Bioanalysis ZONE

eurofins | BioPharma Services

Defining specificity within the ELISpot assay

fsg

Contents

Technology Digest: Defining specificity within the ELISpot assay

by Thomas 'TJ' Beadnell, PhD
(Eurofins Viracor BioPharma Services)

The enzyme-linked immune absorbent spot (ELISpot) assay provides important information on the abundance of low frequency immune cells, with an emphasis on the detection of antigen-specific T-cells or B-cells present within a patient. The assay was originally developed for the detection of antibody-secreting lymphocytes using antigen-coated ELISpot and was later adapted for the analysis of antigen reactive T-cells through the coating of the antibodies to the ELISpot plate. The assay can provide critical information regarding vaccination efficacy, autoimmunity and/or rejection, and cytotoxicity.

Many different cytokines, chemokines and antibody classes have been optimized for use within this assay and have additionally been demonstrated to be effective in multiplexing. In addition, T-cells include many different subsets – the primary being the cytotoxic CD8+ T-cells and the helper CD4+ T-cells – which can be further broken down into CD4 helper subsets, the most relevant from the context of the ELISpot assay being the Th1, Th2 and Th17 subsets. Furthermore, the ELISpot assay can also consider T effector, as well as central memory populations using T-cell expansion protocols.

Many great efforts have been undertaken to help enhance our guidance around the ELISpot assay, including efforts to enhance global harmonization as well as recent recommendations from the GCC [1, 2]. However, as we look at the guidance surrounding flow cytometry, especially regarding the recent release of the H62 CLSI guidelines [3], it is important that we continue to build the established ELISpot guidelines to better inform and unify ELISpot assays and results across laboratories. This is especially true given that the specificity of the ELISpot responses is harder to gauge in comparison to flow cytometry, a technique that allows for the individual detection of cell types, whereas the ELISpot assay does not specifically delineate these populations.

As we enter an era in which even more therapies emerge for which there are limited specificity controls, it will be important to understand how each specificity assessment can guide our understanding of the assay results.

Assay specificity

Recent guidance for assay specificity highlights specificity as “the ability of an analytical method to detect the analyte of interest” or “the extent to which an assay responds to all subsets of a specified analyte” [4, 5]. For the ELISpot assay, the first statement focuses on making sure that the capture antibody along with detection reagents are specific for their

intended target. A key to accomplishing this criterion is the selection of high quality, specific monoclonal antibodies and the selection of quality-controlled kits and reagents. The second statement, referring to all subsets of a specified analyte, becomes more assay-specific and therefore is important to look at in the context of the different subsets surrounding the assay of interest.

In addition, the ICH guidance states “the procedures used to demonstrate specificity will depend on the intended objective of the analytical procedure” [6]. We propose that specificity analysis for T-cell ELISpots can be thought of as falling under three main nodes: donor, T-cell and antigen, which may have varying levels of importance based on the question the experiment is aiming to address.

The antigen determines if the responses are specific for the test peptide of interest. The donor determines background responses and donor variance based on differential haplogroup presentation. Finally, the T-cell determines the immune component of interest, providing a specific picture of T-cell activation. Although analysis will not always be limited to these parameters, they represent a useful first step in the search for specificity.

Donor specificity

For donor specificity, we describe this as a process by which we determine specificity in the assay based on unique responses from each donor. As alluded to above, it may not always be possible to find an appropriate control to measure specificity and therefore strategies need to be employed to determine responses in relation to negative vehicle control responses. Peptide responses can therefore be determined by creating a threshold minimum response following the subtraction of the background signal. In addition, defining a fold increase threshold above background can also function in the determination of specific antigenic responses. This can be further accomplished through false positivity rate-based methods for defining determination criteria [7].

Additional research has also highlighted a criterion by which responses across multiple peptides were assessed for anti-HER2 responsiveness [8]. The number of patients responding to one peptide, the mean number of all reactive peptides and the cumulative response across all peptides were quantified to gauge overall responsiveness. As T-cells may develop immunity to select peptides, it may be important to define specificity and responses based off multiple peptides. In addition, individual donor responses should vary based on the number of antigen specific T-cells within a peripheral blood mononuclear cells (PBMC) pool and the overall responsiveness and binding strength.

Therefore, variability in spot-forming units amongst positive responding PBMCs may additionally provide insight into the specificity of responses indicating that responses are specific for each donor. It can also clarify whether the underlying responses are indicative of

of the overall composition of the PBMC pool and/or the specificity of responsiveness for each antigen specific T-cell clone.

Antigen specificity

There are many ways of thinking about the type of antigen used within the assay, but in the end they can all be summarized as a foreign substance that is used to elicit an immune response. Types of antigens may include, but are not limited to: viruses, allergens, bacteria/commensals, mutated proteins, non-mutated irregularly expressed antigens (cancer)/endogenous tumor-associated antigens and genetic engineering-associated antigens. Historically, majority of the work in the ELISpot field has focused on viruses and allergens, for which serological tests can be utilized to best distinguish controls based on positive and negative responding donors. However, with an onset of new therapeutics and genetic engineering we are observing a greater need for analyzing immune cell responses to these new antigens.

CAR-T cell antigens and commensal proteins make up a part of this new wave of therapeutics for which limited responsiveness is likely to be observed within the general population either from a lack of exposure or tolerance. Using an antigen-based specificity approach researchers can utilize alternative, but similar antigens, to prove specific responses against these antigens. T-cell specificity

The T-cell ELISpot allows for the analysis of the presentation of peptides by antigen-presenting cells and the direct activation of antigen-specific T-cells by the interaction between the cognate major histocompatibility complex (MHC) and T-cell receptor (TCR). Activation results in an increase in cytokine and chemokine production by the T-cells, which can be analyzed through the capture of the analyte of interest by antibodies coated to the ELISpot plate. Key analytes for understanding the functional activation of T-cells include the cytokines IFNγ, TNFα, IL-2, IL-4, IL-17 and IL-21, as well as the cytotoxic factors perforin and granzyme.

These analytes can be measured in combination or individually to provide specificity to the assay in delineating specific T-cell subsets. However, often the analytes are measured singly. A major factor in measurement of a single analyte is likely a result of validation constraints in which not all donor T-cells respond the same and may have varying thresholds of secretion for the selected cytokines/chemokines.

The most common analyte is IFNγ; whereas IFNγ is primarily secreted by T-cells and NK cells, it can also be secreted by most of the cells within the PBMC pool, suggesting a potentially non-specific response [9]. Thus, there are multiple additional ways in which a specific T-cell response can be verified. T-cell depletion represents one such way in which T-cell specific responses can be observed. A challenge to this approach is limiting adverse activation of the

T-cells due to activation of the TCR, which can be somewhat avoided through more specific isolations such as CD4 and CD8 isolations or memory/effector specific isolations [10,11]. As a complement or alternative approach antibodies directed against MHC presentation can also prevent and identify T-cell specific responses [12]. Lastly, engineered controls such as engineered TCR-specific T-cells can provide a further specificity control within the assay.

Summary

Assay specificity remains an important factor when selecting a bioanalytical assay, especially for emerging and advanced therapies such as CAR-T cell therapy. ELISpot shows immense promise for the detecting of low frequency cells such as T- and B-cells. However, it is crucial to fully understand three key variables in ELISpot specificity, as well as continue to develop guidance, to ensure that the appropriate assays are applied to get the desired data.

References

1. R Islam, J Vance and M Poirier. Recommendations on ELISpot assay validation by the GCC. *Bioanalysis*, 14, 187–193 (2022).
2. S Janetzki, KS Panageas, L Ben-Porat et al. Results and harmonization guidelines from two large-scale international ELISpot proficiency panels conducted by the Cancer Vaccine Consortium (CVC/SVI). *Cancer Immunol. Immunother*. 57, 303–315 (2008).
3. Clinical and Laboratory Standards Institute. H62. Validation of Assays Performed by Flow Cytometry (1st Edition).
4. S Piccoli, D Mehta, A Vitaliti et al. 2019 White Paper on Recent Issues in Bioanalysis: FDA Immunogenicity Guidance, Gene Therapy, Critical Reagents, Biomarkers and Flow Cytometry Validation (Part 3-Recommendations on 2019 FDA Immunogenicity Guidance, Gene Therapy Bioanalytical Challenges, Strategies for Critical Reagent Management, Biomarker Assay Validation, Flow Cytometry Validation & CLSI H62). *Bioanalysis*, 11, 2207–2244 (2019).
5. B Corsaro, T-Y Yang, R Murphy et al. 2020 White Paper on Recent Issues in Bioanalysis: Vaccine Assay Validation, qPCR Assay Validation, QC for CAR-T Flow Cytometry, NAb Assay Harmonization and ELISpot Validation (Part 3 – Recommendations on Immunogenicity Assay Strategies, NAb Assays, Biosimilars and FDA/EMA Immunogenicity Guidance/Guideline, Gene & Cell Therapy and Vaccine Assays). *Bioanalysis*, 13, 415–463, (2021).
6. ICH: Guidance of Analytical procedures: Text and Methodology. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/q2r1-validation-analytical-procedures-text-and-methodology-guidance-industry
7. Z Moodie, Y Huang, L Gu et al. Statistical positivity criteria for the analysis of ELISpot assay data in HIV-1 vaccine trials. *J. of Imm. Meth*. 315: 121-132, 2006.
8. J Datta, C Rosemblit, E Berk et al. Progressive loss of anti-HER2 CD4+ T-helper type 1 response in breast tumorigenesis and the potential for immune restoration. *Oncoimmunology*, 4, 10 (2015).
9. F Castro, AP Cardoso, RM Gonçalves, K Serre and MJ Oliveira. Interferon-Gamma at the Crossroads of Tumor Immune Surveillance or Evasion. *Front. Immunol*. 0, 847 (2018).
10. I Cassaniti, SA Calarota, KMG Adzasehoun et al. Memory T cells specific for HBV enumerated by a peptide-based cultured enzyme-linked immunospot assay in healthy HBV-vaccinated subjects. Hum. Vaccin. Immunother. 12, 2927–2933 (2016).
11. I Cassaniti, E Percivalle, F Bergami et al. SARS-CoV-2 specific T-cell immunity in COVID-19 convalescent patients and unexposed controls measured by ex vivo ELISpot assay. *Clin. Microbiol. Infect*. 27 (7) (2021).
12. M Wang, ST Tang, A Stryhn et al. Identification of MHC class II restricted T-cell-mediated reactivity against MHC class I binding Mycobacterium tuberculosis peptides. *Immunology*. 132(4):482-91 (2011).

White Paper

For reprint orders, please contact: reprints@future-science.com

Bioanalysis

Recommendations on ELISpot assay validation by the GCC

Gene therapy, cell therapy and vaccine research have led to an increased need to perform cellular immunity testing in a regulated environment to ensure the safety and efficacy of these treatments. The most common method for the measurement of cellular immunity has been Enzyme-Linked Immunospot assays. However, there is a lack of regulatory guidance available discussing the recommendations for developing and validating these types of assays. Hence, the Global CRO Council has issued this white paper to provide a consensus on the different validation parameters required to support Enzyme-Linked Immunospot assays and a harmonized and consistent approach to Enzyme-Linked Immunospot validation among contract research organizations.

First draft submitted: 23 November 2021; Accepted for publication: 18 January 2022; Published online: 9 February 2022

Keywords: cell therapy • cellular immunity • ELISpot • Enzyme-Linked Immunospot • GCC • gene therapy • global CRO council in bioanalysis • regulated bioanalysis • vaccine • validation

Rafiq Islam[1], Jennifer Vance[2], Martin Poirier[3] *et al.*

*Author for correspondence: wei@global-cro-council.org

Disclosure: Due to the equality principles of Global CRO Council for Bioanalysis (GCC), the authors are presented in alphabetical order of member company name, with the exception of the first author, who provided major contributions to this paper.

The full list of author names and affiliations can be found at the end of the article

Background

The Global CRO Council in Bioanalysis (GCC) was created in 2010 as an independent global consortium bringing together many contract research organization (CRO) leaders to discuss various topics and challenges in scientific and regulatory issues related to bioanalysis [1]. Since its formation, the GCC has held regular meetings and published conference reports to share discussions and opinions [2–10]. White papers on specific topics of widespread interest in bioanalysis have also been published to provide unified GCC recommendations helpful to the global bioanalytical community [11–19].

Introduction

Gene therapy, cell therapy and vaccine research have led to an increased need to perform cellular immunity testing in a regulated environment to ensure the safety and efficacy of these treatments. Cellular immunity assays are more complex than traditional immunoassays due to the fact that they include cell culture and not traditional immuno-sandwich. This can result in assays that are less reproducible. Furthermore, cellular immunity assays must be sensitive enough to reliably detect potentially low levels of T-cell populations [20]. It is also known that the reliability of the results can be dependent on the experience of the operator, especially in the handling of primary blood cells [21]. Finally, the lack of appropriate reference standards and positive control samples, particularly those that mimic test samples, can be a challenge.

The most common method for the measurement of cellular immunity has been Enzyme-Linked Immunospot (ELISpot) assays; however, there is a lack of regulatory guidance available discussing the recommendations for developing and validating these types of assays. The available literature can provide examples of cellular immunity testing assays [22–24], but the Clinical and Laboratory Standards Institute (CLSI) published a request for clear guidance for validating these assays as long ago as 2004 [25]. Historically, bioanalysts have attempted to adapt bioanalytical method validation guidance documents [26,27] into a fit-for-purpose approach to method validation, but these documents do not consider ELISpot assays in scope and many parameters are not applicable. In an effort to provide specific recommendations

newlands press

Table 1. GCC survey on ELISpot.

Question	Response
What is the intended use of your ELISpot assays?	Research exploratory: 59.5% Pre-clinical: 70.3% Clinical: 81.2%
Do you perform regulated ELISpot assays?	Yes: 75.7% No: 24.3% It should be noted that those who answered no to this question were forwarded directly to the question regarding critical reagents.
What are the percentages of ELISpot assays you perform in a regulated environment?	GLP or GCP labs: ~50% CLIA/CAP labs: ~25% ISO/GMP labs: 0%
What do you consider a "critical reagent" for ELISpot assays?	PBMC: 89.3% PVDF-backed 96-well microplate: 28.6% Detection antibody: 96.4% Streptavidin-AP: 32.1% BCIP/NBT: 14.3% Positive control: 89.3% Dilution buffers: 0% Wash buffer: 0% Other: 17.9% Other reagents included the diluent (in some instances), assay specific stimuli (e.g., peptide pools), novel therapeutics, peptides, antigens.
Do you use lot-to-lot bridging protocols for ELISpot critical reagents?	Yes: 89.3% No: 10.7%
Do you use positive controls for each run in ELISpot assays?	Yes: 100.0% No: 0.0%
What validation parameters do you use for ELISpot assays?	Precision: 96.4% Sensitivity (LOD): 85.7% Specificity: 78.6% Dilutional linearity: 46.4% Reportable range: 35.7% Ruggedness and robustness: 71.4% Other: 32.1% Other parameters included PBMC/splenocytes F/T stability, matrix stability, whole blood/frozen PBMC stability, plate imaging and sample stability, selectivity, parameters outlined in Corsaro *et al.* [30]
Do you use patient samples for ELISpot validation?	Yes 57.1% No 42.9%
What assay acceptance criteria do you use for ELISpot validation?	The following criteria were used by multiple organizations: • Criteria outlined in Piccoli *et al.* [29] and/or Corsaro *et al.* [30] (nine respondents) • Criteria outlined in Maecker *et al.* [20] (two respondents) • Criteria outlined in Janetzki *et al.* [28] (two respondents) • Replicate CV: 20% (one respondent) or 30% (three respondents) • Intra-assay precision: 30% (two respondents), 25% (two respondents), 20% (two respondents) • Inter-assay precision: 30% (three respondents), 25% (one respondent), 20% (two respondents) • Stability of whole blood or PBMCs: 30% (one respondent) or 20% (one respondent) • Positive control greater than negative control (two respondents)
What assay acceptance criteria do you use for ELISpot sample analysis?	The following criteria were used by multiple organizations: • Criteria outlined in Piccoli *et al.* [29] and/or Corsaro *et al.* [30] (seven respondents) • Replicate CV: 30% (three respondents), 20% (one respondent) • Positive stimulation greater than pre-defined threshold (ten respondents)

CAP: College of American Pathologists; CLIA: Clinical Laboratory Improvement Amendments; ELISpot: Enzyme-Linked Immunospot; GCP: Good clinical practice; GLP: Good laboratory practice; GMP: Good manufacturing practice; ISO: International Organization for Standardization; PBMC: Peripheral blood mononuclear cell; PVDF: Polyvinylidene fluoride.

for improving assay performance, white papers have been published [20,28–30], which, when considered together, can help bioanalysts who are validating ELISpot assays.

A survey was provided to representatives in the GCC in order to determine if any of the existing white paper recommendations are being applied in industry, or if other approaches are being used. This survey received 52 responses, and 35 respondents confirmed that they perform ELISpot assays at their organization. This white paper provides a summary of the results of the survey containing questions and answers on the different approaches to ELISpot validation (refer to Table 1), as well as a consensus on the different validation parameters required to support these assays and a harmonized, consistent approach to ELISpot validation among CROs.

Discussion

ELISpot assays are no longer used simply for research or exploratory purposes; survey results indicate that 76% of ELISpot methods are used for pre-clinical or clinical regulated bioanalysis. In fact, the majority of laboratories (>50%) that run regulated ELISpot assays follow good laboratory practice (GLP) or good clinical practice (GCP) regulations. Less than 25% of Clinical Laboratory Improvement Amendments (CLIA)/College of American Pathologists (CAP) laboratories use regulated ELISpot assays and no International Organization for Standardization (ISO)/good manufacturing practice (GMP) laboratories use these assays.

When queried on which reagents are considered "critical," overwhelming consensus was reached that these include peripheral blood mononuclear cell (PBMC) (89% of respondents), detection antibodies (96% of respondents) and positive controls (89% of respondents). Other reagents could be considered critical depending on the assay (e.g., Streptavidin-AP, polyvinylidene fluoride [PVDF]–backed 96-well microplates) and should be indicated as such in the validation documentation. Consensus was also reached that wash buffers and dilution buffers are not considered critical. Furthermore, respondents overwhelmingly agreed that lot-to-lot bridging must be performed for ELISpot critical reagents.

Respondents were asked to indicate which parameters are being assessed during ELISpot assay validation. The survey results unanimously demonstrate that positive controls must be included in each run. Although the survey does not delineate the type of positive control, it is important to note that these controls can include mitogens such as calcium ionomycin, phorbol 12-myristate 13-acetate (PMA) or phytohemagglutinin (PHA) to determine PBMC functionality; peptide controls such as CEF or CEFT to determine presentation dependent activation; or superantigens such as *Staphylococcus aureus* enterotoxins. In addition, a responding PBMC donor or cell line can be used as a positive control to verify analytical test peptide responses [33]. Furthermore, consensus was reached that precision, sensitivity (LOD), specificity, ruggedness and robustness are required during validation. Additional parameters suggested, but without overwhelming agreement, included dilutional linearity, reportable range and sample stability. It was also interesting to note that only just over half of respondents use patient samples for ELISpot validation. Since patient samples may be limited or unavailable during early-stage development, alternative approaches may need to be taken to extrapolate and assess the utility of the test for clinical samples.

In order to determine the existing harmonization of criteria among those who perform ELISpot validations, respondents were asked to outline what criteria are applied to the evaluations. Three respondents ensure that the positive control is greater than a pre-defined threshold such as the negative control. One respondent required that the response is ≥30 SFU/well, allowing the reporting of both standard deviation (SD) and % CV. For wells with fewer than 30 spots, only SD should be reported. Precision for samples with a mean spot count of greater than 100 will be <25%. For samples with a mean spot count of >30 spots/well up to 100 spots/well, the % CV should be <50%. The remaining proposals for intra- and inter-assay precision varied between 20 and 30% without any mention of dependency on the number of spots per well, and one respondent used a criterion of ≤25% RSD. Finally, two respondents reported stability criteria of either ≤20% bias between each run or 30% CV between time points. Almost half of the 28 respondents who answered this question (46%) use existing recommendations by Maecker *et al.* [20], Janetzki *et al.* [28], Piccoli *et al.* [29] and/or Corsaro *et al.* [30]. Table 2 summarizes these recommendations.

The last question discussed the criteria for sample analysis. Several specified that positive and negative controls should be assayed on each plate and used for acceptance. Most respondents confirm that the same criteria as assay validation should be used.

Recommendations

Following the survey results, the GCC supports prior recommendations for ELISpot assay validation presented in Maecker *et al.* [20], Janetzki *et al.* [28], Piccoli *et al.* [29] and Corsaro *et al.* [30]. Table 2 summarizes these recommendations. Table 3 contains a summary of the additional GCC recommendations following this survey.

Conclusion

In an attempt to harmonize ELISpot validation, the GCC highly recommends the industry adopt the parameters and acceptance criteria provided in Table 3.

