ionization, it is fundamental to select the relevant charged ion state and, if HRMS is used, for each charged ion state to select the most abundant isotopic peak or the summing of multiple isotopic peaks for the quantitation. If HRMS instrument sensitivity allows it, the quantitation of the intact protein (or large peptide) is always preferred since it eliminates the inherent variability that may be introduced when applying enzymatic digestion to produce **proteotypic peptides** for measurement as surrogates for the intact molecule [27,28]. A wide variety of sample enrichment techniques can be used such as immunoaffinity enrichment, SPE, LLE, and direct dilution, although the choice of such techniques will depend on the analyte and the desired sensitivity. The method of choice may not only depend solely on the analyte but also on the experience of the analysts,

taking into account that sample preparation along with a robust chromatographic separation are the key requirements for a reliable quantitation method. The use of an IS is recommended in a regulated bioavailability setting. SIL intact protein is the preferred choice as it compensates for the variability related to all steps, including the digestion and enrichment processes. However, it is not usually available and the use of the SIL-flanked (containing amino acid sequence extensions beyond the cleavage site on each end that are recognized by the proteolytic enzyme used in the method) peptides may be the next best choice, followed by the SIL surrogate peptides, and finally by analogue peptides. The use of HR accurate mass or triple quadrupole instruments are both feasible, depending on the ionization, sensitivity and selectivity needed. The monitoring of multiple proteotytic peptides is recommended in early phases for confirmation of quantitation data, since similar results obtained from multiple peptide determinations will increase the confidence in the data initially obtained. The monitoring of multiple peptides may also generate additional information on PK/PD, post-translational modifications, drug mechanism of action and so on, which may be relevant to therapeutic protein's characteristics. If the bioanalytical data obtained on multiple peptides early in the process is sufficient to confirm the method's selectivity, sensitivity and precision, it is then recommended to select one specific peptide as the primary signal for reporting quantitation. The choice of the peptide should be scientifically driven and depend on the information required for the study supported.

In general, cross-validation of LBA with LCMS is not suggested or required since the parallel development of an LBA and LCMS assay strategy is not recommended. Taking overall costs for a parallel development into account, a scientifically justified selection of the most appropriate analytical strategy in a timely manner is preferred. In particular cases where data are already generated using LBA and the application of LCMS is desired for complementary and/or additional information purposes, cross-validation is necessary. Significant differences in LBA versus LCMS results for the same samples do not necessarily invalidate either set of data in general but likely indicate that the forms of the analyte measured are different. For example, LCMS methods could be designed to measure 'total' analytes, while LBA could be designed to measure 'free' analytes. Even in the case where both assay types are designed to measure 'free' analytes, the analyte measurements may still be different, as the measurement of 'free' analyte is dependent on the binding re-equilibration due to sample dilution and length of incubation. In the case of using the assays to evaluate the impact of immunogenicity on PK assays, such discrepancies should lead to a sound scientific explanation (e.g., presence of ADAs interfering with the LBA). However, generation of multiple sets of data is not recommended.

## Conclusion

Below is a summary of the 16 recommendations (eight for LCMS, five for LBA, and three for hybrid LBA and LCMS) made during the 7th WRIB.

### ■ LCMS recommendations

1. ISS should not be part of regulatory guidance in terms of experimental design or assignment of predefined acceptance criteria. However, when the potential for ISS issues is indicated by metabolism information or earlier studies, appropriate scientifically driven experiments are recommended. For overcoming the challenges of establishing the Time = 0 point in ISS, it is recommended to use the first analysis of a sample as the reference point for subsequent analyte stability assessments. If a stability issue specific to study samples is detected, then appropriate sample handling and bioanalytical procedures should be established and implemented. These procedures should accompany study sample analysis including repetition of the ISS evaluation experiment throughout the study as appropriate.

2. Using incurred study samples as part of bioanalytical LCMS method development has several important advantages but has potential for misuse and/or inconsistencies with informed consent. Hence, their use should be clearly addressed in the informed consent form and controlled by SOP. Since the use of incurred study samples in method development is closely related to data quality improvement through better bioanalytical methods, and decrease timelines and costs of the development process, their use should be encouraged but be tightly controlled by SOP and supporting documentation.

3. Fit-for-purpose BMV is rarely or not applied to traditional LCMS small-molecule regulated bioanalysis due to concerns in having the data generated accepted by regulatory agencies. However,

fit-for-purpose approach is well-established and used for biomarker assay validation.

4. The DBS technology presents important benefits and a consensus has been reached on the major recommendations from the IQ Consortium Microsampling Working Group, but at present, DBS implementation should be fit-for-purpose and with close communication with regulators.

5. A tiered qualification/validation approach should be used for MIST and the risk of potential issues should be assessed early. Also, *N*-glucuronide metabolites should be included for MIST if they are deemed to be reactive.

6. For whole blood stability evaluations, a single donor representative of the study population is generally sufficient and the blood source should match the intended population when appropriate.