Table 2. Summary of prior recommendations presented in Maecker *et al.* [20], Janetzki *et al.* [28], Piccoli *et al.* [29] and Corsaro *et al.* [30]

	Janetzki *et al.* [28] Piccoli *et al.* [29]	Maecker *et al.* [20]	Corsaro *et al.* [30]
SOP	• Counting method for apoptotic cells • Overnight rest of cells prior to plating and incubation • Human auditing during reading process • Adequate adjustments for technical artifacts • Training requirements	Does not discuss	Definition of a counting template can be useful
Serum	Pre-tested and optimized for low background:high signal ratio	Does not discuss	• Pre-screened to ensure lack of reactivity • One to two positive samples
Training	Only trained personnel to conduct assays	Does not discuss	Recommends Janetski *et al.* [28]
Replicates per sample	Six	Three to six	• Replicate variability criteria <twofold • Well acceptance criteria should be established
Sensitivity/LOD	• Antigen-specific spot counts per 2×10^5 PBMCs >10 • At least 3x as high as the background reactivity [24,31]	Two SDs above the mean of replicate negative control samples	• One sample to assess ULOD; ULOD ≤450 SFC/well • Media-only wells used to assess LOD • Statistical approach, such as with dual criteria or mean + two SDs of pre-existing immunity/background • Positivity criteria for a sample established based on meaningful level of reactivity that is above the background reactivity
Precision	Does not discuss	• Intra-assay: six replicates per assay • Inter-assay: eight assays on different days • CV 4–133% for medium and high responders • Use SD for low responders • Depends on counting statistics – 2×10^5 PBMCs/well gave highest counting efficiency	• Required during validation • Use ≤10 samples • Intermediate % CV ≤40% for ≥80% of samples having SFU/10^6 PBMCs greater than LOQ • Six samples tested at three cell inputs
Ruggedness	Does not discuss	Three different operators on the same day	• Required during validation • Maximum fold difference between assays <twofold • Different parameters that can vary over time during routine operation should be tested
Linearity	Does not discuss	• Serial dilute PBMC from a high responder (triplicate samples) into PBMC from a non-responder • R^2 values >0.97	• Required during validation • Six samples at three cell inputs • Sample reactivity expected to decrease as cell input decreases
LOQ	Does not discuss	Does not discuss	• Required during validation • Lowest value that can be quantified with acceptable precision (intermediate precision % CV ≤40%) • LLOQ ≥ LOD
Specificity	Does not discuss	Does not discuss	• Required during validation • Determined with a negative cutoff determination for vector and transgene peptide pools using a correction factor specific to each peptide pool
Normalization of results	Does not discuss	Does not discuss	Cells can be normalized, that is, PBMCs analyzed by flow cytometry and then PBMCs adjusted to the fixed/pre-defined number of T-cells (executed at the site of the ELISpot analysis)
Critical reagents	Does not discuss	Does not discuss	• Do not to use cells with a viability <80% • High-quality frozen PBMC preparations • Capture and detection antibody pair • Antigen source can be overlapping synthetic peptides or whole proteins • Filter plates may be PVDF or other variety • Side-by-side assessment of new reagent lots against qualified lots is necessary for reagent bridging and trending

ELISpot: Enzyme-Linked Immunospot; PBMC: Peripheral blood mononuclear cell; PVDF: Polyvinylidene fluoride; SD: Standard deviation; SFC: Spot-forming cell; SFU: Spot-forming unit; SOP: Standard operating procedure; ULOD: Upper limit of detection.
Data taken from [20,28–30].

Table 3. Additional GCC recommendations on ELISpot assay validation.

Parameter	Recommendation
SOP	Should include PBMC isolation, counting method, cell handling, plate reading method, training requirements and equipment specific procedures It is suggested that the type of collection tube (heparin), sample collection and handling times, shipping times determined during method development and validation also be included
Normalization of results	• Intra- and inter-subject, as needed • Multiple baselines possible • Normalization against negative control
Critical reagents	• PBMC (refer to Corsaro *et al.* [30] for recommendation for using high-quality PBMC; alternative criteria may be utilized when evaluated during method development and confirmed to support the context of use); also, it is possible to consider the option of CPT tubes for collection as an alternative to removal of granulocytes at point of analysis • Detection antibodies • Positive controls • Lot-to-lot bridging should be performed • In the absence of vendor-provided stability data, stability experiments must be performed to demonstrate critical reagent stability as per usage in the assay
Sample type	• Patient samples reflecting study population should be used during method development and validation, if available • All runs should include positive controls and negative controls • Run all controls and samples in triplicate (three wells per result) during validation and sample analysis
Quality controls	• At least two levels of positive control and one negative control (media only) • Acceptable range should be established during validation • Reference sample/trending control for each day ELISpot is run
Validation parameters based on COU	• Precision • Sensitivity (LLOQ and LOD) • ULOQ/reportable range • Specificity • Ruggedness and robustness • Linearity • Critical reagent stability • Whole blood or PBMC stability
Precision	• Minimum ten donors • Inter-assay: should include a minimum of six runs with three replicates each by two analysts over multiple days • Intra-assay: minimum one run and six replicates • CV $\leq$30% • Total error $<$40% (LLOQ $<$50%)
Sensitivity (LLOQ and LOD)	• Determined based on precision data; the intermediate precision as LLOQ should be based on the acceptable intermediate precision of 40% • LOD is determined based on two SDs above the mean of replicate negative control samples • Due to the mathematical considerations of a high % CV at low spot numbers per well, statistical testing [32] is recommended for samples that are below 30 spots per well and above the LOD • LLOQ $\geq$ LLOD
ULOQ	Defined as the maximum number of individual spots per well the ELISpot plate reader software can discriminate; this can be achieved by counting spots using a series of cell dilutions treated with mitogen, or peptide for a donor with a very strong peptide response
Specificity	Positive control greater than negative control; should also be tested with non-specific peptides such as beta-actin; in the case of non-specific peptides, the response must be less than LLOQ; in addition, specificity should examine the full extent to which an assay responds to all subsets of an analyte [30]; often this may also include an assessment of specificity for the target cell type, especially for assays aiming to measure this component
Ruggedness and robustness	• Maximum fold difference between assay ruggedness factor levels is expected to be less than twofold for tenfold dilution of cells; in the case of analysts and instruments – should meet the % CV criteria • Inter-laboratory comparison studies may be performed to demonstrate assay ruggedness
Linearity	• Serial dilute PBMC from a high responder into PBMC from a non-responder • Use at least six donors (high responders) and at least three dilutions • Sample reactivity expected to decrease as cell input decreases • R^2 values $>$0.97
Selectivity	• Ten different lots/donors of PBMC (refer to Corsaro *et al.* [30] for recommendation for using high-quality PBMC) • Decide on and establish level of response needed for LLOQ from the PBMCs in MD, to determine the level of reactivity needed for selectivity determination of the assay in validation • Controls for selectivity are based on media and PBMC positivity criteria for a selectivity sample established based on meaningful level of reactivity of that which is below the background reactivity [30]; the difference between selectivity samples is the media controls with PBMC responses $\leq$ LLOQ; +/- 40% • $\geq$70% of the lots should pass this criterium

COU: Context of use; ELISpot: Enzyme-Linked Immunospot; LLOD: Lower limit of detection; MD: Method development; PBMC: Peripheral blood mononuclear cell; SD: Standard deviation; SOP: Standard operating procedure.

Future perspective

The GCC as a global organization will continue to provide recommendations on hot topics of global interest in bioanalysis. Please contact the GCC [34] for the exact date and time of future meetings, and for all membership information.

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. Premkumar N, Lowes S, Jersey J *et al.* Formation of a global contract research. *Bioanalysis* 2(11), 1797–1800 (2010).
2. Breda N, Garofolo F, Cruz Caturla M *et al.* The 3rd Global CRO Council for Bioanalysis at the International Reid Bioanalytical Forum. *Bioanalysis* 3(24), 2721–2727 (2011).
3. Lowes S, Jersey J, Shoup R *et al.* 4th Global CRO Council for Bioanalysis: coadministered drugs stability, EMA/US FDA guidelines, 483s and carryover. *Bioanalysis* 4(7), 763–768 (2012).
4. Nicholson R, Lowes S, Cruz Caturla M *et al.* 6th GCC focus on LBA: critical reagents, positive controls and reference standards; specificity for endogenous compounds; biomarkers; biosimilars. *Bioanalysis* 4(19), 2335–2342 (2012).
5. Rocci M, Lowes S, Shoup R *et al.* 7th GCC: use of incurred samples; fit-for-purpose validation; solution stability; electronic lab notebooks; hyperlipidemic matrix testing. *Bioanalysis* 6(20), 2713–2720 (2014).
6. Bower J, Fast D, Garofolo F *et al.* 8th GCC: consolidated feedback to US FDA on the 2013 draft FDA guidance on bioanalytical method validation. *Bioanalysis* 6(22), 2957–2963 (2014).
7. Hayes R, LeLacheur R, Dumont I *et al.* 9th GCC closed forum: CAPA in regulated bioanalysis; method robustness, biosimilars, preclinical method validation, endogenous biomarkers, whole blood stability, regulatory audit experiences and electronic laboratory notebooks. *Bioanalysis* 8(6), 487–495 (2016).
8. Cape S, Islam R, Nehls C *et al.* The 10th GCC closed forum: rejected data, GCP in bioanalysis, extract stability, BAV, processed batch acceptance, matrix stability, critical reagents, ELN and data integrity and counteracting fraud. *Bioanalysis* 9(7), 505–516 (2017).
9. Islam R, Briscoe C, Bower J *et al.* 11th GCC closed forum: cumulative stability; matrix stability; immunogenicity assays; laboratory manuals; biosimilars; chiral methods; hybrid LBA/LCMS assays; fit-for-purpose validation; China Food and Drug Administration bioanalytical method validation. *Bioanalysis* 10(7), 433–444 (2018).
10. Briscoe C, Hughes N, Hayes R *et al.* 12th GCC closed forum: critical reagents; oligonucleotides; CoA; method transfer; HRMS; flow cytometry; regulatory findings; stability; and immunogenicity. *Bioanalysis* 11(12), 1129–1138 (2019).
11. Lowes S, Jersey J, Shoup R *et al.* Recommendations on: internal standard criteria, stability, incurred sample reanalysis and recent 483s by the Global CRO Council for Bioanalysis. *Bioanalysis* 3(12), 1323–1332 (2011).
12. Sangster T, Maltas J, Struwe P *et al.* Recommendations on ISR in multi-analyte assays, QA/bioanalytical consultants and GCP by Global CRO Council for Bioanalysis (GCC). *Bioanalysis* 4(14), 1723–1730 (2012).
13. Boterman M, Doig M, Breda M *et al.* Recommendations on the interpretation of the new EMA guideline on bioanalytical method validation by Global CRO Council for Bioanalysis (GCC). *Bioanalysis* 4(6), 651–660 (2012).
14. Lowes S, Boterman M, Doig M *et al.* Recommendations on bioanalytical method stability implications of co-administered and co-formulated drugs by Global CRO Council for Bioanalysis (GCC). *Bioanalysis* 4(17), 2117–2126 (2012).
15. Hougton R, Gouty D, Allinson J *et al.* Recommendations on biomarker bioanalytical method validation by GCC. *Bioanalysis* 4(20), 2439–2446 (2012).
16. Lowes S, LeLacheur R, Shoup R *et al.* Recommendations on incurred sample stability (ISS) by GCC. *Bioanalysis* 6(18), 2385–2390 (2014).
17. Islam R, Kar S, Ritzén H *et al.* Recommendations for classification of commercial LBA kits for biomarkers in drug development from the GCC for bioanalysis. *Bioanalysis* 11(7), 645–653 (2019).
18. Nels C, Buonarati M, Cape S *et al.* GCC consolidated feedback to ICH on the 2019 ICH M10 bioanalytical method validation draft guideline. *Bioanalysis* 11(18s), 1–228 (2019).
19. Bower J, Zimmer J, McCowan S *et al.* Recommendations for the content and management of certificates of analysis for reference standards from the GCC for bioanalysis. *Bioanalysis* 13(8), 609–619 (2021).
20. Maecker HT, Hassler J, Payne JK *et al.* Precision and linearity targets for validation of an IFNγ ELISPOT, cytokine flow cytometry, and tetramer assay using CMV peptides. *BMC Immunology* 9(9), (2008). https://doi.org/10.1186/1471-2172-9-9

21. Janetzki S, Schaed S, Blachere NE *et al.* Evaluation of ELISpot assays: influence of method and operator on variability of results. *J. Immunol. Methods* 291, 175–183 (2004).
22. Britten CM, Gouttefangeas C, Schoenmaekers-Welters MJP *et al.* The CIMT-monitoring panel: a two-step approach to harmonize the enumeration of antigen-specific $CD8^+$ T lymphocytes by structural and functional assays. *Cancer Immunol. Immunother.* doi:10.1007/s00262-007-0379-z (2007).
23. Cox JH, Ferrari G, Janetzki S. Measurement of cytokine release at the single cell level using the ELISPOT assay. *Methods* 38, 274–282 (2006).
24. Dubey S, Clair J, Fu T-M *et al.* Detection of HIV vaccine-induced cell-mediated immunity in HIV-seronegative clinical trial participants using an optimized and validated enzyme-lined immunospot assay. *J. Acquir. Immune Defic. Syndr.* 45, 20–27 (2007).
25. NCCLS. Performance of single cell immune response assays; approved guidelines. NCCLS document I/LA26-A (2004).
26. US FDA, Center for Drug Evaluation and Research. Draft guidance for industry: Clinical immunogenicity considerations for biosimilar and interchangeable insulin products (2019). https://www.fda.gov/media/133014/download
27. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP). Guideline on Bioanalytical Method Validation. London, UK (2011).
28. Janetzki S, Panageas KS, Ben-Porat L *et al.* Results and harmonization guidelines from two large-scale international ELISpot proficiency panels conducted by the Cancer Vaccine Consortium (CVC/SVI). *Cancer Immunol. Immunother.* 57, 303–315 (2008).
29. Piccoli S, Mehta D, Vitaliti A *et al.* 2019 white paper on recent issues in bioanalysis: FDA immunogenicity guidance, gene therapy, critical reagents, biomarkers and flow cytometry validation (part 3 – recommendations on 2019 FDA immunogenicity guidance, gene therapy bioanalytical challenges, strategies for critical reagent management, biomarker assay validation, flow cytometry validation & CLSI H62). *Bioanalysis* 11(24), 2207–2244 (2019).
30. Corsaro B, Yang TY, Murphy R *et al.* 2020 white paper on recent issues in bioanalysis: vaccine assay validation, qPCR assay validation, QC for CAR-T flow cytometry, NAb assay harmonization and ELISpot validation (part 3 – recommendations on immunogenicity assay strategies, NAb assays, biosimilars and FDA/EMA immunogenicity guidance/guideline, gene & cell therapy and vaccine assays). *Bioanalysis* 13(6), 415–463 (2021).
31. Cox JH, Ferrari G, Kalams SA *et al.* ELISpot Collaborative Study Group. Results of an ELISpot proficiency panel conducted in 11 laboratories participating in international human immunodeficiency virus type 1 vaccine trials. *AIDS Res. Hum. Retroviruses* 21, 68–81 (2005).
32. Moodie Z, Price L, Gouttefangeas C *et al.* Response definition criteria for ELISPOT assays revisited. *Cancer Immunol. Immunother.* 59, 1489–1501 (2010).
33. Hagen J, Kalyuzhny AE. Essential controls for ELISpot assay. *Methods Mol. Biol.* 1808, 31–41 (2018).
34. Global CRO Council for Bioanalysis (2021). http://www.global-cro-council.org

Authors

Rafiq Islam[1], Jennifer Vance[2], Martin Poirier[3], Jennifer Zimmer[4], Ardeshir Khadang[5], Dave Williams[6], Jennifer Zemo[6], Todd Lester[6], Marianne Fjording[6], Amanda Hays[6], Nicola Hughes[7], Fabio Garofolo[8], Curtis Sheldon[9], Rudolf Guilbaud[10], Christina Satterwhite[11], Kelly Colletti[11], Elizabeth Groeber[12], Heidi Renfrew[13], Mathilde Yu[14], Jenny Lin[15], Xinping Fang[16], Mark Wissel[17], Thomas Beadnell[17], John Lin[18], Santosh Shah[18], Wei Garofolo*,[19], Natasha Savoie[19], Roger Hayes[20], John Pirro[21], Cheikh Kane[21], Marsha Luna[21], Allan Xu[22], Stephanie Cape[23], Mark O'Dell[24], Robert Wheller[25], Hanna Ritzen[26], Esme Farley[27], Lisa Kierstead[28], William Mylott[29], Edward Tabler[29], Moucun Yuan[29], Shane Karnik[30], Troy Voelker[31], Ira DuBey[1], Clark Williard[32], Kelly Dong[33], Jing Shi[34] & Jim Yamashita[34]

Affiliations

[1]Smithers, Gaithersburg, MD, USA
[2]AIT Bioscience – a Nexelis Company, Seattle, WA, USA
[3]Altasciences, Everett, WA, USA
[4]Alturas Analytics, Moscow, ID, USA
[5]Axis Clinical, Dilworth, MN, USA
[6]BioAgilytix, Durham, NC, USA
[7]Biopharmaservices, Toronto, ON, Canada
[8]BRI – a Frontage Company, Vancouver, BC, Canada
[9]Celerion, Lincoln, NE, USA
[10]CellCarta, Montreal, Quebec, Canada
[11]Charles River Laboratories, Reno, NV, USA
[12]Charles River Laboratories, Ashland, OH, USA
[13]Charles River Laboratories, Shrewsbury, MA, USA
[14]CIRION, Montreal, Quebec, Canada
[15]CMIC, Hoffman Estates, IL, USA
[16]CR Medicon, Piscataway, NJ, USA
[17]Eurofins Virocor, Lee's Summit, MO, USA
[18]Frontage Laboratories, Exton, PA, USA
[19]GCC, Toronto, ON, Canada
[20]ICON Laboratory Services, Whitesboro, NY, USA
[21]KCAS Bioanalytical & Biomarker Services, Shawnee Mission, KS, USA
[22]Keystone Bioanalytical, Lansdale, PA, USA
[23]Labcorp, Madison, WI, USA
[24]Labcorp, Indianapolis, IN, USA
[25]LGC, Ely, UK
[26]Mercodia, Uppsala, Sweden
[27]Pharmaron, Germantown, MD, USA
[28]PPD Laboratories, Wayne, PA, USA
[29]PPD Laboratories, Richmond, VA, USA
[30]Pyxant Labs, Colorado Springs, CO, USA
[31]Pyxant Labs, Salt Lake City, UT, USA
[32]Syneos Health, Princeton, NJ, USA
[33]United-Power Pharma, Beijing, China
[34]WuXi AppTec, Plainsboro, NJ, USA

IMMUNOGENICITY ASSAY DEVELOPMENT, VALIDATION & TRANSFER IN A VACCINE CANDIDATE CLINICAL TRIAL

INTRODUCTION / OVERVIEW

Vaccine development is a long, complex process, requiring evaluation for safety, immunogenicity, and protective efficacy. With this complexity, many factors can impact the probability of licensure and ultimate public health impact.

A large global pharmaceutical company, with a promising new vaccine candidate, partnered with Eurofins Viracor to perform immunogenicity testing as part of their clinical trial development. The sponsor's goal in developing the vaccine is to create widespread immunity to protect against infection, morbidity, and potential mortality. This objective can be compared to the historical efforts used to control polio, smallpox, HPV, or measles.

SITUATION / CHALLENGE

The sponsor initially recruited Eurofins Viracor for rapid transfer of an ELISpot assay, before recognizing the value and efficiency of being able to rely on the broad technical expertise of Viracor as a single provider for all assay optimization, validation and transfer requirements of the trial. Consequently, the success of the first project was soon followed by an expanded request for combined development and transfer for a Neutralizing Antibody (NAb) assay. Ultimately, Viracor was asked to develop, validate and perform multiple biomarker assays including ELISpot and NAb, with a qPCR assay added later as an endpoint.

SOLUTION

ELISpot Assay:

The ELISpot assay enables assessment of study subjects' adaptive immunity induced by the sponsor's vaccine candidate through quantification of the T cell responses. Combining the results of the ELISpot assay with the NAb assay enables a broad spectrum interpretation of overall immunity induced by the vaccine.

The developed ELISpot assay was performed in tandem at the sponsor's site and at Viracor, with the same PBMC sample panel. The results of this study demonstrated an excellent level of analytical concordance between the two sites, and therefore confidence that the assay was performing well. In addition, assay optimization by the Viracor R&D team improved throughput, maximized signal robustness, improved precision and enhanced analytical specificity of the assay.

Neutralizing Antibody Assay (NAb)

This is a cell-based assay for determining the presence and relative titer of human virus-specific neutralizing antibodies in subjects' serum following investigational treatment with the vaccine. Increases in virus-specific neutralization titer response demonstrates the study subject is producing a positive immune response to the vaccination.

The resulting fully developed and optimized assay was validated following FDA immunogenicity guidance criteria and transitioned to Viracor's clinical lab environment for high throughput testing of trial samples.

qPCR Assay

The purpose of this assay in the study was to assess for a viral infection and therefore determine vaccine efficacy.

Viracor performed four validations as part of the project (virus in matrix 1, virus in matrix 2, human cellular gene in matrix 1, and human cellular gene in matrix 2). In addition to this Viracor developed assay, another assay was also developed to assess qPCR assay specificity.

CASE STUDY

Immunogenicity Assay Development, Validation & Transfer in a Vaccine Candidate Clinical Trial

OUTCOME

- Acceptance criteria for the assay(s) were created in pre-validation, submitted to and approved by CBER without change or comment.
- Several specific areas of our expertise were called upon to enable the Viracor team to deliver on this project for the sponsor:

1. Our PCR expertise was needed to optimize and validate the quantitative viral detection and PCR specificity assays, as well as troubleshooting specific obstacles during nucleic acid extraction.
2. Our equipment allowed high through-put of samples, which was of high importance for the timely processing of the high number of samples expected in this study.
3. Our expertise in cell culture and ELISpot assay techniques proved invaluable to optimization and validation of a sensitive, robust, precise and specific ELISpot assay method.
4. Leveraging our excellent project management with Viracor's scientific expertise and openness to collaboration allowed timelines to exceed sponsor expectations.

- Our commitment to quality, technical precision, maintenance of critical cell lines and serum samples allowed Viracor to optimize, validate, and deliver multiple precise, robust and sensitive assays for the sponsor's trial.

Viracor is an integrated arm of Eurofins, offering complex/esoteric testing for clinical research. For more than 30 years, Viracor has been dedicated to helping clients by providing high quality, accurate results across multiple phases of drug development. Offering our partners broad experience in molecular infectious disease testing, immune response monitoring, vaccine safety/efficacy assessment, allergy and hypersensitivity testing.

Viracor is passionate about delivering value to our clients by providing timely, actionable information — never losing sight of the connection between the testing we perform and the goals of your study.