7. Nonspecific analyte binding to various materials is one of the major challenges faced in urine assays as well as in peptide analysis. All aspects of sample collection, processing and chromatography must be thoroughly evaluated early in method development. Appropriate anti-adsorptive agents should be chosen and proper sample collection procedure should be shared with the clinical sites to avoid bias in measured analyte concentrations.

8. For LCMS, a hyperlipidemic matrix test is useful to assess potential matrix effects due to the presence of lipids. To be scientifically meaningful, the matrix to employ for this test should be representative as much as possible of the incurred samples expected to be analyzed with the method. The use of artificial lipidemic matrix such as plasma spiked with fat emulsion is likely not representative and should be avoided. A naturally lipidemic matrix obtained from donors with abnormally high levels of triglycerides (either consistently high-level donors or donors following a high-fat meal) should be used.

## ■ LBA recommendations

1. For PK assays, parallelism should not be expected to be routinely included in the validation. Instead, it should be evaluated on a case-by-case basis in order to determine when it is necessary or relevant based on the study sample data in comparison to QC performance. There should be a scientific rationale to justify the evaluation. For biomarker assays, where the calibrators are sufficiently different than the endogenous analyte, parallelism should be assessed in the assay development phase and the relevant information included in the validation.

2. When trying to determine the effect of ADAs on PK assays, a risk-based approach should be implemented as not all abnormal PK is inherently due to ADA. Knowledge of the expected drug therapeutic window and information regarding possible interference by ADA in the PK assay should be applied.

3. A risk-based approach should be used when determining the need to develop and validate cell-based neutralizing assays. Consideration should be given to the mechanism of action for the molecule. If cell signaling is involved, a cell-based NAb assay is recommended. If the mechanism is based on a binding event, it may be appropriate to establish a noncell-based NAb assay.

4. Biotherapeutics development can benefit greatly from evaluation of emerging technologies. In order to promote new emerging technologies, it is beneficial, where appropriate, that resources be allocated to the exploration of new promising technologies.

5. For large molecule LBAs, ISS should not be required as standard practice.

## ■ Hybrid LBA & LCMS recommendations

1. During the initial discovery phase evaluation of ADCs, many analytes will need to be tested, requiring several assays to be developed and qualified. The type of analytes to track during later nonclinical and clinical phases will be determined by the exposure type that gives the signals that provide the best correlation with the clinically relevant patient outcomes. Assays for these analytes will require full validation.

2. Changes to an imperfect commercial assay kit can be made to enable validation. For example, only critical reagents from the kits may be used to build an assay. A strong consensus was reached on the fact that the data provided in the kit insert should not be used to support assay validation.

3. The acceptance criteria for the quantitation of large molecules by LCMS should be similar to or stringent/tighter than those of LBA. The choices of instruments (HRMS or triple quadrupole), ISs (SIL-intact protein or SIL-flanked peptide), extraction techniques, and signature peptide should all be thoroughly evaluated to optimize method performance. As the parallel development of LBA and LCMS assay strategies is not preferred, cross-validation between LBA and LCMS is generally not recommended.

## Acknowledgements

*The authors would like to acknowledge the US FDA, Europe EMA, Health Canada, The Netherlands MEB, Brazil ANVISA, Japan MHLW and UK MHRA for supporting this workshop. L Stevenson (Biogen Idec Inc.) for her major contribution in coordinating this White Paper for the biotherapeutics part. B DeSilva (Bristol-Myers Squibb), L Amaravadi (Biogen Idec Inc.), M Brudny-Kloeppel (Bayer Pharma AG), A Musuku, (Pharmascience) and F Garofolo (Algorithme Pharma Inc.) for chairing the workshop. L Stevenson (Biogen Idec Inc.) and M Rocci (ICON Development Solutions) for chairing the White Paper discussions. All the workshop attendees and members of the bioanalytical community who have sent comments and suggestions to complete this White Paper. I Dumont, S Martinez, C Dicaire, L Wright, L Mayrand-Provencher, M Losauro for collecting the workshop minutes and preparing the first draft of this White Paper. W Garofolo (Meeting Coordinator) for the logistic organization of the event. Future Science Group as a trusted partner.*

## Disclosure

*The views expressed in this article are those of the authors and do not reflect official policy of the US FDA, EU EMA, Health Canada, MEB (The Netherlands), ANVISA (Brazil), MHLW (Japan) and MHRA (UK). No official endorsement by the FDA, EU EMA, Health Canada, MEB, ANVISA, MHLW or MHRA is intended or should be inferred.*

## Financial & competing interests disclosure

*The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.*

*No writing assistance was utilized in the production of this manuscript.*

## References

1 Savoie N, Booth BP, Bradley T *et al.* 2008 White Paper: the 2nd calibration and validation group workshop on recent issues in good laboratory practice bioanalysis. *Bioanalysis* 1(1), 19–30 (2009).

2 Savoie N, Garofolo F, van Amsterdam P *et al.* 2009 White Paper on recent issues in regulated bioanalysis from the 3rd calibration and validation group workshop. *Bioanalysis* 2(1), 53–68 (2010).