Contact us today to discover how the Eurofins Viracor team can make the difference in your projects.

eurofins | Viracor BioPharma

Trusted Laboratory Partner

CAP Accredited CLIA Licensure

Over 2,800 validated tests

Eurofins Viracor Biopharma
1001 NW Technology Dr.
Lee's Summit, MO 64086
USA
Tel +1 800 305 5198
Fax +1 816 347 0143

clinicaltrials@viracor-eurofins.com
www.eurofins-viracor.com/biopharma

For reprint orders, please contact: reprints@future-science.com

Bioanalysis

2020 White Paper on Recent Issues in Bioanalysis: Vaccine Assay Validation, qPCR Assay Validation, QC for CAR-T Flow Cytometry, NAb Assay Harmonization and ELISpot Validation (Part 3 – Recommendations on Immunogenicity Assay Strategies, NAb Assays, Biosimilars and FDA/EMA Immunogenicity Guidance/Guideline, Gene & Cell Therapy and Vaccine Assays)

Bart Corsaro[‡,1], Tong-yuan Yang[‡,2], Rocio Murphy[‡,3], Ivo Sonderegger[‡,10], Andrew Exley[‡,§,7], Sylvie Bertholet[‡,1], Naveen Dakappagari[‡,5], Francis Dessy[‡,6], Fabio Garofolo[‡,§,*,8], Lisa Kierstead[‡,§,9], Holger Koch[‡,10], Ghanashyam Sarikonda[‡,5], Natasha Savoie[‡,§,11], Richard Siggers[‡,§,12], Therese Solstad[‡,§,13], Yanmei Lu[§,19], Mark Milton[§,22], Jean-Claude Marshall[§,21], Jason DelCarpini[§,14], Boris Gorovits[§,¶,17], Swati Gupta[§,15], Lynne Jesaitis[§,16], John Kamerud[§,17], Arno Kromminga[§,18], Anna Ma[§,20], Jim McNally[§,23], Haoheng Yan[#,¶,4], Bonnie Wu[#,¶,2], Daniela Verthelyi[#,¶,4], Susan Kirshner[#,¶,4], Joao Pedras-Vasconcelos[#,¶,4], Manoj Rajadhyaksha[#,¶,24], Roland F. Staack[#,¶,25], Elana Cherry[#,¶,12], Isabelle Cludts[#,¶,26], Madeleine Dahlbäck[#,27], George R. Gunn[#,28], Akiko Ishii-Watabe[#,29], Vibha Jawa[#,30], Robert Kubiak[#,31], Michael Partridge[#,24], Marco Petrillo[#,32], Samuel O. Pine[#,33], Johann Poetzl[#,34], Sam Song[#,35], Chris Stebbins[#,36], Yuling Wu[#,31], Lucia Zhang[#,¶,12], Sumit Kar[¶,11], Meina Liang[¶,37], Mohsen Rajabi Abhari[¶,4], Becky Schweighardt[¶,16], Kay Stubenrauch[¶,25] & Yuanxin Xu[¶,38]

[1]GlaxoSmithKline, Rockville, MD, USA
[2]Janssen R&D, Spring House, PA, USA
[3]Merck, West Point, PA, USA
[4]US FDA, Silver Spring, MD, USA
[5]Navigate, a Novartis Subsidiary, San Diego, CA, USA
[6]GlaxoSmithKline, Rixensart, Belgium
[7]UK MHRA, London, UK
[8]BRI, Vancouver, BC, Canada
[9]PPD, Wayne, PA, USA
[10]Takeda, Opfikon, Switzerland
[11]CFABS, Montreal, QC, Canada
[12]Health Canada, Ottawa, ON, Canada
[13]Norwegian Medicines Agency, Oslo, Norway
[14]bluebird bio, Cambridge, MA, USA
[15]Abbvie, Irvine, CA, USA
[16]BioMarin, Novato, CA, USA
[17]Pfizer, Andover, MA, USA
[18]BioAgilytix, Hamburg, Germany
[19]Sangamo Therapeutics, Brisbane, CA, USA
[20]CRISPR Therapeutics, Cambridge, MA, USA
[21]Pfizer, Groton, CT, USA
[22]Novartis, Cambridge, MA, USA
[23]BioAgilytix, Durham, NC, USA
[24]Regeneron Pharmaceuticals, Tarrytown, NY, USA
[25]Roche Pharma Research & Early Development, Roche Innovation Center, Munich, Germany
[26]UK MHRA-NIBSC, London, UK
[27]Novo Nordisk, Maaloev, Denmark
[28]GlaxoSmithKline, Collegeville, PA, USA
[29]Japan MHLW-NIHS, Tokyo, Japan
[30]Merck, Kenilworth, NJ, USA
[31]AstraZeneca, Gaithersburg, MD, USA
[32]Biogen, Cambridge, MA, USA
[33]Ablynx, Zwijnaarde, Belgium
[34]Hexal AG, Oberhaching, Germany
[35]Takeda, Cambridge, MA, USA
[36]Immunologix Labs, Tampa, FL, USA
[37]AstraZeneca, South San Francisco, CA, USA
[38]Intellia Therapeutics, Cambridge, MA, USA

[‡]SECTION 1 – Vaccine Clinical Assays and Cell Therapy (Authors in Section 1 are presented in alphabetical order of their last name, with the exception of the first five authors who were major contributors).
[§]SECTION 2 – Gene Therapy, qPCR, NGS and ELISpot Validation (Authors in Section 2 are presented in alphabetical order of their last name, with the exception of the first five authors who were major contributors).
[#]SECTION 3 – NAb Assay Harmonization, Biosimilars and FDA/EMA Guidance/Guideline (Authors in Section 3 are presented in alphabetical order of their last name, with the exception of the first seven authors who were major contributors).
[¶]SECTION 4 – Immunogenicity Assay Strategies (Authors in Section 4 are presented in alphabetical order of their last name, with the exception of the first six authors who were major contributors).

Disclaimer: The views expressed in this article are those of the authors and do not reflect official policy of the US FDA, Europe EMA, France ANSM, UK MHRA, Austria AGES, Brazil ANVISA, Health Canada, Japan MHLW, and Finland Fimea. No official endorsement by the FDA, EMA, ANSM, Health Canada, MHRA, AGES, ANVISA, MHLW or Fimea is intended or should be inferred.

*Author for correspondence: fabiogarofolo@hotmail.com

newlands press

The 14th edition of the Workshop on Recent Issues in Bioanalysis (14th WRIB) was held virtually on June 15-29, 2020 with an attendance of over 1000 representatives from pharmaceutical/biopharmaceutical companies, biotechnology companies, contract research organizations, and regulatory agencies worldwide. The 14th WRIB included three Main Workshops, seven Specialized Workshops that together spanned 11 days in order to allow exhaustive and thorough coverage of all major issues in bioanalysis, biomarkers, immunogenicity, gene therapy and vaccine. Moreover, a comprehensive vaccine assays track; an enhanced cytometry track and updated Industry/Regulators consensus on BMV of biotherapeutics by LCMS were special features in 2020. As in previous years, this year's WRIB continued to gather a wide diversity of international industry opinion leaders and regulatory authority experts working on both small and large molecules to facilitate sharing and discussions focused on improving quality, increasing regulatory compliance and achieving scientific excellence on bioanalytical issues.

This 2020 White Paper encompasses recommendations emerging from the extensive discussions held during the workshop and is aimed to provide the Global Bioanalytical Community with key information and practical solutions on topics and issues addressed, in an effort to enable advances in scientific excellence, improved quality and better regulatory compliance. Due to its length, the 2020 edition of this comprehensive White Paper has been divided into three parts for editorial reasons.

This publication (Part 3) covers the recommendations on Vaccine, Gene/Cell Therapy, NAb Harmonization and Immunogenicity). Part 1 (Innovation in Small Molecules, Hybrid LBA/LCMS & Regulated Bioanalysis), Part 2A (BAV, PK LBA, Flow Cytometry Validation and Cytometry Innovation) and Part 2B (Regulatory Input) are published in volume 13 of Bioanalysis, issues 4 and 5 (2020), respectively.

First draft submitted: 7 January 2021; Accepted for publication: 8 January 2021; Published online: 3 February 2021

Keywords: bioanalysis • biomarkers • cell therapy • gene therapy • immunogenicity • vaccine • WRIB

Acronyms	
AAV:	Adenovirus-associated virus
Ab:	Antibody
ADA:	Anti-drug antibody
ADC:	Antibody-drug conjugates
ADCC:	Antibody-dependent cell-mediated cytotoxicity
ADHS:	Antibody-depleted human serum
BAV:	Biomarker assay validation
BLA:	Biologics license application
BLAST:	Basic Local Alignment Search Tool
BMV:	Bioanalytical method validation
CAR-T:	Chimeric antigen receptor T cells are T cells that have been genetically engineered to express one or more receptors targeting specific proteins for use in immunotherapy
CBA:	Cell-based assays
cDNA:	Complementary DNA
CDR:	Complementarity-determining regions
CDx:	Companion diagnostics
cGMP:	Current Good Manufacturing Practices
CIC:	Circulating immune complexes

CLBA:	Competitive ligand binding assays
CLIA:	Clinical laboratory improvements amendments
Clinically Relevant ADA:	ADA impacting PK, PD, efficacy and/or safety of the biotherapeutic in patients
CLSI:	Clinical and Laboratory Standards Institute
CMC:	Chemistry, Manufacturing, and Controls
COU:	Context of use
Ct:	Threshold cycle
CTL:	Cytotoxic T lymphocytes
CV:	Coefficient of variation
dPCR:	Digital polymerase chain reaction
ddPCR:	Droplet digital polymerase chain reaction
DNA:	Deoxyribonucleic acid
DoE:	Design of experiments
Dx:	Diagnostic
FFP:	Fit-for-purpose
FIH:	First-in-human
FPR:	False-positive rate
GCT:	Gene and cell therapy
gDNA:	Genomic DNA
GTx:	Gene therapy
HDR:	Homology directed repair
IC:	Immune complex
IND:	Investigational new drug
Indel:	Insertion–deletion mutations
IQR:	Inter-quartile range
ISI:	Integrated Summary of Immunogenicity
ISR:	Incurred sample reproducibility
ITI:	Immune tolerance induction
KOL:	Key opinion leader
LBA:	Ligand binding assay
LCMS:	Liquid chromatography mass spectrometry
LDT:	Laboratory developed test
LLOQ:	Lower limit of quantitation
LOB:	Limit of blank
LOD:	Limit of detection
mAb:	Monoclonal antibody
MDB:	Multi-domain biotherapeutics
MFI:	Mean fluorescent intensity
MIQE:	Minimum Information for Publication of Quantitative Real-Time PCR Experiments
MoA:	Mechanism of action
MOI:	Multiplicity of infection
MRD:	Minimum required dilution
mRNA:	Messenger RNA
Multiplex:	A type of assay that simultaneously measures multiple analytes in a single experiment.
NAb:	Neutralizing antibody
NGS:	Next generation sequencing

NHEJ:	Non-homologous end joining
OD:	Optical density
PBMC:	Peripheral blood mononuclear cell
PCR:	Polymerase chain reaction
PD:	Pharmacodynamic
PK:	Pharmacokinetic
PMR:	Post-marketing request
PMT:	Photomultiplier tube
PVDF :	Polyvinylidene difluoride
QA:	Quality assurance
QC:	Quality control
qPCR:	Quantitative (real-time) polymerase chain reaction
RNA:	Ribonucleic acid
RSD:	Relative standard deviation
RT:	Reverse transcription
SD:	Standard deviation
SEC:	Size exclusion chromatography
SOP:	Standard operating procedure
ss/ds:	Single stranded/double stranded
SSM:	Spillover spreading matrix
TI:	Transduction inhibition
ULOQ:	Upper limit of quantitation
WRIB:	Workshop on Recent Issues in Bioanalysis

Index

Introduction

The 14th edition of the Workshop on Recent Issues in Bioanalysis (14th WRIB) was held virtually between 15–29 June 2020 with an attendance of over 1000 representatives from pharmaceutical/biopharmaceutical companies, biotechnology companies, contract research organizations, and regulatory agencies worldwide. The 14th WRIB included three main workshops, seven Specialized Workshops that together spanned 11 days to allow exhaustive and thorough coverage of all major issues in bioanalysis, biomarkers, immunogenicity, gene therapy, cell therapy and vaccine.

Moreover, a comprehensive vaccine assays track; an enhanced cytometry track, and updated Industry/Regulators consensus on bioanalytical method validation (BMV) of biotherapeutics by mass spectrometry (hybrid assays, LCMS and HRMS) were special features in 2020.

As in previous years, this year's WRIB continued to gather a wide diversity of international industry opinion leaders and regulatory authority experts working on both small and large molecules to facilitate sharing, reviewing, discussing and agreeing upon best approaches aimed to achieve scientific excellence and increase regulatory compliance on bioanalytical issues.

The active contributing chairs included Dr. Stephen C Alley (Seattle Genetics), Dr. Anna Edmison (Health Canada), Dr. Chris Evans (GSK), Dr. Christine Fandozzi (Merck), Dr. Sally Fischer (Genentech), Dr. Fabio Garofolo (BRI), Dr. Christine Grimaldi (Boehringer Ingelheim), Dr. Lindsay King (Pfizer), Dr. Rocio Murphy (Merck), Dr. Hendrik Neubert (Pfizer), Dr. Manoj Rajadhyaksha (Regeneron), Dr. Catherine Soo (Health Canada), Dr. Susan Spitz (Incyte), Dr. Roland Staack (Roche), Dr. Scott Summerfield (GSK), Dr. Alessandra Vitaliti (Novartis), Dr. Jan Welink (EU EMA), Dr. Haoheng Yan (US FDA), Dr. Tong-yuan Yang (Janssen), Dr. Hongbin Yu (Boehringer Ingelheim), Dr. Yan Zhang (BMS).

The participation of regulatory agency representatives continued to grow at WRIB [1–25] including the below:

- **Regulated Bioanalysis and BMV Guidance/Guidelines:** Dr. Arindam Dasgupta (US FDA), Dr. Sam Haidar (US FDA), Dr. Mohsen Rajabiabhari (US FDA), Dr. Tahseen Mirza (US FDA), Dr. Nilufer Tampal (US FDA), Dr. Suman Dandamudi (US FDA), Dr. Diaa Shakleya (US FDA), Dr. Jinhui Zhang (US FDA), Dr. Patrick Faustino (US FDA), Dr. Jan Welink (EU EMA), Mr. Stephen Vinter (UK MHRA), Mr. Michael McGuinness (UK MHRA), Dr. Anna Edmison (Health Canada), Dr. Catherine Soo (Health Canada), Dr. Susan Stojdl (Health Canada), Mr. Gustavo Mendes Lima Santos (Brazil ANVISA)
- **Immunogenicity, Gene Therapy, Cell Therapy and Vaccines:** Dr. Susan Kirshner (US FDA), Dr. Daniela Verthelyi (US FDA), Dr. Joao Pedras-Vasconcelos (US FDA), Dr. Haoheng Yan (US FDA), Dr. Meiyu Shen (US FDA), Dr. Mohsen Rajabi Abhari (US FDA), Dr. Isabelle Cludts (UK MHRA), Dr. Elana Cherry (Health Canada), Dr. Lucia Zhang (Health Canada), Dr. Akiko Ishii-Watabe (Japan MHLW), Dr. Sara Gagneten (US FDA), Dr. Andrew Exley (UK MHRA), Dr. Therese Solstad (EU EMA/Norway NoMA), Dr. Richard Siggers (Health Canada)
- **Biomarkers:** Dr. Yow-Ming Wang (US FDA), Dr. Abbas Bandukwala (US FDA), Dr. Kevin Maher (US FDA), Dr. Yoshiro Saito (Japan MHLW)

All the traditional "*working dinners*" attended by both industry key opinion leaders (KOL) and regulatory representatives were held in a virtual format this year, and the extensive and fruitful discussions from these "*working dinners*" together with the lectures and open panel discussions amongst the presenters, regulators and attendees culminated in consensus and recommendations on items presented in this White Paper.

A total of 167 recent issues ('hot' topics) were addressed and presented in this White Paper, which are the background on each issue, exchanges, consensus and resulting recommendations on these one hundred and sixty-seven topics.

Due to its length, this comprehensive White Paper has been divided into three parts for editorial reasons. This publication covers **Part 3** recommendations.

Part 1 – Issue 4 – February 2021

Hybrid Assays and HRMS

- BMV of Biotherapeutics by LCMS and Hybrid Assays: Regulatory Rigor & Acceptance Criteria While Waiting for the ICH M10 Guideline (six topics)

- Hybrid Assays for adeno-associated virus (AAV) Gene Therapy & Extracellular Vesicles: Advanced Applications (four topics)
- Hybrid Assays for Target Engagement: Novel Applications (three topics)
- High Resolution Mass Spectrometry for Protein Therapeutic Bioanalysis: Current Developments (three topics)

Small Molecules Innovation, Peptides and Oligos

- Microbiome Contributions to Small Molecule Drug Metabolism and its Impact on Bioanalytical Assays (six topics)
- Acoustic-Mass Spectrometry (MS) for Bioanalytical Applications (four topics)
- High Resolution Mass Spectrometry (HRMS): Small Molecule Method Development Strategies (four topics)
- Design of Experiments for Therapeutic Peptides: Modern Discovery Bioanalytical Laboratories (three topics)
- Oligonucleotides and Chain-Shorted Metabolites: Advanced Strategies (four topics)

Regulatory Challenges in Mass Spectrometry

- Data Integrity and Regulatory Factors to Consider when Using Cloud Computing (three topics)
- Impact of Excipients on Bioanalytical Methods: "What are regulators asking?" (two topics)
- Parallelism Evaluation in Small Molecule Endogenous Compounds – New Considerations on ICH M10 Guideline (three topics)
- Abnormal Internal Standard Response: Compliance with the 2018 FDA Guidance and ICH M10 Guideline (four topics)
- Microsampling in Regulated Bioanalytical Juvenile & Pediatric Studies (three topics)

Part 2A – Issue 5 – March 2021

Biomarker Assay Validation (BAV)

- Need for a BAV Guidance (three topics)
- When Clinical Biomarker Assay Should be Under Clinical Laboratory Improvements Amendments (CLIA)? (three topics)
- Current Applications of Context of Use (COU) in Fit-for-Purpose (FFP) BAV (six topics)
- Advancements in Extracellular Vesicles (EV) (two topics)

PK LBA Regulated Bioanalysis

- Stability Testing of Biotherapeutics: FDA, EMA, & ICH M10 Guidance/Guideline (two topics)
- Critical Reagents: Latest Approaches (six topics)
- Bispecific Monoclonal Antibodies & Bispecific T-cell Redirectors: Unique Challenges in Pharmacokinetic (PK) Assays (three topics)
- Common Issue with Laboratory Information Management System (LIMS) Based Software for Ligand Binding Assay (LBA) Support (two topics)
- Parallelism Evaluation in Regulated Bioanalysis for PK LBA: FDA, EMA, & ICH M10 Guidance/Guideline (four topics)

Flow Cytometry Validation

- Flow Cytometry Validation: Applicability of Clinical and Laboratory Standards Institute (CLSI) H62 Guideline to Regulated Bioanalysis (four topics)
- Flow Cytometry Validation: Target Engagement and Receptor Occupancy (three topics)
- Flow Cytometry Validation Strategies for Assays Using Challenging Sample Types (two topics)
- Validation Strategies for Image Cytometry Based Assays (three topics)
- Validation of Lower Limit of Quantitation (LLOQ) in Flow Cytometry (two topics)

Cytometry Innovation

- New Insights in Automated Gating (three topics)
- Advantages & Challenges in using Mass Cytometry (CyTOF) (five topics)
- High Dimensional/High Parameter Flow Cytometry (five topics)

Part 2B – Issue 5 – March 2021

- Input from Regulatory Agencies on Bioanalysis & BMV
- Input from Regulatory Agencies on Immunogenicity & Biomarkers

Part 3 – Issue 6 – March 2021

Vaccine Clinical Assays and Cell Therapy

- Clinical Vaccine Assay Validation (six topics)
- Quality control (QC) Samples in Chimeric Antigen Receptor T Cells (CAR-T) & Vaccine Flow Cytometry Assays: Current Industry Standards (three topics)
- Quantitative (real-time) Polymerase Chain Reaction (qPCR) Assays for CAR-T Programs (three topics)
- Immunogenicity Strategy for CAR-T Products (three topics)

Gene Therapy, qPCR and ELISpot Validation:

- qPCR, Droplet Digital Polymerase Chain Reaction (ddPCR), and Next Generation Sequencing (NGS) Assay Development & Validation: Best Practices (five topics)
- AAV Capsid Neutralizing Antibody (NAb) Assays Development and Validation (three topics)
- ELISpot & Single Cell Western Blot Assay Validation (three topics)
- Application of Current FDA/EMA Immunogenicity Guidance/Guideline to Gene Therapy (two topics)

NAb Assay Harmonization, Biosimilars and FDA/EMA Guidance/Guideline:

- Cell-based NAb Assays – Sensitivity and Drug Tolerance and the Relevance for Clinical Outcome (four topics)
- NAb Assay Harmonization: Recent Trends and Expectations (four topics)
- Biosimilar Immunogenicity: Current Industry Standards (three topics)
- The 2019 US FDA Immunogenicity Guidance: Reflections a Year Later (three topics)

Immunogenicity Assay Strategies:

- Lessons Learned from Late-Stage Clinical Studies (four topics)
- Circulating Immune Complexes (three topics)
- Multi-Domain Biotherapeutics: Immunogenicity Assay Strategies (four topics)
- Definition of Persistent Anti-Drug Antibody (ADA) Responses and its Clinical Relevance (four topics)

SECTION 1 – Vaccine Clinical Assays and Cell Therapy

Bart Corsaro[1], Tong-yuan Yang[2], Rocio Murphy[3], Ivo Sonderegger[10], Andrew Exley[7], Sylvie Bertholet[1], Naveen Dakappagari[5], Francis Dessy[6], Fabio Garofolo[8], Lisa Kierstead[9], Holger Koch[10], Ghanashyam Sarikonda[5], Natasha Savoie[11], Richard Siggers[12], Therese Solstad[13]

Authors in Section 1 are presented in alphabetical order of their last name, with the exception of the first five authors who were major contributors.

The affiliations can be found at the beginning of the article

DISCUSSION TOPICS & CONSOLIDATED QUESTIONS COLLECTED FROM THE GLOBAL BIOANALYTICAL COMMUNITY

The topics detailed below were considered as the most relevant 'hot topics' based on feedback collected from the 13th WRIB attendees. They were reviewed and consolidated by globally recognized opinion leaders before being submitted for discussion during the 14th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Clinical Vaccine Assay Validation

What validation parameters should be evaluated when transitioning a previously validated vaccine assay into a multiplexed format? How are quality control samples used to monitor assay trending performance? What data is needed when bridging to a new critical reagent in vaccine assays? What are the recommendations on the best practice for standardizing processes for vaccine clinical assay qualification and validation of established immunoassays? How can we overcome linearity issues in vaccine assays, which are sometimes a mixture of different aspects like parallelism,

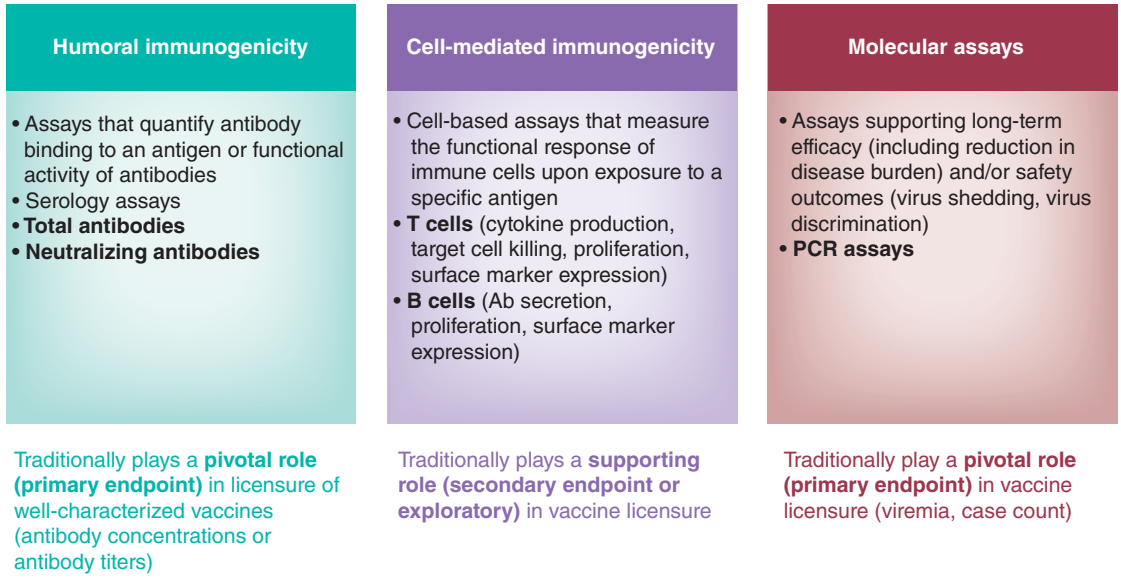


Figure 1. Different Types of Bioanalytical Assays in Vaccine Development

relative accuracy, and dilution? What are strategies for short and long-term stability assessment for vaccine induced antibodies?

Assessment of QC Samples in CAR-T & Vaccine Flow Cytometry Assays: Current Industry Standards

What are the options when using test specific controls? Are they practical and meaningful? Are system level controls routinely used in clinical diagnostic testing applicable for CAR-T monitoring and appropriate for clinical trial submissions? What can CAR-T flow cytometry assays learn from experience with vaccine assay setup?

qPCR Assays for CAR-T Programs

What are considered critical reagents in qPCR assay? How do we perform stability for critical reagents for qPCR? Can we work with WHO to issue a universal human genomic DNA (gDNA) standard so we can use it to calibrate our qPCR assay across the industry by measuring the house keeping gene?

Immunogenicity Strategy for CAR-T Products

Is there agreement that it is straightforward to monitor host humoral responses considering that chronic impact of ADAs on CAR-T persistence is largely unknown or that data have yet to be seen on whether human scFv can result in lower ADAs and better product regarding persistence? Positive host cellular immune responses have been confirmed in treated subjects and seem to be correlated with clearance. Will the use of human scFv as CAR reduce cellular immune response rate? Can lessons on immunogenicity from autologous CAR-T products be applied to allogenic CAR-T products? What are the "right" assay(s) to monitor host cellular immune responses? How do we conduct fit for purpose validation?

DISCUSSIONS, CONSENSUS AND CONCLUSIONS

Clinical Vaccine Assay Validation

Immunogenicity measurements are key assessments for the clinical development of biotherapeutics and vaccines. While similar assay technologies are applied by the two fields, the goal of the clinical testing and the strategies applied are different. While the immunogenicity assessment for biotherapeutics is focused on the detection of very small amounts of ADA (sensitivity), the immunogenicity assessment for vaccines is more focused on the reproducible quantitation of the antibody response to the vaccine. For this reason, vaccine immunogenicity assessments typically do not use the tiered approach, which is an essential element for biotherapeutic immunogenicity assessments. Many different types of bioanalytical assays are used to assess endpoints in vaccine clinical development (refer to Figure 1); this white paper will focus on providing recommendations on serology assays of total antibodies to

Table 1. Parameters for vaccine serology assay qualification and validation.

Parameter	Assay qualification	Full validation
LOB, LOD	X	X
Precision	X	X
Linearity	X	X
LLOQ and ULOQ	X	X
Accuracy	X	X
Specificity	X	X
Analytical range	X (only for quantitative assay)	
Robustness	†	
Ruggedness	†	
Interferences	†	
Sample stability	†	

†It is recommended that parameters like robustness, ruggedness, interference, and sample stability are assessed as early as possible during the validation process (optimally during qualification). They should be included in the validation if they were not assessed during the qualification.

measure humoral immunogenicity.

Regulatory guidelines and multiple white papers are available to guide the pharmaceutical industry to adequately validate ADA assays [18,21,25–27]. Guidelines and white papers specific for vaccine serology assays, on the other hand, are limited and would be very useful to understand general regulatory expectations regarding assay development, assay validation and clinical immunogenicity testing for vaccine clinical candidates early in development. The 2019 White Paper in Bioanalysis [25] began building the foundation of a framework to define the expectations for vaccine assay qualification, validation, and life cycle management. The 2020 consensus developed the 2019 framework and further focused on providing more detailed and practical recommendations for clinical vaccine assay development and validation. Together, the 2019 & 2020 White Papers were designed to provide a harmonized reference for vaccine assay validation.

Regulatory agencies may request the demonstration of assay consistency over the life of the development program from applicants. In response to this regulatory request, it is recommended to use a phased approach to clinical vaccine assay development, separated into 3 distinct phases: 1) assay setup (establish assay format and run parameters), 2) qualification (determine assay performance), and 3) validation (confirm performance in "real life" conditions, with pre-defined acceptance criteria).

Assay Setup

The goal of the initial (set-up) phase of assay development is to identify the format of the assay and potential critical reagents. Understanding the pathogen and the immune response to the vaccine will lead to selection of relevant assays predictive of clinical benefit. The intended use of the assay needs to be clearly defined before selecting and optimizing the methodology. During this phase, development can be limited by sample availability. A barrier to the development of these assays is the lack of appropriate reference standards and positive control samples, particularly those which mimic the matrix and antibody profile of the samples to be tested prior to the first-in-human studies. Serum from pre-clinical studies (e.g., vaccinated laboratory animals) may provide similar breadth of the expected response and high titered serum for understanding the initial assay parameters. However, serum from naturally infected individuals, although not mimicking the antibody profile, may be a better representation of the anticipated human sample matrix. While neither sample is ideal for in-depth clinical assay development, they provide useful and necessary tools. This development phase is fit for a Phase I clinical trial.

Assay Qualification

Qualification is planned and conducted when sufficient human incurred samples are available. During qualification, all relevant assay performance criteria are evaluated. The selection of the qualification parameters is dependent on the use of the assay. These attributes typically include precision, linearity, specificity, accuracy (as applicable), limit of detection (LOD), and analytical range, i.e., LLOQ and upper limit of quantitation (ULOQ). In addition, assay robustness (acceptable variations in incubation times and temperatures) and ruggedness (impact of days, analysts

and reagent lots) are ideally evaluated during this phase. Furthermore, sample stability (e.g., short-term stability) may also be evaluated (see analyte stability assessment below).

Linearity

It is recommended to perform the linearity assessment for quantitative vaccine assays. Linearity is a mixture of different aspects like parallelism, relative accuracy, and dilution, depending on the assay format. There is a difference between linearity over the signal range (i.e., from one starting dilution, similar to parallelism) and linearity over the assay concentration range (i.e., from several starting dilutions). To assess dilutional linearity, a dose proportionality approach is recommended. Samples covering the range of interest are serially diluted (independent dilutions) in multiple replicates. For each sample, the dose proportionality is assessed assuming a power model ($10^{\alpha} Dilution^{\beta}$)

$$log_{10}(Result) = \alpha + \beta \log 10\,(Dilution)$$

Based on a criterion for the ratio between the dilution corrected extremes of the range considered, an acceptance range is computed for the slope in which the 90% confidence interval must fit. The range over which dose proportionality is demonstrated is obtained by a recursive search [28]. Linearity can be determined using 3–5 samples covering the range of interest, serially diluted 4–5 times (independent dilutions) with negative sera. The samples should be run by a minimum of 2 analysts on 2 days at least in duplicate on each day. The expectations for the variability of back calculated titers should be determined.

Specificity

Analytical specificity is the ability of an analytical method to detect the analyte of interest; only the component it purports to measure or the extent to which the assay responds to all subsets of a specified analyte and not to other substances present in the sample. Specificity can be tested with pre-absorption experiments, where the test-serum is pre-incubated with either homologous or heterologous antigen. Initial experiments may use a limited sample set with multiple competitor protein concentrations. Final experiments should use a greater number of samples (5–10 samples) with a single competitor protein concentration. Optical density (OD) or Mean Fluorescent Intensity (MFI) of the sample tested needs to be within the linear range of the assay. The percent inhibition can be used to report specificity results. Multiplexed assays add a second layer of complexity to the specificity testing as the assay must be specific for the homologous protein, but not compete with other measured, heterologous antigens (see multiplex assays).

Limit of Detection

The LOD of an assay is defined as the lowest concentration that has a high probability of producing a response that will be distinguished from the background response (i.e., the response at zero concentration) as determined in a limit of blank (LOB) experiment. It can be determined by spiking 2-fold serial dilutions of a reference standard (where it exists) into negative or antibody-depleted human serum (ADHS). The LOD can also be determined considering the assay precision in the very low concentration range near the LOB.

Precision (Intra-, Inter- and Total Assay Variability)

Intra-assay variability (repeatability) represents within-run variation, while inter-assay variability represents the between-run variation (intermediate precision) attributable to different days, analysts, reagents, etc. The estimate of assay precision may be used to establish acceptance ranges for control samples, and to calculate a statistically meaningful fold-increase in antibody titers for an individual sample. The precision evaluation should test multiple incurred samples a minimum of 6 times.

Analytical Range

The LLOQ and ULOQ define the antibody concentration range over which the assay is acceptably accurate and precisely quantitates samples. LLOQs can be determined by evaluating the precision profile and assay relative accuracy. A recommendation is that at least 80% of the samples within the LLOQs must have variability estimates $<$20% relative standard deviation (RSD) for a standard antibody binding assay. Ideally, incurred samples are used for the evaluation of precision and relative accuracy because they best reflect the variability of the polyclonal immune response after vaccine administration.

Reference Standards

Assay maintenance activities need to ensure reproducible and precise titer determination over multiple years. A reference standard aligned with an international standard (when available) should be considered. When an international reference standard is unavailable, a high titered pool of samples can be used instead. Arbitrary units to define the standard can be used (e.g., C-value of the 4PL curve). However, it is important that test and reference sera have a parallel dilutional response curve. This parallelism allows sample concentration calculations over the widest range of the standard curve and gives assurance to the relationship of the sample values.

Qualification Reporting and Use

Following the completion of the qualification phase, it is recommended to write a full analytical development report, detailing experiments, experimental design and results (passed and failed). The report should conclude with recommendations for assay validation acceptance ranges. A detailed standard operating procedure (SOP) based on the final procedure used in the assay qualification is required and must be used in an assay validation. A qualified assay is suitable for the analysis of primary and secondary clinical endpoints of Phase I and Phase II clinical trials. Qualified assays may also be used for exploratory clinical endpoints of late-stage trials (Phase III). The interpretation of data from the qualified assay will help establish the design of the late Phase II/Phase III clinical trials.

Assay Validation

If the assay setup and qualification work is completed with a high level of quality, the final development step, assay validation, should be relatively easy and short compared to the initial development phases. Assay validation requires the testing of the assay performance against predefined acceptance criteria optimally using incurred samples spanning the entire analytical range, which are representative of the Phase III program. These criteria are established based on the results from the assay qualification data and the intended use of the assay. Parameters evaluated during validation include: precision, linearity, specificity, accuracy, LOD, quantification range (LLOQ and ULOQ). If any of the validation parameters fail, the validation is considered a failure and the cause of the failure should be reported and investigated. Validated assay support testing for late-stage clinical trials may be reviewed by regulatory agencies prior to Phase III testing.

Regulatory Interactions

Health authorities often request to review validation plans and/or qualification reports before validation and/or to confirm that assays are suitable for testing pivotal clinical study samples. Specific guidance on vaccine immuno-assay validation, issued by regulators, would help decrease the need for such pre-validation regulatory interactions.

Assay Life-Cycle Management

Since data for licensure of vaccines are generated throughout the clinical development program, it is necessary to demonstrate the stability of the assay performance over multiple years. To accomplish this need, many assays will use three quality control systems: implementation of an assay standard, trending, and proficiency panel testing. Assays may develop over time due to changes in conditions or reagents. Therefore, as part of the life cycle management of the assay, the validation should be periodically reassessed to determine if any additional validation work is required. Changes to the assay that may affect assay performance (e.g., new testing laboratories and changes in test samples, new age groups, specific disease populations) may require a partial validation or a full assay revalidation.

Analyte Stability Assessment

Short-term stability experiments assess whether sample handling and storage affect the assay results. The experiments should mimic conditions that are encountered during clinical testing (e.g., freeze/thaw cycles, short term storage at 2–8°C or room-temperature storage). Short-term stability can be assessed early during assay development (e.g., during qualification). If the assay is planned to be further developed, stability may be assessed at a later stage. It has been shown that antibodies are stable in serum or plasma stored at -20°C and -80°C beyond 3–4 years. For this reason, long-term stability studies for frozen matrices may not be required [29,30].

Ideally, samples used for stability experiments would cover the quantifiable range of the assay because incurred samples best reflect the heterogeneity and the matrix of samples used for clinical testing. If insufficient sample volume is available from clinical studies, spiking of negative samples or pooling of positive samples may be considered. When taking the complexity of multiplex assays into account, the number of stability samples per assay subtype

might be reduced focusing on concentrations close to the LLOQ of the assay. It is, however, expected that every assay subtype is evaluated as part of the stability study.

Other stability parameters related to the setup and design of the assay should also be considered in a stability program. Critical reagents such as antigen coated plates or beads, and sample predilutions which can be created to enhance assay efficiency must be evaluated to demonstrate that routine experimental practices do not impact assay results.

Long-Term Assay Control

Long-term assay control is important in order to guarantee comparability of test results of long clinical trials and to allow the comparison of data between multiple clinical trials. Two tools may be implemented in order to achieve long-term assay control: assay trending and proficiency panel testing.

Assay trending is performed by analyzing run acceptance QC samples that are measured during routine clinical testing. Trending limits are stricter than run acceptance limits and serve as an early indicator for assay drift. It is recommended to pre-define the trending analysis in a trending plan. This document should describe the trending limits for the QCs, how often the trending is analyzed, the minimal amount of runs per trending-period, and how to react in case of out-of-trend events.

Proficiency panels are additional tools to monitor long-term assay performance. It is recommended to pre-specify the setup, characterization, and the testing of the proficiency panel in a plan. Ideally, a sufficient number of incurred samples that span the quantification range are analyzed multiple times on a regular basis. The volume of samples must be sufficient to run the proficiency panel over multiple years. Results from the panel should be reported in a formal report.

Critical Reagents

Critical reagents are reagents that may impact assay performance. The critical reagents should be identified during assay development and documented as part of the qualification and validation. A process for the bridging of critical reagents prior to use in clinical sample testing is important to guarantee a constant assay performance and avoid drift.

Ideally a new (candidate) critical reagent lot should be qualified by comparing its performance head-to-head to a qualified reagent lot. Minimally, the performance of QC samples should be assessed to qualify a new reagent lot. Optimally a panel of incurred samples is tested over multiple independent runs with the candidate and the qualified reagent lot. In some instances, the qualified lot is not available anymore (e.g., if expired). In this case, a comparison to historically generated data may be considered. The acceptance criteria for the critical reagent qualification should take into consideration the assay performance and the intended use of the assay. Keeping some of the reagents used during qualification/validation can help to assess whether the assay has changed its performance. This is only possible if the critical reagents have a long shelf-life.

Multiplex Testing

Multiplex testing allows the reduction of sample volume. This volume reduction is critical for pediatric studies but may also be beneficial in adult studies. Transitioning from a single-plex to a multiplexed format poses some challenges. Meaningful oversight on robust assay performance and successful critical reagent bridging strategies are more difficult to establish when working with multiplexed assay formats.

It is recommended that all validation parameters are redeveloped and revalidated when transitioning a previously validated vaccine assay into a multiplexed format. Specificity and cross-reactivity will have a major impact on a multiplex format and necessitate re-assessing antigen concentrations and potentially to establish a new reference standard. However, if the change is restricted to the read-out only, then partial validation may be sufficient.

If bridging needs to be established to the previous assay format there should be pre-defined acceptance criteria. Clinical endpoints and data interpretation may be relevant in defining the acceptance criteria of the assay bridging.

A change of platforms is not recommended during Phase III because equivalency may not be possible to demonstrate.

Assessment of QC Samples in CAR-T & Vaccine Flow Cytometry Assays: Current Industry Standards

CAR-T therapies present unprecedented opportunities and challenges for bioanalytical scientists. We are witnessing an explosion in the numbers of next generation CAR-T agents and new clinical trials which is generating an increasing interest in standardizing assays and ensuring quality control. Flow cytometry assays are a key methodology

Table 2. System Level QC Practices for CAR-T Monitoring Assays.

Parameter	Quality control practice
Cytometer	Acceptable cytometer setup bead performance 'releases' the instrument for clinical use
	Target channel (MFI) verification (e.g., Euroflow beads) confirms standardized MFI output across all cytometers in-use
Antibodies	'New' Individual and 'cocktailed' antibodies should demonstrate comparable specificity and MFI output against 'In-use' antibodies prior to clinical trial implementation. Only used within the specified expiry period
Reagents	Only pre-qualified critical buffers such as Lysis & Staining reagents are used. Examples of pre-qualification using 'test' specimens include verification of acceptable viability or pH

used for monitoring PK/cellular kinetics and efficacy of CAR-T therapies in clinical trials and are also growing in popularity for vaccine trials. Flow cytometry can document vaccine-induced versus natural immunogenicity; vaccine take and response rate (efficacy); support the justification of the final vaccine formulation; and demonstrate non-inferiority versus other vaccines. High parametric flow cytometry can be applied to clinical trial T-cell exploratory endpoints but requires rigorous fit-for-purpose instrument optimization, antibody (Ab) panel design, sample preparation, assay setup and data analysis.

There is currently no finalized regulatory guidance for general flow cytometry assay validation or specifically for the measurement of CAR-T levels by flow cytometry, prompting a consortium of experts from the International Clinical Cytometry society and CLSI to develop guidelines [31]. The main regulatory concern is to demonstrate that the CAR-T assays are actually measuring the desired analyte and QC being used to prove this is representative of samples. To address this, the 2019 White Paper in Bioanalysis recommended the use of QC samples relevant to the cell population of interest such as stabilized whole blood, cryopreserved PBMCs, or "spiked" QCs. Implementation of QCs was recommended to at least periodically track that the assay is performing consistently [25].

Based on real-world experiences encountered with the first approved CAR-T therapy, the 2020 recommendations are focused on describing the practical limitations associated with implementation of traditional QCs, while proposing alternate approaches to ensure the quality of these high complexity flow cytometry assays. These approaches include instrument standardization, appropriate panel design to exclude unwanted cells, personalized gating controls, viability dyes, and bead-based or volumetric approaches for cell counting. Regulatory considerations and future perspectives were also discussed. Two types of controls are currently utilized: test specific controls and system level controls.

Test specific controls are QCs with varying CAR-T levels including transfected cell lines, transduced healthy donor cells, or patient drug product, each with their own benefits and limitations. Transfected cell lines are an excellent option for creating homogenous controls expressing known levels of CAR-T construct. However, these do not contain additional T cell markers required for developing gating approaches for patient samples. While transduced healthy donor T cells do not exhibit light scatter properties of patient specimens exposed to chemotherapy, they are the most practical option for establishing initial gating methods and validation parameters. Patient specific T cells or drug product are ideal controls, but their availability is limited for CAR-T monitoring assays due to their prioritized use in product release assays and treatment of patients. Reliability of test specific controls is based on 3 key factors: (1) instrument, (2) stability of detection reagent and (3) stability of test specific control that should be monitored. A central lab should be used to standardize protocol, using the same instrument type and optical bench configuration, the same cell line, and the same lots of critical reagents to harmonize results across all sites.

System level controls are routinely used in clinical diagnosis and are thus reliable for monitoring CAR-T PK (see Table 2); these approaches have received support from subject matter experts representing FDA and MHRA. Fluorescence Minus One (FMO) controls are used as gating controls to identify the gating boundary for the one antibody that is missing and used to identify background staining due to fluorescence spillover. Isotype controls may be used for identifying nonspecific staining. Instrument QC controls and experimental controls (positive, negative) are important to ensure the assay is performing as expected. Certified participation in a relevant external quality assessment scheme is essential. The use of a single central laboratory is recommended if possible, or evidence to support comparability of results across participating laboratories is required. QCs are not necessary on every sample run but should be used often enough to demonstrate reliable results and should undergo QA oversight. For cocktail antibodies, if they are prepared in-house, stability needs to be demonstrated. If they are sourced from a

vendor, the provided stability can be leveraged. In addition, lot to lot bridging study is crucial to qualify the release of new reagent lot.