3 Savoie N, Garofolo F, van Amsterdam P *et al.* 2010 White Paper on recent issues in regulated bioanalysis and global harmonization of bioanalytical guidance. *Bioanalysis* 2(12), 1945–1960 (2010).

4 Garofolo F, Rocci M, Dumont I *et al.* 2011 White Paper on recent issues in bioanalysis and regulatory findings from audits and inspections. *Bioanalysis* 3(18), 2081–2096 (2011).

5 DeSilva B, Garofolo F, Rocci M *et al.* 2012 White Paper on recent issues in bioanalysis and alignment of multiple guidelines. *Bioanalysis* 4(18), 2213–2226 (2012).

6 Lowes S, Jersey J, Shoup R *et al.* Recommendations on: internal standard criteria; stability; ISR and recent 483s by the Global CRO Council (GCC) for Bioanalysis. *Bioanalysis* 3(12), 1323–1332 (2011).

7 Lee JW, Devanarayan V, Barrett YC *et al.* Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm. Res.* 23(2), 312–328 (2006).

8 Breda M. Garofolo F, Caturla MC *et al.* The 3rd Global CRO Council (GCC) for bioanalyis at the international reid bioanalytical forum. *Bioanalysis* 3(24), 2721–2727 (2011).

9 Hougton R, Gouty D, Allinson J *et al.* Recommendations on biomarker bioanalytical method validation by GCC. *Bioanalysis* 4(20), 2439–2446 (2012).

10 O'mara M, Hudson-Curtis B, Olson K *et al.* The effect of hematocrit and punch location on assay bias during quantitative bioanalysis of dried blood spot samples. *Bioanalysis* 3(20), 2335–2347 (2011).

11 Jonsson O, Villar RP, Nilsonn LB *et al.* Capillary microsampling of 25 µl blood for the determination of toxicokinetic parameters in regulatory studies in animals. *Bioanalysis* 4(6), 661–674 (2012).

12 Hamilton RA, Garnett WR, Kline BJ. Determination of mean valprolic acid serum level by assay of a single pooled sample. *Clin. Pharmacol. Ther.* 29(3), 408–413 (1981).

13 Ji AJ, Jiang Z, Livson Y, Davis JA *et al.* Challenges in urine bioanalytical assays: overcoming nonspecific binding. *Bioanalysis* 2(9), 1573–1586 (2010).

14 Ewles M, Goodwin L. Bioanalytical approaches to analyzing peptides and proteins by LCMS/MS. *Bioanalysis* 3(12), 1379–1397 (2011).

15 ANVISA. *Manual for Good Bioavailability and Bioequivalence Practices.* Brazilian Sanitary Surveillance Agency, Brazil (2003).

16 EMA, Committee for Medicinal Products for Human Use. *Draft Guideline on Validation of Bioanalytical Methods.* London, UK (2009).

17 EMA, Committee for Medicinal Products for Human Use. Guideline on bioanalytical method validation. London, UK (2011).

18 ANVISA. Resolution RDC no. 27, May 17, 2012. Brazilian Health Surveillance Agency, Brazil (2012).

19 DeSilva B, Smith W, Weiner R *et al.* Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm. Res.* 20, 1885–1900 (2003).

20 Rodbard D, Feldman Y, Jaffe ML *et al.* Kinetics of two-site immunoradiometric (sandwich) assays-II. *Immunochem.* 15, 77–82 (1978).

21 Kelley M, Ahene AB, Gorovits B *et al.* Theoretical considerations and practical approaches to address the effect of anti-drug antibody (ADA) on quantification of biotherapeutics in circulation. *AAPS J.* 15(3), 646–658 (2013).

22 Lee J, Kelley M, King L *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.* 13(1), 99–110 (2011).

23 Gupta S, Devanarayan V, Finco D *et al.* Recommendations for the validation of cell-based assays used for the detection of neutralizing antibody immune responses elicited against biological therapeutics. *J. Pharm. Biomed. Anal.* 55, 878–888 (2011).

24 Kaur S, Xu K, Saad OM *et al.* Bioanalytical assay strategies for the development of antibody-drug conjugate biotherapeutics. *Bioanalysis* 5(2), 201–226 (2013).

25 Gorovits B, Alley SC, Bilic S *et al.* Bioanalysis of antibody-drug conjugates: American Association of Pharmaceutical Scientists Antibody-Drug Conjugate Working Group position paper. *Bioanalysis* 5(9) 997–1006 (2013).

26 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research. *Guidance for Industry, Investigators, and Reviewers. Exploratory IND Studies.* FDA, Rockville, MD, USA (2006).

27 Ramagiri S, Garofolo F. Large molecule bioanalysis using Q-TOF without predigestion and its data processing challenges. *Bioanalysis* 4(5), 529–540 (2012).

28 Morin LP, Mess JN, Garofolo F. Large-molecule quantification: sensitivity and selectivity head-to-head comparison of triple quadrupole with Q-TOF. *Bioanalysis* 5(10), 1181–1193 (2013).