CAR-T Flow Cytometry assays can also benefit from vaccine experience with instrument and assay setup. It is recommended to use latest peer-reviewed guidelines for the use of flow cytometry and cell sorting in immunological studies to support CAR-T assay development [32]. Instrument setup should be based on the stain index measurement for each detector to determine the best sensitivity and minimize the spillover/spread matrix (SSM). Daily performance checks of instrument precision and sensitivity [33,34] using 3 sets of beads (neutral comp, single peak and rainbow beads) are also needed. It is important to check linearity of the photomultiplier tube (PMT) response. Assay setup defines protocol optimization, background evaluation, gating strategy, specificity, precision, linearity and LOB/LOD. CAR-T detection is usually done together with other cell surface markers in a multicolor, multiplexing fashion. Potential interference from other staining reagents should be investigated thoroughly. In the absence of SRMs for particular flow cytometric applications, it is challenging to perform some performance measures. It is recommended to design appropriate approach and ask regulatory agency for early feedback.

qPCR Assay for CAR-T Programs

Because of its high sensitivity and sampling convenience, qPCR is the most commonly used methodology (even if the use of ddPCR is increasing) for monitoring the fate of CAR-T cells and is especially useful for monitoring low quantities of CAR-T cells as part of long-term studies. Optimal qPCR primers can detect the CAR-T inserted transgene. There is limited regulatory guidance and qPCR method development and validation to support regulated bioanalysis for CAR-T therapies in clinical studies. Discussions built upon the 2019 White Paper recommendations on qPCR validation [25] which recommended following scientifically-led method development and validation strategies, with support from the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) and CLSI guidance [35,36]. A two-phased approach was suggested with initial qualification to determine what performance characteristics can be achieved, followed by validation against pre-defined criteria.

The recommended validation parameters for qPCR assays for CAR-T programs includes sensitivity or limit of quantitation (LLOQ; 50 copies/μg gDNA) and LOD, intra- and inter-assay precision, accuracy, DNA extraction efficiency from tissues, and engineering controls to ensure there is no cross contamination, PCR efficiency, PCR linearity, specificity and selectivity, and robustness. Currently, there is no requirement for performing incurred sample reproducibility (ISR).

For accuracy and precision, QCs can be plasmid spiked into human gDNA for transgene or plasmid only for reference gene or spiked into surrogate nonhuman gDNA. CAR-T from normal donors spiked into diseased whole blood can demonstrate intra- and inter-assay variability while accuracy can be inferred from reference gene measurement from a qualified lot of human gDNA.

Recommendations were also provided for stability assessment of critical reagents which were identified as primers, probes, reference standards, master mix, and positive control (cell lines). To assess their stability, it was recommended to establish a critical reagent qualification program for lot-to-lot bridging, stability, etc.

Finally, the question was raised whether the industry should work with WHO to issue a universal human genomic DNA standard to calibrate qPCR assays across the industry by measuring the house keeping gene. Two options were suggested for these standards. The first was to consider NIBSC wild-type standard 18/164 [widely used as a standard in cancer genome testing] to calibrate CAR-T assays once the copy-number of the reference gene of choice has been characterized. The second proposed the use of WHO 1st International Reference Panel (19/158) for the quantitation of Lentiviral Vector Integration Copy Numbers which will be released shortly.

Immunogenicity Strategy for CAR-T Products

Persistence of CAR-T cells in the subject's circulation plays a critical role in long term efficacy while it could also pose a potential long-term safety risk during treatment and after remission. Immunogenicity to CAR-T products is expected to be more complex when compared to protein biologics. It can be generated from host humoral and cellular responses due to the unique CAR-T product structure and design. Information on immunogenicity to CAR-T products is quite limited due to the fact that most of CAR-T products are still under clinical development [37]. Potential clinical consequences of immunogenicity are largely unknown and are currently still being monitored in clinical studies for each product.

The discussion centered around the bioanalytical strategies and fit for purpose experimental methodologies applied to monitor clinical immunogenicity of CAR-T products based on risk factors and product structure/design.

Host humoral immune responses can be measured with LBA-based or cell-based formats [38,39]. FDA still recommends that this assessment be performed for autologous CAR-T cells. Allo-CARs may induce cellular response. There are limited data on the impact of humoral/cellular response on CAR-T cells safety, efficacy and persistence. Many other transgenes are expressed; humoral responses to all the transgenes are not studied and the impact is unknown. There is a low incidence of cellular immune response against CAR, and no clear relationship between cellular immunity and clinical outcomes exists.

Positive host cellular immune responses have been confirmed in treated subjects and seem to be correlated with clearance [40]. However, supporting data is lacking as many studies do not assess cellular immunogenicity. It was discussed if the use of human scFv as CAR could reduce cellular immune response incidence. It was concluded that CAR can be immunogenic, regardless of the species of origin of the scFv, because foreign sequences are expressed, or novel epitopes are created. Therefore, an immunogenicity assessment is recommended for humanized CAR-T cell therapy. The risk lies in the full length of the CAR construct.

It was also discussed whether lessons on immunogenicity from autologous CAR-T products can be applied to allogenic CAR-T products and how these assays are developed and validated. It was recommended to focus on the clinical problems when developing an assay, such as the context of adverse events or failure of efficacy and how immunogenicity could be implicated in those issues. Generally, immunogenicity concerns are greater for allo-CARs. Therefore, it is recommended to adopt similar approaches to autologous CAR-T cells. For cellular immune response assays (for CAR T and AAV-based gene therapies), both ELISPOT and flow cytometry have been used [41,42].

RECOMMENDATIONS

Below is a summary of the recommendations made during the 14th WRIB:

Clinical Vaccine Assay Validation

- Specific Industry/Regulator **recommendations for vaccine serology assays** would be very helpful to understand general regulatory expectations regarding assay development, assay validation and clinical immunogenicity testing for vaccine clinical candidates early in development
- **Assay consistency** should be demonstrated over the life of the development program
- It is recommended to use a phased approach to clinical vaccine assay development, separated into 3 distinct phases: 1) **assay setup** (establish assay format and run parameters), 2) **qualification** (determine assay performance), and 3) **validation** (confirm performance in "real life" conditions with pre-defined acceptance criteria). Refer to Table 1 for parameters recommended for each phase.
- The intended use of the assay needs to be clearly defined before selecting and optimizing the methodology during **assay setup**.
- **Assay qualification** can only be planned and conducted once sufficient human incurred samples are available.
    - Perform the **linearity** assessment for quantitative vaccine assays using 3–5 samples covering the range of interest, serially diluted 4–5 times (independent dilutions) with negative sera. The samples should be run by a minimum of 2 analysts on 2 days at least in duplicate on each day. The expectations for the variability of back calculated titers should be determined.
    - To assess **dilutional linearity**, a dose proportionality approach assuming a power model ($10^{\alpha} Dilution^{\beta}$) is recommended.
    - Test **specificity** on a minimum of 5–10 samples covering the analytical range in a competition experiment using homologous and heterologous protein. OD or MFI of the sample tested needs to be within the linear range of the assay. The percent inhibition can be used to report specificity results.
    - **LOD** is set at the lowest concentration that has a high probability of producing a response that will be distinguished from the background response as determined in a LOB experiment by spiking 2-fold serial dilutions of a reference standard (where it exits) into negative or ADHS or determined considering the assay precision in the very low concentration range near the LOB
    - **LLOQ** can be determined by evaluating the precision profile and assay relative accuracy. At least 80% of the samples within the LLOQs must have variability estimates <20% RSD for a standard antibody binding assay.
    - **Precision** should test multiple incurred samples a minimum of 6 times.

  - Write a full **analytical development report**, detailing experiments, experimental design and results (passed or failed). The report should conclude with recommendations for assay validation acceptance ranges.
  - A **detailed SOP** based on the final procedure used in the assay qualification is required and must be used in an assay validation.

- **Assay validation** requires the testing of the assay performance against predefined acceptance criteria optimally using incurred samples spanning the entire analytical range, which are representative of the Phase III program.
  - If any of the validation parameters fail, the validation is considered a failure and the cause of the failure should be reported and investigated.
  - Assay validation may **periodically be reassessed** to determine if any additional validation work is required. Changes to the assay could trigger assay revalidation (e.g., new testing laboratories and changes in test samples).
- Standard and test samples need a parallel response from several points on the curve and use of arbitrary units to define the standard can be used (e.g., C-value of the 4PL curve).
- Formal **trending plans** and **proficiency panel testing** are recommended elements of the serology assay maintenance strategy.
- **Short-term stability** experiments should mimic conditions that are encountered during clinical testing (e.g., freeze/thaw cycles, short term storage at 2–8°C or room-temperature storage).
- **Long-term stability** of antibodies is generally accepted in serum or plasma stored at -20°C and -80°C beyond 3–4 years. For this reason, long-term stability studies for frozen matrices may not be required.
- **Regulatory agencies** may be consulted to confirm agreement on the design and acceptance criteria proposed for validation.
- Trending plans
  - Performed by analyzing run acceptance QC samples
  - Trending limits are stricter than run acceptance limits
  - Trending plan defines limits, frequency, and when investigation is required
- Proficiency panels
  - Titers must cover the entire analytical range of the assay.
  - Volume of samples must be sufficient to perform multiple runs.
  - Results from the panel should be reported in a formal report.
- To define acceptance criteria for bridging **critical reagent** lots, consider both the intended use of the assay and the assay variability.
  - Minimally QC samples should be used to qualify a new reagent lot.
  - Testing should be completed across multiple independent runs with the candidate and qualified lot
  - If the qualified lot is not available anymore (e.g., if expired), a comparison to historically generated data may be considered
- It is recommended that all validation parameters should be redeveloped and revalidated when transitioning a previously validated vaccine assay into a **multiplexed format**.
  - Re-assess antigen concentrations and potentially to establish a new reference standard.
  - If bridging needs to be established to the previous assay format there should be pre-defined acceptance criteria.
  - A change of platforms is not recommended during Phase III because equivalency may not be possible to demonstrate.

Assessment of QC Samples in CAR-T & Vaccine Flow Cytometry Assays: Current Industry Standards

- High parametric flow cytometry can be applied to clinical trial T-cell exploratory endpoints but requires fit-for-purpose albeit rigorous instrument optimization, Ab panel design, sample preparation, assay setup and data analysis.
- Two types of controls are currently suggested: **test specific controls** and system level controls.
- Patient specific T cells or drug products are the most ideal option for use as test specific controls but have limited availability.
    - Reliability of test specific controls is based on 3 key factors: (1) instrument, (2) stability of detection reagent and (3) stability of test specific control that should be monitored.
    - A central lab should be used to standardize protocol, using same instrument type and optical bench configuration, the same cell line, and the same lots of critical reagents to harmonize results across all sites.
- **System level controls** are reliable for monitoring CAR-T PK/cellular kinetics. Refer to Table 2 for recommendations.
- It is recommended to use latest peer-reviewed guidelines for the use of flow cytometry and cell sorting in immunological studies to support CAR-T assay development.
- **Instrument setup** should be based on the stain index measurement for each detector to determine the best sensitivity and minimize the spillover spread matrix (SSM).
- Daily performance checks of precision and sensitivity using 3 sets of beads (neutral comp, single peak and rainbow beads) are needed.
- Potential interference from other staining reagents should be investigated thoroughly.

qPCR assay for CAR-T Programs

- The recommended **validation parameters** for qPCR assays for CAR-T programs includes:
    - Sensitivity or limit of quantitation (LLOQ; 50 copies/μg gDNA) and LOD (range of response),
    - Intra- and inter-assay precision: CAR-T from normal donors spiked into diseased whole blood can demonstrate intra- and inter-assay variability
    - Accuracy: can be inferred from reference gene measurement from a qualified lot of human gDNA
    - DNA extraction efficiency from tissues
    - Engineering controls to ensure there is no cross contamination
    - PCR efficiency
    - PCR linearity
    - Specificity and selectivity
    - Robustness
    - ISR does not need to be performed
- **Critical reagents** were identified as primers, probes, reference standards, master mix, and positive control (cell lines).
- It was recommended to establish a critical reagent qualification program for lot-to-lot bridging, stability, etc.
- NIBSC wild-type standard 18/164 or WHO 1st International Reference Panel (19/158) can be used as a universal human genomic DNA standard to calibrate qPCR assays across the industry by measuring the house keeping gene

Immunogenicity Strategy for CAR-T Products

- Host humoral immune responses should be measured with LBA-based or cell-based formats for autologous CAR-T cells
- CAR can be immunogenic, regardless of the species of origin of the scFv, because foreign sequences are expressed, or novel epitopes are created. Therefore, an immunogenicity assessment is recommended for humanized CAR-T cell therapy. The risk lies in the full length of the CAR construct

- It was recommended to focus on the clinical problems when developing an assay for allogenic CAR-T products, such as the context of adverse events or failure of efficacy and how immunogenicity could be implicated in those issues
- Immunogenicity concerns are greater for allo-CARs. Therefore, it is recommended to adopt similar approaches to autologous CAR-T cells; however, additional considerations may be needed for allogeneic cells

SECTION 2 – Gene Therapy, qPCR, NGS and ELISpot Validation

Yanmei Lu[19], Mark Milton[22], Lisa Kierstead[9], Jean-Claude Marshall[21], Andrew Exley[7], Jason DelCarpini[14], Fabio Garofolo[8], Boris Gorovits[17], Swati Gupta[15], Lynne Jesaitis[16], John Kamerud[17], Arno Kromminga[18], Anna Ma[20], Jim McNally[23], Natasha Savoie[11], Richard Siggers[12], Therese Solstad[13]

Authors in Section 2 are presented in alphabetical order of their last name, with the exception of the first five authors who were major contributors.

The affiliations can be found at the beginning of the article

DISCUSSION TOPICS & CONSOLIDATED QUESTIONS COLLECTED FROM THE GLOBAL BIOANALYTICAL COMMUNITY

The topics detailed below were considered as the most relevant 'hot topics' based on feedback collected from the 13th WRIB attendees. They were reviewed and consolidated by globally recognized opinion leaders before being submitted for discussion during the 14th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

qPCR, ddPCR, and NGS Assay Development & Validation: Best Practices

Do we need to update the 2019 recommendations [25] on qPCR validation to further harmonize industry best practice? Are different companies still using different levels of qPCR validation? Are the principles of FFP validation well understood for qPCR or is there still confusion with the assays to be "characterized, qualified or fully validated"? What is required for those validation levels for clinical testing (CLIA laboratory developed tests (LDT) versus Good Clinical Laboratory Practices)? What assay parameters need to be evaluated for NGS and qPCR-based quantitative insertion–deletion mutations (indels) and integration assays? Is there anything we need to add to the 2019 recommendations on assay parameters or generation and characterization of reference standards and QCs for assay development/qualification and assay performance monitoring [25]? NGS computational analysis provides quantitative %indel results based on sequences without the need of using a standard curve, but back calculating %indel against a standard curve increases assay robustness. Should %indel be adjusted using a standard curve? What are the current best practices for managing ddPCR technical challenges? Does guidance on ddPCR need to be implemented? What is the best practice for manually setting the positive/negative threshold when the analysis software is unable to? Frequently the negative control is zero (no positive droplets detected) with multiple independent analyses. If it is zero, what is the best method to establish the LOD? Is less than 90% amplification efficiency acceptable if the assay is quantitative with an otherwise acceptable performance? Would you use ddPCR to determine both biodistribution and shedding of the vector/transgene?

AAV Capsid NAb Assays Development and Validation

Should anti-capsid antibody assays (anti-drug antibody (ADA) and NAb) be used as exclusion criteria for clinical trials of AAV-based gene therapy (GTx)? How is a clinically relevant cut off determined? What regulatory framework is used for these assays: typical ADA guidance or CLIA? Would these assays be considered companion diagnostics (CDx) when the therapeutic is approved? Should cellular immunity assays (such as enzyme-linked immunospot (ELISpot)) for capsid or transgene product be used as exclusion criteria for clinical trials of AAV-based GTx?

ELISpot & Single Cell Western Blot Assay Validation

Do we need to update the 2019 recommendations on ELISpot validation to further harmonize the industry best practice, duplicate or triplicate analysis, cytokine read out, qualification level and definition of positive ELISpot responses [25]? What assay parameters need to be evaluated for ELISpot validation? How does the amount of variability for these assays factor into a potential validation plan? Is there anything we need to add to the 2019 recommendations on assay parameters of precision, assay range (LOQs), ruggedness, linearity [25]? What are the FFP acceptance criteria for Single Cell Western Assays? Are LBA criteria (20%–25%) suitable?

Application of Current FDA/EMA Immunogenicity Guidance/Guideline to Gene Therapy

Are the current Immunogenicity Guidance/Guidelines (FDA/EMA) for biotherapeutics fully applicable to gene and cell therapy? Are we seeing the implementation of clinically relevant immunogenicity strategies via inclusion criteria? Do we have enough data to know what titers are relevant to impact transduction? What is current industry experience?

DISCUSSIONS, CONSENSUS AND CONCLUSIONS

qPCR, ddPCR, and NGS Assay Development and Validation: Best Practices

qPCR Validation

The number of GTx in development has grown significantly in recent years. AAVs have become the most prevalent delivery vector among viral based GTx. Many of these GTx are aimed at treating rare genetic conditions by introducing a functional copy of the gene to restore the function of the protein. Several regulatory guidelines have been published by the EMA [43], FDA [44] and PMDA [45] to provide industry with general considerations related to AAV GTx clinical development, including discussions and recommendations related to the diverse bioanalytical support needed for these unique therapies.

One of the requirements for AAV GTx clinical development is to understand the shedding kinetics and potential infectivity of viral particles by patients after a single dose of AAV GTx. Vector shedding is a concern as exposure to naïve individuals could induce a NAb response rendering future treatment ineffective. Shedding assays are a requirement of EMA, FDA, and PMDA [43–45]. qPCR assays that quantitatively detect the product specific nucleic acid are highly specific, sensitive, reproducible and high throughput. Due to these advantages, qPCR has been the primary assay even though it cannot differentiate intact versus non-infectious or degraded virus. Secondary infectivity assay development is recommended by EMA [43] and PMDA [45] to ensure that rare recombination events do not occur which can enable infection by these vector genome replication incompetent AAVs. The need for the quantification of shed viral vector in a variety of matrix types is clear; however, the level to which these assays should be validated and the specific development criteria for these assays are less clearly defined. Per FDA [44], often an assay with a quantitative readout, like qPCR is used because of the ease of performing and standardizing the assay, high throughput format, rapid turnaround time, and assay sensitivity.

Regulatory agencies require measurement of shed AAV particles from patients in a variety of matrix types, including whole blood, plasma, saliva, urine, semen, stool and potentially tears depending on route of delivery. FDA and EMA have consistently required urine, stool, and saliva [43,44]. Each matrix type can pose its own unique challenge in terms of qPCR and/or ddPCR assay development and should be considered in the overall assay development and validation [46].

Previous White Papers [21,25] have provided the basis on to approach qPCR assay validation in bioanalytical laboratories where the sensitivity requirements may be different in preclinical vs. clinical assays. Similar to LBA validations, assay performance controls and QCs should be established to monitor accuracy, precision, range of quantification, analytical sensitivity and specificity using clinical material if at all possible. Previous discussions have also given strategies for performing qPCR and infectivity assays in challenging matrices such as urine [47] by using a surrogate marker by staining for a late stage viral replication protein in a cell-based assay. Different companies are using different levels of validation, indicating the need for clear industry/regulator recommendations on which parameters should be evaluated. Regulators recommended the application of general principles such as performing FFP validation based upon COU but provided few details. The following 2020 recommendations are aimed to give further practical guidance into the design of qPCR assays for GTx.

Assay Qualification and Validation

qPCR assays should be **qualified and validated** by the sponsor, with a clear understanding of the assay sensitivity, specificity, reproducibility and variability in each matrix to be tested (Table 3). Fit for purpose validation is generally to verify interference and avoid cross-contamination. Spike recovery with qualified internal control in samples is preferable to assess for interference.

For PCR primer probe set efficiencies, most laboratories use criteria of 90–110%, but some go as low as 85% efficiency. Lower than 85% amplification efficiency is not recommended because low efficiency is generally caused by poor primer/probe designs and assay conditions. These are often associated with poor assay performance and may not be suitable for long-term use for sample analysis.

Table 3. qPCR Assay Validation Considerations for Cell and Gene Therapy.

Parameter	Considerations
Standard Calibrators	• Ct (threshold cycle) values vs back calculated copy numbers • Avoid inhibitory factors (DNA quality and quantity)
Accuracy and Precision	• Define assay variability for both transgene and reference gene measurement • QCs: plasmid spiked into human gDNA for transgene; plasmid only for reference gene or spiked into surrogate nonhuman gDNA using the clinical material or construct • CAR-T from normal donors spiked into diseased whole blood: repeat testing to demonstrate intra- and inter-assay variability while accuracy can be inferred • Reference gene measurement from a qualified lot of human gDNA • Define an internal control for bioanalysis if not multiplexing • Three levels of QCs • Three replicates per sample: duplicate reactions plus one replicate spiked with qualified internal control (no need if multiplexing) or multiplexed into each sample
LLOQ and LOD	• Y intercept using a standard calibration curve (theoretical LOD) • ≤50 copies of transgene per μg gDNA (149,925 cells, 0.03%) while reference gene can be much higher (~300,000 copies/μg gDNA)
Specificity	Need to distinguish gene of interest sequence from other interfering endogenous sequence (check for cross reactivity)
Stability	Storage for each matrix and DNA using spike in controls and processing
Incurred Sample Reanalysis	Limited data is available for ISR to understand the suitable criteria
Amplification Efficiency	• $[10^{(-1/\text{Slope})} -1]$ • Efficiency ≈ 100% when slope ≈ -3.32 • Slope should be between -3.58 and -3.10 (corresponding to 90 and 110% efficiency)
Factorial Optimization	Examine the factors that may impact primer/probe performance across a variety of interactions: Factors: Forward Primer Concentration, Reverse Primer Concentration, Probe Annealing Temperature, Master Mix Response: slope, earliest Ct, highest fluorescence and guard banding

Acceptance Criteria

Acceptance criteria should be established before validation. Specific criteria should be evaluated case by case since specifics will differ with assay and COU. Laboratories should refer to the previously mentioned papers discussing qPCR validation. Most of assays may not meet criteria for accuracy and precision for small molecule assays with %Bias and %CV at 15% (LLOQ at 20%), respectively. Some assays can meet LBA criteria for precision and accuracy criteria with %Bias and %CV at 20% (LLOQ at 25%), respectively but this should be defined on an assay by assay basis and in terms of COU.

Primer/Probe Selection

The primer and probe selection are critical for assay performance [48]. No template controls (NTC) and baseline areas should be clear and clean with proper instrument calibration for the dye. The exponential phase needs to show a strong, straight growth; 100% efficient PCRs show 1 Ct difference between successive 2-fold dilutions with less than 10% CV difference between technical replicates. The plateau should be as close to horizontal as possible (e.g., within +/- 1 Ct difference) and individual replicates should plateau at the same fluorescence level. If the plateau is not horizontal, this indicates that the PCR is not efficient. This can happen if the denaturation midpoint (Tm) of primers in the reaction differs by more than 5°C, especially for primer/probes designed in GC (guanine-cytosine) rich genomic regions. Different concentration levels should plateau at the same level otherwise this indicates reduced assay sensitivity. It was also recommended to design and test multiple oligomer sets per target and perform basic local alignment search tool (BLAST) search to avoid primer dimers for specificity. In addition to a BLAST search, assessment of the primer probe set using gDNA from the same species of intended use to ensure no cross reactivity should also be done during development. Target isolation procedures should be tested for extraction efficiency and potential assay interference using A280/260 or other methods to quantify gDNA/cDNA to normalize results (or use reference targets). Critical reagents (primer, probe, master mix) need to be optimized for assay performance and stability. Each step of thermal cycling should also be optimized for an amplification efficiency above 90% using independently prepared replicate controls and samples. A factorial approach should be used to test factors that impact performance such as primer concentrations, probe concentrations, annealing temperatures, and master mix type. Slope, earliest Ct, and highest fluorescence responses should be assessed for optimization.

Calibrator Material and Assay Controls

Calibrator material should be equivalent or very similar to test samples. If possible, use clinical grade material although it is acknowledged that using clinical grade material may preclude some patients from being treated due to their limited supply. During the design phase, ss/ds (single stranded/double stranded) DNA template can be used (whole or partial). The extractable material for development should demonstrate extraction efficiency with encapsulated ss/ds DNA or cloned/synthesized DNA. If the assay is intended to be used for an extended period of time, it is preferred to use current Good Manufacturing Practices (cGMP) material, but this may not be feasible due to manufacturing constraints. Research lots may be used as long as they are within the same context of use, but bridging to a clinical lot may be required. Long term stability should be performed.

NGS Assay Development and Qualification

Gene editing is advancing rapidly toward clinical applications. For example, AAV-mediated *in vivo* delivery of gene editing reagents together with a transgene is being developed to introduce a functional gene into the albumin locus of patient hepatocytes. Transgene integration is achieved through engaging homology directed repair (HDR) as well as non-homologous end joining (NHEJ) DNA repair pathways. In addition, the NHEJ DNA repair process can also lead to the introduction of short insertions and deletions (indels) of nucleotides, without transgene integration. Measurement of indels may be used as a surrogate of genome editing efficiency. The evaluation of indels at previously identified off-target sites is, furthermore, essential for monitoring patient safety in clinical trials.

NGS is the technology of choice for the quantification of indels because of its high discovery power for heterogenous indel variants, the ability to multiplex samples and analytes, and the requirement for small DNA input material. For transgene integration, the HDR and NHEJ DNA repair mechanisms generate different DNA sequences at the target integration site, but they produce identical albumin-transgene fusion messenger RNA (mRNA) as a result of pre-mRNA splicing. Reverse transcription (RT) followed by qPCR is a suitable method to quantify transcripts with known sequences.

It was recommended for assay qualification to include precision, accuracy, and sensitivity; reference standards and quality controls were prepared by mixing genomic DNA extracted from a clonal liver cell line carrying 100% indels with wild type unedited genomic DNA at varying ratios. Although next generation sequencing computational analysis pipelines provides quantitative %indels data for patient samples, back calculating %indels against the standard curve makes the assay more robust. This can be especially important in situations where small biopsy samples yield limited amounts of genomic DNA that do not allow for assay failure and repeat sample testing. Discussions uncovered that most sponsors do not adjust %indel using a standard curve but it was agreed that %indel can be adjusted using a standard curve when only limited gDNA material is available.

Regulatory guidance for quantitative molecular assays to monitor editing in patients is lacking. Assay parameters to be evaluated for NGS and qPCR-based quantitative indel and integration assays during validation for intended purpose were discussed. It was recommended that qualification parameters to be evaluated for NGS-based clinical indel assays include gDNA input, specificity of PCR reactions, sequencing coverage, quantifiable range and LOD, precision and accuracy, selectivity, and gDNA stability. In addition, bridging and linearity should be considered when applicable.

It was recommended that in absence of a universal reference standard, a well characterized/qualified reference standard can be used; consideration should be taken for lot-to-lot bridging to ensure consistency. It was also recommended that reference standards and QCs for NGS based clinical indel assays should be generated with different levels of indels such as mixing cell lines of known indel levels with unedited parental cells, followed by gDNA extraction. If possible, the known indel levels would be best confirmed by using an orthogonal method. Stability of reference standards and QCs should be tested, and a qualification/bridging program should be established.

It was recommended that parameters to be evaluated for RT-qPCR gene integration assays include RNA input, specificity of RT-qPCR reactions, quantifiable range and LOD, precision and accuracy, selectivity, tissue storage and shipping conditions to ensure RNA integrity, and QC RNA stability. Reference standards and QCs for RT-qPCR gene integration assay can be generated and characterized by, for example, cloning the fusion cDNA and wild type DNA into a plasmid under the control of a T7 promoter. *In vitro* transcription can be used to synthesize RNA reference material. The use of A260 to calculate copy number based on size of transcript was recommended. A qualification/bridging program should be established for reference standards and QCs.

ddPCR Assay Development: Best Practices

The efficacy and safety evaluation of gene/cell therapies during development requires measuring target nucleotide sequence levels following their delivery/editing and characterizing their biological distribution and potential to release into the environment. qPCR has historically been the bioanalytical workhorse in studies of nucleic acid quantitation and characterization. More recently a newer technology, droplet digital PCR (ddPCR), has been added to the gene/cell therapy bioanalytical assay toolkit. ddPCR is a dPCR method based on water-oil emulsion partitions or droplets, Similarities between the two PCR methods include the common use of TaqMan chemistry comprised of target-specific primers and fluorescent probes to amplify and detect target sequences in samples. A key difference is the process by which source amplicons are quantitated. In real-time qPCR, the magnitude of fluorescence is measured continuously. The cycle at which the signal is detected above threshold (Ct) or reaches maximal increase in released fluorophore (Cp) is used to determine the concentration of the target sequence by interpolation from a standard calibrator curve. In contrast, dPCR involves the partitioning of target and background sequences to an estimated single copy prior to amplification. End-point reactions are analyzed in thousands of partitions for the presence (positive) and absence (negative) of fluorescence to determine the absolute number of target sequences, without the need for a standard curve and with increased tolerance to variable efficiency of amplification and matrix interference. ddPCR enables specific, accurate quantification of a vector/transgene construct. It readily enables multiplexing, which can be challenging to establish in a qPCR environment where amplicons compete for resources and efficiency close to 2 must be maintained for both amplicons. Multi-color ddPCR facilitates assays that are challenging on the qPCR platforms due to amplicon structure, such as quantitative measurement of AAV vector ITR fusions in cells/tissues over time. It also enables assays that are essentially impossible on the qPCR platform, such as determination of linkage between two different sites on a vector as a measure of vector integrity.

The utility of ddPCR to determine vector shedding and biodistribution as well as transgene biodistribution and expression was discussed. It was concluded that ddPCR can be used for both biodistribution and shedding studies but it was recommended that the same technology be used for both. With respect to both q and ddPCR, it was recommended that a yield threshold be determined for samples with low levels of total DNA (e.g., urine). There is a lack of guidance on collection of spot samples or collection of samples over a defined period of time. Hence, it was recommended for low DNA yield, samples can be stored in single-use aliquots.

Due to the newness of the technology, the current recommendations represent the best practices for managing ddPCR technical challenges. Moreover, another resource is the digital MIQE guidelines: Minimum Information for the Publication of Quantitative Digital PCR Experiments [49].

The 2020 recommendations describe the best practices for manually setting the positive/negative threshold when the analysis software is unable to. It was recommended to use positive controls as guides and/or evaluate 1 or 2D scatter plots and histograms for separation of single from double positives and positives from negatives and to draw appropriate thresholds. The event is documented, verified by a second scientist, and the original and manual raw data and analysis files are saved.

Frequently the negative control is zero (no positive droplets are detected) with multiple independent analyses. The best method to establish the LOD is to determine the lowest concentration at which 95% of the positive samples are detected. In other words, within a group of replicates containing the target at concentrations at the LOD, no more than 5% failed reactions should occur [35]. Alternatives to calculating the LOD include probability analysis or the maximum likelihood estimation method.

If the assay is quantitative with acceptable performance, less than 90% amplification efficiency due to amplicon structure between the primers (e.g., hairpin loops, high GC content), may be acceptable if the assay is FFP and with proper justification supported by other measures (accuracy, precision, linearity). ddPCR is less impacted by amplification efficiency – with reduced efficiency, the droplet is still positive though the signal may be reduced. As long as there is a clear separation of positive/negative, there is no impact to ddPCR quantification. With respect to primer/probe sets, design considerations are similar to those for qPCR and reaction optimization is recommended. Annealing temperature gradients can be used to optimize the separation of positives and negatives. Optimization may include confirmation of quantitative results with an independent nearby primer/probe set [49].

AAV Capsid NAb Assays Development & Validation

Development of GTx utilizing AAV vectors requires a number of unique approaches for bioanalytical and immunogenicity evaluation. Among the most important of these are assays to detect antibodies which may inhibit

the ability of the viral capsid to transfect target cells, also known as NAb or transduction inhibition (TI) assays. These cell-based assays are routinely used to measure the ability of patient derived antibodies to block cellular infection. The presence of NAbs has been associated with poor therapeutic efficacy, thus positivity in the assays has sometimes been used as exclusion criteria for clinical trials. Induction of NAb by the initial administration of AAV limits re-administration of the same serotype capsid. Exclusion of anti-AAV NAb seropositive subjects is common in clinical studies [50]. The use of anti-capsid NAb or TI assays and the methodology to identify parameters which may impact performance was discussed. In addition, the regulatory requirements for such testing and the need for developing CDx devices prior to marketing approval of the therapeutic product were discussed.

LBA approaches are used to detect total anti-AAV antibodies and cell-based assays are used for anti-AAV NAb assays. A bridging assay format detecting both non-neutralizing and neutralizing antibodies and using cut point as in other ADA assays gives better analytical sensitivity and specificity than cell-based assays. Cell-based NAb assays measure AAV capsid transduction efficiency in cultured cells using an AAV capsid engineered with reporter gene products. NAb reduces transduction efficiency of AAV capsids and reduces assay signals. There are serum components and compounds other than NAbs present in blood that can inhibit the transduction of cells by AAV capsids. Thus, it is more accurately called transduction inhibition assay although some factors may even enhance transduction. Often a 50% inhibition cut-off has been used to define sample positive/negative score and the reported titer. A statistically defined cut-point value is recommended going forward to align with the approach applied for protein based biotherapeutics. The titer used as threshold for NAb positivity is empirically determined and may differ from one product or assay to the next. Parameters to be optimized include: cell line (HEK 293, HeLa, CHO, Huh-7), cell number, serum (negative control, heat inactivation), multiplicity of infection (MOI) per cell (10^2 to 10^5), helper adenovirus or chemicals to enhance AAV transduction and expression, reporter gene (luciferase, GFP, LacZ), incubation time, and cut point assignment (50% inhibition or statistically derived cut point factor) [51].

The regulatory considerations for GTx NAb assays were discussed. During the clinical study, a qualified test for supportive studies (such as Phase I/II) and a validated test for pivotal studies (such as Phase III studies) can be used. During clinical development, these assays may be run in a single lab as LDTs; however, upon marketing of the therapeutic product it may be preferable to have an FDA-approved CDx available. In the EU, country specific regulations or recommendations may apply. Based on current GTx BLA submissions, Phase I/II studies were used as pivotal studies to support approval. Therefore, the fit-for-purpose concept applies to GTx NAb assays. It is recommended to consult regulatory agencies early on for guidance. In other countries, an ISO 15189 accreditation and a diagnostic permit is usually needed. Further discussion on this topic is recommended at next WRIB.

Another concern was the approach to determining exclusion criteria for clinical trials of AAV-based GTx, as discussed in draft guidance provided by the FDA [52]. Options include using anti-capsid antibody assays (ADA and/or NAb) or cellular immunity assays (such as ELISpot) for a capsid or transgene product. The utility of cellular immunity assays for capsid or transgene product as exclusion criteria has not been evaluated by regulators and the current consensus is that this is not required. These assays may be better used as an investigative tool or to guide preventive measures if a problem is suspected. Clinically relevant efficacy and safety issues should be thoroughly investigated, including the use of cellular immunity assays.

When considering the use of anti-capsid antibody assays (ADA and NAb), it was recommended to understand the impact of the presence of anti-capsid antibodies on the clinical outcome. Exclusion criteria for clinical trials of AAV-based GTx require scientific and clinical justification; anti-capsid antibody assays should not be universally used to exclude patients without knowing the impact on safety and efficacy. The precautionary principle is recommended for first-in-human (FIH) studies which may require excluding ADA/NAb positive subjects from AAV-based GTx trials. Where ADA/NAb positive subjects are excluded from pivotal studies, consideration should be given to performing separate studies in these subjects, which might also address the role of immunomodulation including Ab depletion. More data are needed before providing specific guidance. Drug developers should collect data on initial efficacy, duration of efficacy, early adverse events and long-term safety to inform future clinical trials.

ELISpot & Single Cell Western Blot Assay Validation

ELISpot Validation

Detection and monitoring of the immune response triggered by biologic drugs is of paramount importance for the safety and efficacy assessment of large molecule drugs. As the immune responses against these biologics include the

induction of CD8-positive cytotoxic T lymphocytes (CTLs), the analysis of activated T-cells is required. Activated CTLs will trigger the killing of the transduced cells, resulting in a decreased efficacy of the gene therapy and in putative organ damage.

ELISpot assays allow the detection of low frequency cells secreting various molecules such as cytokines. ELISpot has recently seen a resurgence and has become the technique of choice for the assessment of the cellular immunity triggered by gene therapies, via the assessment of the activation of cytotoxic T cells. While bioanalytical scientists have gathered extensive expertise with the cell culture, immunoassay and image analysis aspects of ELISpot, the quality of the cells initially used in the assay and their relevant stimulation is one of the key factors impacting the quality and the biological relevance of the results. It is indeed critical that peripheral blood mononuclear cells (PBMCs) are isolated in a timely and consistent manner from patients' whole blood, across multiple clinical sites and that cell stimulation with peptide pools and positive controls is carried out in a biologically relevant fashion. This need can pose logistical challenges for global studies and may require significant effort to ensure that PBMCs are isolated within a pre-defined time from whole blood collection at a clinical site.

Gene and cell therapy (GCT) programs have unique characteristics that differentiate them from vaccine programs, but the need to identify relevant T-cell immune responses is shared. The experience gleaned from the use of T-cell ELISpot in vaccine clinical trials provides many lessons learned that can be applied to GCT programs.

Within the vaccine industry, T-cell immune responses are induced in humans by vaccination to infectious diseases such as HIV, VZV, HPV and others. However, the determination of a correlate of protection has been elusive for clinical vaccine trials using the ELISpot assay. Several cancer immunotherapy modalities have shown a correlation of IFN-γ T-cell responses with survival however, the studies are very limited in subject number and positive clinical responses making interpretation difficult. Immunogenicity assays used in vaccine clinical trials are expected to be qualified or validated depending upon the phase of the clinical trial. It was agreed that certain parameters used to qualify or validate an antibody binding assay cannot readily be applied to an ELISpot assay. There are key approaches that can be used to optimize the performance of the ELISpot assay so that it can be qualified for clinical application. As an indicator of T-cell immunity, the ELISpot assay is sensitive, relatively robust and can be optimized to reduce variability. Vaccine clinical programs generally run over an extended period of time, resulting in the need for maintaining consistent assay performance. The ELISpot is no exception and brings its own challenges to assay maintenance.

White Papers are available to provide clarity on harmonization of practices and analysis of the quality of ELISpot method results [25,53]. The 2019 White Paper discussed challenges for ELISpot assays including a complex workflow from sample collection to testing, especially for larger multicenter studies.

The current recommendations on ELISpot validation build on the 2019 White Paper in Bioanalysis and ongoing industry experience comprised of performing ELISpot assays according to standard harmonized procedures including automated reading for spot counts [54]. It was recommended that laboratories performing ELISpot assays participate in external QA assessment, for example, T cell ELISpot Proficiency Panels [55].

Critical factors when developing the method are focused on the reagents. High quality frozen peripheral blood mononuclear cell (PBMC) preparations are critical (see below), as are the capture and detection antibody pair. The antigen source can be overlapping synthetic peptides or whole proteins, but purity is a key factor. Filter plates may be polyvinylidene difluoride (PVDF) or some other variety.

Before validation, techniques can be used to optimize the ELISpot process to reduce variability. This optimization includes trained staff, a standardized thaw process, resting of PBMC overnight, and using triplicate vs. duplicate wells. Multifactorial design of experiment (DoE) should be used to assess and optimize the robustness of the critical steps. Automated ELISpot readers are also important tools for optimization. The scan and count function for spot imaging and counting reduces processing time and integrates consistency. Key imager parameters include spot size and shape, which are important for differentiating background from true spots, and spot contrast/intensity. A consistent approach to spot counting through method development and optimization should be established. Definition of a counting template can be useful; counting parameters (e.g., min/max spot size, sensitivity) that have been defined using a set of relevant ELISpot well images would maintain consistency within a study. Maintaining consistency is important for large studies, although there remain some challenges around the correlation of ELISpot positive signals and serious adverse events in patients.

The traditional validation parameters for LBA and vaccine immunogenicity assays can be adapted for ELISpot assay qualification using a sample panel (Table 4) and assay qualification parameters (Table 5).

Table 4. Sample Panel for ELISpot Assay Qualification.

Sample Use	Total No. Samples	Description
Assay Control	1 to 2 Positive 1 Negative	PBMC pre-screened to ensure reactivity to relevant antigen and/or mitogen
LOD/LLOQ, Ruggedness, Precision	Minimum 10 donors	PBMC from normal human donors tested over a period of 2-3 weeks by multiple analysts
Linearity/Relative Accuracy	6 samples tested at different cell inputs	6 samples selected from Panel R based on sample volume

Table 5. ELISpot Assay Qualification Parameters.

Parameter	Total No. Samples	Description	Expectations
Control Sample	1-2 Positive	The acceptable range of responses for the control will be established	N/A
Limits of Detection	1	Assess the instrument's upper limit of detection and the media only wells will be used to assess the assay's lower limit of detection	The assay upper limit of detection (ULOD) is expected be ≤450 spot forming cells (SFC) per well.
Limits of Quantitation	N/A	Establish LLOQ. LLOQ specifies the lowest value that can be quantified with acceptable precision (i.e., intermediate precision %CV ≤40%)	LLOQ ≤ LLOD (lower limit of detection)
Precision Ruggedness	≤10	Provide estimates of the intra-run-, inter-run, and total assay precision. Assess the ruggedness of the assay to factors that will vary over time during routine operation (analyst).	Intermediate %CV estimate must be ≤40% for ≥80% of the samples having mean spot forming units (SFU)/106 PBMCs > LLOQ. The maximum fold difference between assay ruggedness factor levels is expected to be less than 2-fold.
Linearity/Relative Accuracy	6 (each tested at 3 cell inputs)	Assess linearity/relative accuracy.	The sample reactivity is expected to decrease as cell input decreases.
Specificity	N/A	Ability to detect a positive reaction to a specific stimulation and to differentiate this reaction from background or negative response	Determined with a negative cut-off determination for vector and transgene peptide pools using a correction factor specific to each peptide pool.

Assay Suitability Criteria

The assay suitability criteria are derived from the assay qualification and should include QC sample ranges. The negative control has no antigen, mock, and determines the background. The positive control is a mitogen, peptide or responder PBMC. The replicate variability criteria should be <2-fold, and well acceptance criteria should be established.

Defining a Positive Response

Recommendations were provided to define a positive ELISpot response. The critical parameters to assess are ELISpot responses from an existing positive responder population (if available) and a large data set from non-responders. A statistical approach with dual criteria or mean + 2 standard deviations of pre-existing immunity/background is one option that is reasonable. In the absence of an established, clinically relevant positivity criterion for reactivity to peptide antigens, the positivity criteria for a sample must be established based on the level of reactivity that is meaningful above the background reactivity. For example, positive samples are defined by a positive read out 3 times the background when assay variability is twice the background. A different cut-off may be used based on the nature of the assay.

Assay Maintenance

Recommendations were given on how to maintain ELISpot assay performance. A proficiency panel of PBMC samples with expected ELISpot responses (optimally to the antigen of interest) should be established. New analysts should be qualified on the assay and their proficiency should be maintained. Side by side assessment of new reagent lots against qualified lots is necessary for reagent bridging and trending. Both real-time trending of controls during active testing and proficiency panel which monitors assay performance over time were recommended. In the EU, assays need to be confirmed every year. In other countries, the quality control unit of the laboratory will ensure compliance.

Robustness/Ruggedness

Different parameters that can vary over time during routine operation should be tested. The first tier of tests should include detection antibody incubation time, dilutions of conjugated antibodies, and culture period of cells. The second tier of tests can include kit lots (n≥2), and operators (n≥3). Validation experiments should be conducted depending upon COU.

Peripheral Blood Mononuclear Cells

The most critical factor for ELISpot assay performance is PBMC sample quality and availability. It is recommended to use multiple donors (relevant population), and a minimum of 10 vials per donor (1×10^7 PBMC per vial) as part of assay development and validation. Preparation protocols from blood are based on differences in blood cell density: blood cells with round nuclei are lymphocytes (T cells, B cells, natural killer cells) which are the majority of PBMC population, monocytes, and dendritic cells (small percentage). Erythrocytes and platelets contain no nuclei and granulocytes (neutrophils, basophils, and eosinophils) contain multi-lobed nuclei.

It is important to plan PBMC collection in collaboration with the clinical team to mitigate inter-site differences in sample collection and handling. Blood should be collected at clinical sites according to optimized protocol. Central labs should be provided the optimized protocol for isolation, counting, aliquoting, and cryo-preservation. A mock PBMC isolation from blood should be performed to mimic blood storage and shipment delay. After thawing, the bioanalytical lab should assess cell viability (% recovery) and cytokine synthesis. Viability <80% indicates a need to troubleshoot and optimize protocols. If viability ≥80%, cytokine production can be tested.

Prolonged storage of blood (>3h) results in accumulation of granulocytes in the lymphocyte ring of lymphoprep isolation. Increased amounts of granulocyte result in a dramatic dilution of target cells (T cells) causing a high risk to generate false negative ELISpot results. RosetteSep™ granulocyte depletion removes accumulating granulocytes from PBMC preparation and restores T cell frequency.

Cell numbers should be normalized, i.e., PBMCs can be analyzed by flow cytometry and then normalized to the fixed/pre-defined number of T cells (executed at the site of the ELISpot testing). It is acknowledged that limited sample volume, for example, in pediatric samples, might be a limiting factor.

Single Cell Western Blot

Advances in gene therapy are providing benefits to patients by addressing needs that have not been met by traditional biotherapeutics. Expression of the transgene products will be an important pharmacodynamic marker, and new techniques and platforms may need to be employed to measure this analyte in a variety of matrices. For example, a multiplexed western blot assay can measure expression of a transgene product as well as the diseased state biomarker in single red blood cells. The diseased state biomarker has been shown to cause unhealthy cellular morphology in subjects, while the transgene product has been shown to improve cellular morphology. By applying this assay to measure both proteins at the single cell level, a better understanding of the pan-cellularity of the transgene is gained.

Design of experiment studies used to characterize assay variability have been meaningful to establish validation acceptance criteria for this novel approach. In assay development, electrophoresis, cell occupancy, lysis time, antibody selection and concentration, and fluorescence detection instrument settings need to be considered. Electrophoresis optimization will provide better resolution of the analyte, making data analysis easier to automate. Not all fluorophores are created equal – some analytes will be better resolved due to channel and fluorophore selection. In the absence of good control antibody and marker, maintaining consistent PMT voltage becomes critical for tracking assay performance. Analyst technique can impact the number of occupied wells.

The fit-for-purpose validation acceptance criteria for single cell western assays were discussed and whether LBA criteria (20–20%) are sufficient for use. Consensus was reached that FFP validation criteria should be based on context of use and established prior to performing the experiments. Assays should be validated if the data will be used as a crucial parameter for patient safety and efficacy of the treatment. In this case, validation studies should be based on ICH Q2R1 guidance [56]. When only used as supporting data, assay qualification is sufficient.

Application of Current FDA/EMA Immunogenicity Guidance/Guideline to Gene Therapy:

Gene and cell therapies present unique challenges in the assessment of immunogenicity. The potential for host immune responses to impact the efficacy and durability of these therapeutics has led to an emphasis on characterizing the immune response more thoroughly than typically performed for other biotherapeutics. Many of the current principles that are captured in the FDA [26] and EMA [57] guidance documents are applicable to immunogenicity

risk assessment and assay development and validation for GTx. However, there are areas where the guidelines do not specifically address the unique concerns surrounding this new area of drug development. Specifically, there are concerns about pre-existing immunity and its impact on transduction of gene therapies, a greater emphasis on cell-based assays, and immunogenicity against transgenes versus the vector.

Unlike the "typical" biotherapeutic which is one molecule, gene and cell therapies have the potential to deliver many foreign proteins into the patient. It is important to consider all foreign proteins that may be introduced/produced. This includes anti-capsid antibodies, anti-transgene antibodies, anti-chimeric antigen receptor (CAR) antibodies, and anti-CAS9 antibodies, etc. Performing an early risk assessment is essential if only to get started on the different assays that may be needed. A typical screen, confirm, titer approach for testing can be used. Screening and NAb assays are likely to have significant numbers of baseline positives which will require far more individuals than typical for cut point setting. It may be difficult in rare disease or pediatric populations to obtain sufficient number of individuals and sample volumes. Significant geographic differences have also been observed [58–60]. Cell based assays should be implemented earlier than typically seen for biotherapeutics including transduction-based neutralization measures and determination of cellular immunity to AAV as inclusion/exclusion criteria.

It was agreed that the current immunogenicity guidance documents for biotherapeutics [26,57] are broadly applicable to gene/cell therapy [51], even though the FDA guidance specifically states that it does not apply to GTx. Although neither the FDA nor EMA have issued detailed guidance on the immunogenicity assessment of gene/cell therapy, consideration should be given to developing immunogenicity assays for, for example, anti-vector, anti-transgene product, etc. with timely measurements. Risks associated with gene therapies are not clear, so a product specific risk assessment should be considered.

Consensus was reached to base immunogenicity assessments on several general principles. First, initiate an immunogenicity risk assessment and update as needed during development along with justification of the details of the assessment. Given significant uncertainty, store samples in order to perform additional analyses later on, if needed. Ideally, execute the immunogenicity assay at a single, central laboratory. Investigate the impact of immunogenicity on clinically relevant issues such as lack of efficacy, loss of efficacy, and adverse events.

Current data do not support a general recommendation on what pre-existing titers are relevant in terms of clinical safety and efficacy. For protein-based biologics, it has been proposed that pre-existing antibodies may, but do not necessarily lead to, boosted immune responses. A clinically relevant cut point needs to be determined, anticipating that high titers may have a higher clinical relevance. Other factors like serotype, dose, product type, intended target organ, etc. may impact inclusion criteria. Due to challenges in comparing assays across sponsors, it is impossible to select a titer for exclusion/inclusion for all serotypes.

RECOMMENDATIONS

Below is a summary of the recommendations made during the 14th WRIB:

qPCR, ddPCR, and NGS Assay Development and Validation: Best Practices

qPCR Validation

- qPCR assays should be qualified and validated by the sponsor, with a clear understanding of the assay sensitivity, reproducibility and variability in each matrix to be tested. Parameters to consider include: standard calibrators, accuracy, precision, LLOQ and LOD, stability, amplification efficiency and factorial optimization. Refer to Table 3
- **Acceptance criteria** should be established before validation. Specific criteria should be evaluated case base case since specifics will differ with assay and COU
- **Primer/probe selection:**
    - NTCs and baseline areas should be clear and clean with proper instrument calibration for the dye
    - The exponential phase needs to show a strong, straight growth (target 1 Ct difference between successive 2-fold dilutions)
    - The plateau should be as horizontal as possible and individual replicates should plateau at the same fluorescence level
    - Different concentration levels should plateau at the same level otherwise this indicates reduced assay sensitivity

- Design and test multiple oligomer sets per target and perform BLAST search to avoid primer dimers and for specificity
- Target isolation procedures should be tested for extraction efficiency and potential assay interference using A280/260 or other methods to quantify gDNA/cDNA to normalize results (or use reference targets)

- **Calibrator material** should be equivalent or very similar to test samples. If possible, use the clinical material
- For long term calibrators, it is preferred to use cGMP material but research lots may be used with bridging
- **Extraction efficiency** can be evaluated with encapsulated ss/ds DNA or cloned/synthesized DNA

NGS indel assays and RT-qPCR integration assays

- Measurement of indels may be used as a surrogate of **genome editing efficiency**
- Evaluation of indels at previously identified off-target sites is essential for **monitoring patient safety** in clinical trials
- NGS is the technology of choice for the **quantification of indels** because of its high discovery power for heterogenous indel variants, the ability to multiplex samples and analytes, and the requirement for small DNA input material
- Reverse transcription followed by qPCR is a suitable method to quantify transcripts with known sequences
- In absence of a universal reference standard, a well characterized/qualified reference standard can be used; consideration should be taken for lot-to-lot bridging to ensure consistency
- **Reference standards and QCs for NGS based clinical indel assays**

  - Should be generated with different levels of indels such as mixing cell lines of known indel levels with unedited parental cells, followed by gDNA extraction
  - If possible, the known indel levels would be best confirmed by using an orthogonal method
  - Stability should be tested
  - Qualification/bridging program should be established

- **Parameters to be evaluated for NGS based clinical indel assays** include:

  - gDNA input
  - Specificity of PCR reactions
  - Sequencing coverage
  - Quantifiable range, linearity and LOD
  - Precision and accuracy
  - Selectivity
  - gDNA stability
  - Bridging and linearity should be considered when applicable

- **Reference standards and QCs for RT-qPCR integration assays**

  - Clone the fusion cDNA and wild-type DNA into plasmids under the control of a T7 promoter
  - *In vitro* transcription can be used to synthesize RNA reference material
  - Copy number of purified RNA can be calculated using A260 and size of transcript

- **Parameters to be evaluated for RT-qPCR gene integration assays include:**

  - RNA input
  - Specificity of RT-qPCR reactions
  - Quantifiable range
  - Precision and accuracy
  - Sensitivity and LOD
  - Selectivity
  - Tissue storage and shipping conditions to ensure RNA integrity
  - QC RNA stability

- Qualification/bridging program should be established for reference standards and QCs

- Back calculating %indel against a standard curve increases assay robustness when only limited gDNA material is available

ddPCR Assay Development: Best Practices

- ddPCR enables specific, accurate quantification of a transgene construct and readily enables multiplexing
- ddPCR can be used for both biodistribution and shedding studies but it was recommended that the same technology be used for both
- For low DNA yield, samples can be stored in single-use aliquots
- Best practices for manually setting the positive/negative threshold when the analysis software is unable to:
  - Use positive controls as guide
  - Evaluate 1 or 2D plots and histograms for separation of positive and negative (single/double positives)
  - Draw threshold
  - Document event and second scientist confirms
  - Save original and manual raw data and analysis files
- Establish the LOD by determining the lowest concentration at which 95% of the positive samples are detected. Within a group of replicates containing the target at concentrations at the LOD, no more than 5% failed reactions should occur
- If the assay is quantitative with acceptable performance, less than 90% amplification efficiency due to amplicon structure between the primers (e.g., hairpin loops, high GC content), may be acceptable if the assay is FFP and with proper justification supported by other measures (accuracy, precision, linearity)

AAV Capsid NAb Assays: Development & Validation

- **LBA approaches** are used to detect total anti-AAV antibodies and cell-based assays are used for anti-AAV NAb assays
- **A bridging assay** format detects both non-neutralizing and neutralizing antibodies and uses cut point as in other ADA assays giving better assay sensitivity and specificity than cell-based assays
- **Parameters** to be optimized include:
  - Cell line (HEK 293, HeLa, CHO, Huh-7)
  - Cell number
  - Serum (negative control, heat inactivation)
  - MOI per cell (10^2 to 10^5)
  - Helper adenovirus or chemicals to enhance AAV transduction and expression
  - Reporter gene (luciferase, GFP, LacZ)
  - Incubation time
  - Cut point assignment (50% inhibition or statistically derived cut point factor)
- During the clinical study, a qualified test for supportive studies (such as Phase I/II) and a validated test for **pivotal studies** (such as Phase III studies) can be used
- For commercial products, an ADA/NAb assay may be considered a **CDx** and should be performed in CLIA-certified lab in US. In the EU, country specific regulations or recommendations may apply. In other countries, an ISO 15189 accreditation and a diagnostic permit is usually needed
- When considering the use of **anti-capsid antibody assays** (ADA/NAb)
  - Understand the impact of the presence of anti-capsid antibodies on the clinical outcome
  - Exclusion criteria require scientific and clinical justification
  - Anti-capsid antibody assays should not be universally used to exclude patients without knowing the impact on safety and efficacy
  - The precautionary principle is recommended for FIH studies which may require excluding ADA/NAb positive subjects from AAV-based GTx trials

- Where ADA/NAb positive subjects are excluded from pivotal studies, consideration should be given to performing separate studies in these subjects, which might also address the role of immunomodulation including corticosteroids and Ab depletion

- The utility of **cellular immunity assays for capsid or transgene product** as exclusion criteria has not been evaluated by regulators. Consensus from industry is that that this is not required. These assays may be better used as an investigative tool or to guide preventive measures if a problem is suspected

ELISpot & Single Cell Western Blot Assay validation

ELISpot Validation

- Multifactorial DoE can be used to assess and optimize the robustness of the critical steps
- **Reagents:**
  - High quality frozen PBMC preparations are critical (see below)
  - Capture and detection antibody pair
  - Antigen source can be overlapping synthetic peptides or whole proteins
  - Filter plates may be PVDF or other variety
- Automated ELISpot readers:
  - Scan and count function for spot imaging and counting reduces processing time and integrates consistency
  - Key imager parameters include spot size and shape, and spot contrast/intensity
- A consistent approach to spot counting through method development and optimization should be established. Definition of a counting template can be useful
- Validation experiments should be conducted depending upon COU
- Laboratories performing ELISpot assays participate in external QA assessment
- Assay parameters for an **ELISpot validation**:
  - Precision, assay range (LLOQs), ruggedness, linearity and specificity
  - Use high purity grade peptides
  - PBMC preparation harmonization between sites is recommended
  - Do not to use cells with a viability <80%
  - Cells can be normalized, i.e., PBMCs need to be analyzed by flow cytometry and then the PBMC should be adjusted to the fixed/pre-defined number of T-cells (executed at the site of the ELISpot analysis)
  - Different parameters that can vary over time during routine operation should be tested to determine ruggedness
- Assay suitability criteria are derived from the assay qualification and should include QC sample ranges. The replicate variability criteria should be <2-fold, and well acceptance criteria should be established
- **Positive Response:**
  - Critical parameters are the existing positive responder population (if available) and a large data set from non-responders
  - A statistical approach, for example, with dual criteria or mean + 2 standard deviations of pre-existing immunity/background can be used
  - The positivity criteria for a sample must be established based on the level of reactivity that is meaningful above the background reactivity
- for example, a positive read out equals 3 times the background when assay variability is twice the background. A different cut off may be used based on the nature of the assay
- **Assay Maintenance:**

- A proficiency panel of PBMC samples with expected ELISpot responses (optimally to the antigen of interest) should be established
- New analysts should be qualified, and their proficiency maintained
- Side by side assessment of new reagent lots against qualified lots is necessary for reagent bridging and trending
- Real-time trending of controls during active testing and proficiency panel, which monitors assay performance over time, were recommended for trending
- In the EU, assays need to be confirmed every year. In other countries, the quality control unit of the laboratory ensures compliance

- **PBMC:**
  - Use multiple donors (relevant population), and a minimum of 10 vials per donor (1×10^7 PBMC per vial)
  - Blood should be collected at clinical sites according to optimized protocol. Central labs should be provided the optimized protocol for isolation, counting, aliquoting, and cryo-preservation
  - After thawing, the bioanalytical lab should assess cell viability (% recovery) and cytokine synthesis
  - RosetteSep™ granulocyte depletion removes accumulating granulocytes from PBMC preparation and restores target cell (T cell) frequency
  - PBMCs can be analyzed by flow cytometry and then the PBMC adjusted to the fixed/pre-defined number of T cells

Single Cell Western Blot

- Design of experiment studies used to characterize assay variability have been meaningful to establish validation acceptance criteria for this novel approach
- In assay development, electrophoresis, cell occupancy, lysis time, antibody selection and concentration, and fluorescence detection instrument settings need to be considered.
- In the absence of good control antibody and marker, maintaining consistent PMT voltage becomes critical for tracking assay performance
- FFP validation criteria should be based on context of use and established prior to performing the experiments
- Assays should be validated if the data will be used as a crucial parameter for patient safety and efficacy of the treatment. In this case, validation studies should be based on ICH Q2R2 guidance
- When only used as supporting data, assay qualification is sufficient

Application of Current FDA/EMA Immunogenicity Guidance/Guideline to Gene Therapy

- Current Immunogenicity Guidance/Guideline (FDA/EMA) for biotherapeutics [26,57] are broadly applicable to gene/cell therapy
- All foreign proteins that may be introduced/produced by the patient should be considered. This includes anti-capsid antibodies, anti-transgene antibodies, and anti-CAS9 antibodies
- Perform an early **risk assessment**
- A typical screen, confirm, titer approach for testing can be used
- **Screening and NAb assays** are likely to have significant numbers of baseline positives which will require far more individuals than typical for cut point setting
- **Cell based assays** should be implemented earlier than typically seen for biotherapeutics including transduction-based neutralization measures and determination of cellular immunity to AAV as inclusion/exclusion criteria
- Consideration should be given to developing **immunogenicity assays** for example, anti-vector, anti-transgene product, etc. with timely measurements
- Risks associated with gene therapies are unclear, so a product-specific risk assessment should be considered
- Base immunogenicity assessments on several general principles:
  - Initiate an immunogenicity risk assessment and update as needed during development
  - Justify the details of the risk assessment
  - Store samples in order to perform additional analyses later on, if needed

- Ideally, execute the immunogenicity assay at a single, central laboratory
- Investigate the impact of immunogenicity on clinically relevant issues such as lack of efficacy, loss of efficacy, and adverse events

- A clinically relevant cut point needs to be determined, expecting that high titers may have a higher clinical relevance

SECTION 3 – NAb Assay Harmonization, Biosimilars and FDA/EMA Guidance/Guideline

Haoheng Yan[4], Bonnie Wu[2], Daniela Verthelyi[4], Susan Kirshner[4], Joao Pedras-Vasconcelos[4], Manoj Rajadhyaksha[24], Roland F. Staack[25], Elana Cherry[12], Isabelle Cludts[26], Madeleine Dahlbäck[27], George R. Gunn[28], Akiko Ishii-Watabe[29], Vibha Jawa[30], Robert Kubiak[31], Michael Partridge[24], Marco Petrillo[32], Samuel O. Pine[33], Johann Poetzl[34], Sam Song[35], Chris Stebbins[36], Yuling Wu[31], Lucia Zhang[12]

Authors in Section 3 are presented in alphabetical order of their last name, with the exception of the first seven authors who were major contributors.

The affiliations can be found at the beginning of the article

DISCUSSION TOPICS & CONSOLIDATED QUESTIONS COLLECTED FROM THE GLOBAL BIOANALYTICAL COMMUNITY

The topics detailed below were considered as the most relevant 'hot topics' based on feedback collected from the 13th WRIB attendees. They were reviewed and consolidated by globally recognized opinion leaders before being submitted for discussion during the 14th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Cell-Based NAb Assays – Sensitivity and Drug Tolerance and the Relevance for Clinical Outcome

When do cell-based NAb assays add value if ADAs do not have an apparent effect on clinical outcome? When do they add value to measure *in vitro* neutralization in samples during treatment if the drug is not of high immunogenicity risk and NAb results may not be needed during treatment to ensure patient safety? Would regulators expect two NAb assay platforms (one NAb assay during treatment, and one NAb assay for follow up samples) when high drug trough levels are impacting assay sensitivity? What are regulatory and industry opinions on biotherapeutic mechanism of action (MoA) as the primary driving consideration when deciding if a functional cell-based NAb assay is needed?

NAb Assay Harmonization: Recent Trends and Expectations

Are NAb assays required for low immunogenic compounds known to only elicit very low titer ADA responses? Do we need to test NAbs for baseline ADA positive samples, placebo ADA positive samples, or samples with low risk molecules which were demonstrated to contain low-level ADA incidence in Phase I/II studies? What additional value do NAbs provide, given the availability of ADA status, titer, transient/persistent, PK, pharmacodynamics (PD), and efficacy? For bi- or multi-specific biotherapeutics, when should the NAb assay be implemented and what are the viable assay platforms? What is the risk level of immunogenicity for the T-cell redirected cancer cell killing biotherapeutics? When should NAb assay be implemented for this category of drug products?

Biosimilar Immunogenicity: Current Industry Standards

Are cell-based functional assays sensitive and drug tolerant enough to detect potential differences in the immunogenic profile of a proposed biosimilar compared to its reference? Regarding anti-IgG Fc response, how likely is the development of an ADA response towards this highly conserved region of antibody drug molecules? With 8 mg/mL IgG present in blood circulation, would assays be able to detect anti-IgG Fc ADAs specific for the drug? Should the interpretation of immunogenicity results in light of the assay used be included when presenting or publishing immunogenicity data?

The 2019 US FDA Immunogenicity Guidance: Reflections a Year Later

A relatively high false positive rate (FPR) of 5% in the ADA screening assays is expected to ensure high sensitivity. According to industry experience, assay sensitivity is dictated by the choice of analytical platform and not by artificially inflating the FPR. Is there a scientific reason to continue using a 5% FPR instead of a lower value

(e.g., 1% or 0.1%)? The 2019 FDA guidance [26] requires 1% of false positives in the ADA confirmatory assays while the 2017 EMA guidance [57] states that ADA confirmatory assays should eliminate any false positives. Do 1% and 0% FPR, respectively, refer to the confirmatory assay only or to the final data set? There are disadvantages to using box-plots with 1.5- or 3-inter-quartile range (IQR) to identify outliers in cut point data. What alternative approaches could be used to ensure conservative cut point values without unnecessarily excluding ADA-negative samples and preserving normal biological variability within the subject population? What comprises a good risk-based strategy that can be written as part of the Investigational new drug (IND)? What would be needed to perform an ongoing risk assessment that could then be summarized as part of Integrated Summary of Immunogenicity (ISI) for the BLA submissions?

DISCUSSIONS, CONSENSUS AND CONCLUSIONS

Immunogenicity assessments are an essential element of safety and efficacy evaluations in the product development of biopharmaceuticals. NAb assays characterize detected ADAs for their neutralizing activities as a component of an immunogenicity testing strategy. Analysis of ADA and NAbs can contribute to the understanding of drug PK, PD, efficacy, and safety observed during product development and patient treatment.

Cell-Based NAb Assays – Sensitivity and Drug Tolerance and the Relevance for Clinical Outcome

Assays used during clinical development to detect NAbs typically fall into one of two categories, either cell-based assays (CBA) or competitive ligand binding assays (CLBA). Assay formats are typically selected in alignment with drug MoA, assay performance, and the immunogenicity risk assessment; cell-based assay formats being generally preferred by most regulatory agencies for complex MoAs. Ideally, cell-based bioassays provide experimental systems that reflect all or a key portion of the pharmacologic activity of a biotherapeutic. Cell-based NAb assays often provide more appropriate platforms for assessment of ADA-mediated neutralization for agonists and many biotherapeutics targeting cell-surface receptors. However, they are generally more complex and challenging to develop. CLBAs may be more robust, sensitive and drug tolerant when used to support antagonistic biotherapeutics or receptor-Fc fusion proteins. The sensitivity of a CLBA typically reaches the 100–200 ng/mL range for the positive control detection, while CBAs may be less sensitive. In addition, CBAs can be adversely impacted by sample pre-treatment procedures such as those aiming to improve drug tolerance. Consequently, CBAs can have poorer drug tolerance than CLBAs [61,62].

EMA and FDA immunogenicity guidelines [26,57] present the current regulatory views on NAb assay format. The EMA recommends developing a NAb assay which responds well to the biotherapeutic and is sufficiently drug tolerant. The risk of NAb incidence and potential clinical consequences should also be considered. The EMA is open to the use of CLBA for therapeutics that exert their activity only through direct binding to their target (e.g., antagonistic mAbs), as long as the assay "reflects neutralizing capacity/potential in an appropriate manner" [57]. The FDA guideline indicates a preference for CBAs; however, there are bioanalytical considerations such as assay sensitivity, drug tolerance and selectivity that can lead to the use of CLBAs.

USP and industry consensus of opinion is that the biotherapeutic MoA is a primary consideration when deciding whether a functional cell-based NAb assay is to be used. According to Wu, et al., [63] and USP-NF <1106.1> [64], the CBA NAb assay format is preferred for agonist and antibody-drug conjugates (ADC). For an agonist or enzyme-replacement therapy, a CBA or an enzyme activity assay can be used; USP also allows for a CLBA. For an antagonistic MoA, there is disagreement on the preferred approach. Wu et al stated that CLBA is generally preferred, but CBAs can be useful for some biotherapeutics targeting cellular receptors, while USP advocates the use of CBA for soluble receptor biotherapeutics and for antibody-dependent cell-mediated cytotoxicity (ADCC)+ modalities.

Recommendations on the timing of NAb assays in clinical programs were provided in the 2013 White Paper in Bioanalysis [6], which advised taking a risk-based approach to decide when a NAb assay should be implemented, a recommendation that is still supported. In 2017, this recommendation was expanded to state that NAb assays should be included at all clinical phases for high risk molecules [18]. This year, the discussion focused on products of low immunogenicity risk and the possibility of using alternative measures such as PD markers to assess neutralizing activity of samples during treatment when drug interference presents a problem. Indeed, a PD marker or biomarker that correlates with clinical outcome could be part of an integrated approach that uses PK, ADA, and PD data to assess NAb activity over time during treatment. The success of an integrated approach would rely on the data quality from PK, ADA, and PD assays. Similar positions have previously been recommended [15,18,25].

For products where NAb assessment needs to be conducted during treatment but drug interferes with the sensitivity, it could be possible to use two different kinds of NAB assays, although this is not expected by the regulators: one NAb assay with a higher drug tolerance but lower sensitivity during treatment and one NAb assay with lower drug tolerance but higher sensitivity for the follow-up samples in the wash-out period. However, the two methods would need to be bridged for use in the same study and the rationale for using two assay platforms should be clearly described. As it is difficult to compare results from different NAb assays, the use of two methods in the same study is not common and would need to be discussed with regulators prior to implementation.

Since NAb positive controls are only surrogates and do not represent NAbs in clinical samples, a NAb assay with drug tolerance lower than the trough concentration may still detect NAbs in clinical samples. From an industry perspective, sponsors could use the actual drug concentration in NAb positive samples to reassess drug tolerance to help justify the adequacy of the assay; these results can be provided as supplemental data to support the validation report. Regardless of the assay format, NAb assays should be developed to characterize the neutralizing capacity of ADA in the samples collected, being mindful of potential drug and matrix interference. Hence, a certain degree of pragmatism is appropriate when selecting the NAb assay format.

NAb Assay Harmonization: Recent Trends and Expectations

Nab activity assessment for low Immunogenicity risk molecules

Regulatory agencies recommend immunogenicity assessment be conducted for all classes of biotherapeutics. NAb assays would characterize the neutralizing capacity of ADAs and may correlate with clinical outcomes. Some products can be considered to be of low immunogenicity risk, if they are not endogenous mimics, are administered to immunosuppressed patients, or have been previously shown to have low immunogenicity risk. Therefore, it was discussed whether it was possible to reduce the scope of NAb testing for low risk drug molecules. It is important to point out that a low risk molecule could attain a higher risk profile when patient population changes or the drug has a critical quality attribute that could make it more immunogenic. The industry perspective regarding the need for NAb testing for low risk molecules was varied. While some considered that testing for NAb places an unnecessary burden on drug development of low risk biotherapeutics, others voiced that the immunogenicity risk, Ab titers, and the potential clinical consequence of developing NAbs inform the immunogenicity risk; currently they are not sufficient to predict neutralizing activity and thus assessment of NAb activity is needed. Regarding the timing of NAb assay implementation, the need for development, validation and implementation of ADA and NAb assays early in product development is dictated by the risk of the product. To avoid holding up the drug development program, for low-risk products, NAb assays can be developed in parallel with early studies and implemented based on the evolving incidence and titer of ADA.

NAb assessment for multi-domain biotherapeutics

Bispecific and multi-domain biotherapeutics could elicit immune responses with multiple specificities to different domains. Risk factors derived from the drug product of such molecules could include novel amino acid sequences or epitopes in each functional domain, neo-antigens in the linker sequences or in the interface between domains, and post-translational modifications, for example, glycosylation, oxidation, deamidation. This increased complexity presents unique challenges to method development and characterization for immunogenicity. As with other therapeutics, the timing for implementing a NAb assay would depend on the risk of the molecule. The NAb assay should capture all the activities of the product which might be accomplished by single or multiple assays.

Biosimilar Immunogenicity: Current Industry Standards

A biosimilar is a biological product that is highly similar to and has no clinically meaningful differences with an approved reference product. Immunogenicity data for the proposed biosimilar product and the reference product is essential to support a demonstration of biosimilarity or interchangeability for a proposed biosimilar product.

Regulatory feedback on the bioanalytical strategy and method requirements to evaluate immunogenicity of biosimilars has evolved in the last few years, and the current thoughts are described in FDA guidance to industry [65,66] and method validation guidelines [26,57,67].

Clinical Immunogenicity Considerations for Biosimilar and Interchangeable Insulin Products

Most recently in 2019, FDA proposed an exception for the insulin program, which is discussed in the draft FDA guidance for industry "Clinical Immunogenicity Considerations for Biosimilar and Interchangeable Insulin

Products" [68]. This guidance states that "... generally, if a comparative analytical assessment based on state-of-the-art technology supports a demonstration of high similarity for a proposed biosimilar or interchangeable insulin product, there would be little or no residual uncertainty regarding immunogenicity; in such instances, a comparative clinical immunogenicity study generally would be unnecessary to support a demonstration of biosimilarity or interchangeability". The recommendation for insulin products is based on a number of factors, including "the relatively small, structurally uncomplicated and well-characterized nature of insulin products"; "extensive experience and literature survey that confirm minimal or no clinical relevance of immunogenicity with insulin product use"; and "scientific thinking on the lack of clinical impact of immunogenicity with insulin product use", as reflected from multiple sources detailed in the guidance. In conclusion, "current analytical tools used to evaluate quality attributes for insulin products can support a comprehensive analytical comparison thorough enough to support a conclusion that a particular proposed biosimilar insulin product that is "highly similar" to its reference product generally would have little or no residual uncertainty regarding immunogenicity and would be expected, like the reference product, to have minimal or no risk of clinical impact from immunogenicity. In such cases, a comparative clinical immunogenicity study would generally not be necessary to support licensure of a proposed biosimilar or interchangeable product [68]".

In practice, FDA recommends that a 351(k) BLA for a biosimilar or interchangeable insulin product covered by the guidance contains, among other things: adequate Chemistry, Manufacturing, and Controls (CMC) information, a comprehensive and robust comparative analytical assessment, a comparative clinical pharmacology study, and an immunogenicity assessment justifying why a comparative immunogenicity clinical study is not necessary. A comparative clinical immunogenicity study may be needed in some circumstances where there is residual uncertainty about immunogenicity, such as differences in certain impurities giving rise to questions related to immunogenicity.

Immunogenicity Assays for Biosimilar Programs

A one-assay approach for ADA assays, recommended in the FDA guidance [66] and 2016 White Paper in Bioanalysis [15], is widely used for the detection and characterization of ADAs in comparative clinical trials.

Immunogenicity assay validation parameters are the same for innovator and biosimilar programs, and include assay cut points, sensitivity, drug tolerance, specificity, selectivity, precision and reproducibility, robustness and reagent stability and system suitability controls. However, when the one-assay approach is used in a biosimilar program, additional data are needed to demonstrate the assay performs similarly with both products. This was exemplified in a talk by FDA with case studies, including the assay parameters from 5 approved biosimilars to US-licensed Humira. For ADA binding assays, all programs used the one-assay approach. Similarly, the one-assay approach was used for NAb assays, regardless of assay format being CBAs or CLBAs.

It is possible that differences in assay performance result in higher rates of ADAs or NAbs in the comparative immunogenicity testing when compared to the published data. Therefore, immunogenicity data from biosimilar studies need to be assessed using parallel arm studies and be interpreted in the context of the assays used and the totality of the data, including the clinical relevance of the detected ADAs. Although the ADA incidence for a compound may be increased when using highly sensitive ADA assays compared to historical data, the immunogenicity risk class most likely has not changed. While health authorities are aware of this, attention needs to be paid when discussing immunogenicity data in the literature as patients or physicians may not properly interpret the varying immunogenicity rates for the same originator product.

The concept for biosimilar immunogenicity evaluation is tailored for biosimilar development. Industry representatives presented their considerations specific for biosimilar immunogenicity assessment: the potential immunogenicity of the innovator product and published data are evaluated. Since the proposed biosimilar should match the reference product, the risk for new and not already described immune reactions is low. Depending on the immunogenicity risk assessment and product class (e.g., monoclonal antibodies or fusion proteins with human IgG Fc), alternative assay strategies to assess immunogenicity should be considered. In order to allow a sensitive and meaningful comparison of the immunogenic profile of a proposed biosimilar to its reference product, highly sensitive assays are required, including NAb assays. CLBAs are considered as powerful tools in biosimilar immunogenicity assessments due to their high sensitivity, while CBAs usually generate lower sensitivity and drug tolerance. In case anti-drug Fc antibodies may be of interest and potentially neutralize the efficacy of the drug, high levels of endogenous IgG Fc (approximately 8 mg/mL) in circulation, which compete for binding to anti-drug Fc antibodies, need to be factored in for assay development and data interpretation. Furthermore, the drug Fc region is highly conserved and unlikely to elicit ADA responses on its own. NAbs are often specific to the CDR

region. For mAb biotherapeutics with point substitutions on the Fc region (e.g., to enhance effector function), a NAb assay can be developed to assess antibodies specific to such mutations. However, non-neutralizing polyclonal antibodies to an Fc mutation have been reported. Use of anti-Fc positive controls may be challenging in the presence of high concentrations of endogenous IgG Fc in the sample matrix. Depending on the immunogenicity risk assessment, a NAb assay focusing on the antigen binding region of the biosimilar molecule could be sufficient to assess neutralizing capacity of a proposed biosimilar irrespective of its MoA.

2019 US FDA Immunogenicity Guidance: Reflections a Year Later

Building on the 2019 White Paper recommendations [25] on the interpretation and implementation of the 2019 US FDA Immunogenicity Guidance [26], further discussions were held on the impact of those recommendations on current industry practice.

FDA guidance on immunogenicity assay development and validation stresses the importance of selecting appropriate minimum required dilution (MRD) for the assay, but there is no industry perspective put forth in the literature. Z' factor was proposed for MRD selection but its utility is not ideal [64,69]. There is agreement on the importance of MRD but not on the definition of what constitutes an "appropriate" MRD per guidance.

Current EMA [57] and FDA [26] guidance documents display a fair degree of overlap in terms of scientific approach for immunogenicity testing [70]. Both agencies recommend using a tiered strategy for sample testing with sequential screening and confirmatory assays followed by semi-quantitation and characterization of ADA. Assessment of false negative (sensitivity) and false positive (specificity) rates are recommended by both EMA and FDA. Both agencies agree that 5% of false positives in the screening assay is recommended to minimize false negatives. Some concern was expressed regarding an interpretation of EMA guideline as recommending avoiding any false negatives; this would suggest that 5% of false positives can ensure not only high but infinite sensitivity. Industry representative discussed the value of 5% screening FPR vs. lower values (e.g. 1% or 0.1%). According to industry experience, assay sensitivity is dictated by the choice of analytical platform and not by increasing FPR. The consensus from both regulators and industry representatives was that the 5% FPR is reasonable for minimizing false negative and it allows for an initial assessment of clinical samples.

EMA guideline expects the confirmatory tier to eliminate any false positives while FDA guidance recommends 1% of false positive classifications in the confirmatory assay. It was unclear if the 0% and 1% FPR, respectively, referred to the confirmatory assay only or to the final data set. Industry consensus indicated that these FPR refer to the confirmatory cut point determination, which includes all samples. Confirmatory assays aim to identify true ADA positives; however, it is challenging to eliminate all false positives in the confirmatory assay. An FPR lower than 1% could be implemented if the assay demonstrates high sensitivity, specificity, and selectivity, and the data suggest the assay allows for a meaningful interpretation of the clinical data. The regulators recommend the industry consult with the regulatory agencies when alternative FPRs are desired.

FDA 2019 guidance suggests a 1% false positive rate for NAb assays performed on confirmed ADA positive samples. A confirmatory step for NAb assays is not routinely performed. However, there are cases when screening and confirmatory steps are needed for NAb assays. In that situation, the cut points for screening and confirmatory NAb assays should be determined case-by-case.

Industry representatives stated that ideal immunogenicity assays where both false negative and positive rates equal zero are impractical and not necessary and state-of-the-art immunogenicity assays can deliver reasonably high sensitivity without adjusting false positive rates. At the same time, immunogenicity assays can sometimes generate significantly more than 1% of false positives when applied to clinical samples. This could be caused by conservative outlier exclusion during determination of assay cut points [71] and failure to differentiate between analytical and biological variability [72]. The use of non-orthogonal confirmatory assays can also contribute to excessive false positives [73,74].

Regarding the impact of outliers, outlier removal is typically based on Tukey's box-plot [27], but there are disadvantages to using a box-plot with 1.5- or 3-IQR to identify outliers in cut point data. Industry representative stated that the box plot outlier removal assumes the distribution of assay responses for baseline samples is symmetrical, which is often not the case. A box-plot suggests which results should undergo more rigorous statistical analysis, but it cannot easily distinguish between analytical and biological outliers. It has been shown that eliminating outliers using a box-plot can result in cut points that are not appropriate for testing subjects from different study populations [71]. The regulators highlighted that box plot is the most common outlier removal methods seen in the

current submissions and appropriately justified in most cases. When in doubt, industry is encouraged to discuss statistical approaches for outlier removal and cut point determination with the regulatory agencies.

Regarding immunogenicity risk assessments, the FDA guidance [26] suggests performing a risk assessment during early development of the therapeutic drug. This could be initiated early, and include data collected during lead selection as well as rank ordering of the multiple variants of the candidate to bring the least risky sequence forward for further development. To accomplish a comprehensive risk assessment, appropriate tools and methods are needed to identify, characterize and mitigate the risks. A sound risk-based strategy requires a risk versus benefit analysis. A comprehensive end to end risk assessment requires an evaluation of both intrinsic and extrinsic risks spanning the life cycle of the therapeutic protein and includes product, process, patient and disease/treatment related factors [75]. Several tools have been developed and qualified for their ability to identify risks related to the therapeutic protein. Multiple algorithms that can evaluate the different aspects of antigen uptake, processing and presentation can be implemented early on to develop the ability to identify the intrinsic sequence-based risks. Additionally, the post translational modifications and other extrinsic risks related to the process, e.g., changes in product quality attributes (PQAs) that are noted during the later stage of process development, can be assessed using specialized human immune cell-based assays. Such assays can also capture risk of impurities like host cell protein related contaminants, excipient interaction with the therapeutic protein, and any immune modulatory target engagement related liability [76].

Additionally, the preclinical and clinical experience related to exposure, efficacy and safety can be captured through the course of development and the risk designation identified as low, moderate or high. The bioanalytical strategy for clinical trials can also be streamlined based on the identified risk and provided in the IND. An ongoing effort by industry and regulators for further harmonization of the risk assessment tools and their outputs will help support the adoption of these strategies. The risk assessment activities can be performed at key touch points in pipeline development and the assessments should be summarized in the form of key deliverables at various stages of development (IND, Clinical study reports, ISI/BLA etc.).

RECOMMENDATIONS

Below is a summary of the recommendations made during the 14th WRIB:

Cell-Based NAb Assays – Sensitivity and Drug Tolerance and the Relevance for Clinical Outcome

- It is recommended to use the biotherapeutic MoA as the primary consideration when deciding whether a functional cell-based NAb assay is used
    - **Cell-based NAb assay** are considered more reflective of mechanism action for assessment of ADA-mediated neutralization for agonists and many biologic therapeutics targeting cell-surface receptors. In these cases, CBA formats are generally preferred by most regulatory agencies
    - **CLBAs** may be more robust, sensitive and drug tolerant compared to cell based NAb assay, and can be used to support antagonistic biotherapeutics or receptor-Fc fusion proteins, when appropriately justified
- For **low risk molecules**:
    - A PD marker or biomarker that correlates with clinical outcome could be part of an integrated approach that uses PK, ADA, and PD data to assess NAb activity over time during treatment. This would rely on quality data from PK, ADA, and PD assays
- For **NAb assays** with good drug tolerance but compromised sensitivity:
    - Two NAb assay platforms can be used: one NAb assay with a higher drug tolerance and low sensitivity during treatment and one NAb assay with a higher sensitivity and lower drug tolerance for the follow-up samples during the wash-out period
    - The two methods must be bridged for use in the same study and the rationale for using two assay platforms should be clearly described

- Sponsors could use actual drug concentrations in NAb positive clinical samples to reassess assay drug tolerance in the study phase to help justify the adequacy of the assay; these results can be provided as supplemental data to support the validation report

NAb Assay Harmonization: Recent Trends and Expectations

- Immunogenicity risk, ADA titers and incidence, and the potential clinical consequence of NAb response are not currently sufficient to predict neutralizing activity and thus assessment of NAb activity may be needed
- The need for development, validation and implementation of ADA and NAb assays early in product development is dictated by the risk of the product
- To avoid holding up the drug development program, for low-risk products, NAb assays can be developed in parallel with early studies and a decision to implement made later based on the results on incidence and titer of ADA

Biosimilar Immunogenicity: Current Industry Standards

- "If a comparative analytical assessment based on state-of-the-art technology supports a demonstration of "**high similarity**" for a proposed biosimilar or interchangeable insulin product, there would be little or no residual uncertainty regarding immunogenicity. In such instances, a comparative clinical immunogenicity study would generally not be necessary to support licensure of a proposed biosimilar or interchangeable insulin product [68]"
- A one-assay approach for ADA binding assays and NAb assays is widely used for the detection and characterization of ADAs in comparative clinical studies
    - If this approach is used instead of the two-assay approach additional data are needed to demonstrate the assay performs similarly with both products (antigenic equivalence)
- The generated **immunogenicity data** from biosimilar studies need to be interpreted in the context of the assays used and the totality of the data, including the clinical relevance of the detected ADAs
- Attention needs to be paid when discussing immunogenicity data in the literature as patients or physicians may not properly interpret the varying immunogenicity rates for the same originator product

The 2019 US FDA Immunogenicity Guidance: Reflections a Year Later

- Assay sensitivity is influenced by the choice and performance of analytical platform
- Adjusting FPR should not be used to manipulate sensitivity
- The **5% FPR** is widely accepted by industry and the regulatory agencies. **It** is a reasonable choice for ADA assays and allows for an initial assessment of clinical samples
- An **FPR lower than 1%** for the confirmatory assay could be implemented if the assay demonstrates high sensitivity, specificity, and selectivity and allows for a meaningful clinical immunogenicity data interpretation
- In rare cases when there is a need for separate screening and confirmatory steps for NAb assay. The cut points for **screening and confirmatory NAb assays** should be determined case by case
- A risk assessment strategy starting from early discovery and progressing through the entire life cycle of the product development is preferred. The risk assessment can be adjusted based on knowledge obtained, summarized at development milestones and submitted through key regulatory submissions. It is recommended that the entire risk assessment be provided as part of integrated summary of immunogenicity in the marketing application

SECTION 4 – Immunogenicity Assay Strategies

Roland F. Staack[25], Bonnie Wu[2], Haoheng Yan[4], Daniela Verthelyi[4], Susan Kirshner[4], Joao Pedras-Vasconcelos[4], Elana Cherry[12], Isabelle Cludts[26], Boris Gorovits[17], Sumit Kar[11], Meina Liang[37], Mohsen Rajabi Abhari[4], Manoj Rajadhyaksha[24], Becky Schweighardt[16], Kay Stubenrauch[25], Yuanxin Xu[38], Lucia Zhang[12]

Authors in Section 4 are presented in alphabetical order of their last name, with the exception of the first six authors who were major contributors. Author affiliations can be found at the beginning of the article.

The affiliations can be found at the beginning of the article

DISCUSSION TOPICS & CONSOLIDATED QUESTIONS COLLECTED FROM THE GLOBAL BIOANALYTICAL COMMUNITY

The topics detailed below were considered as the most relevant 'hot topics' based on feedback collected from the 13th WRIB attendees. They were reviewed and consolidated by globally recognized opinion leaders before being submitted for discussion during the 14th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Lessons Learned from Late Stage Clinical Studies

How long should immunogenicity responses be monitored in the clinic until considered fully characterized? If anti-drug immune responses are not associated with clinical impact in the trials, should immunogenicity monitoring be required in the post-marketing setting? If so, how long should it be? Should clinical monitoring be emphasized, rather than immunogenicity monitoring, in the post-approval setting? Should immunogenicity testing only be triggered post-approval in the event of a new safety signal, reduced efficacy, or altered PD marker? What type of immunogenicity monitoring is relevant in the post-market setting versus being unwarranted and of limited clinical value?

Circulating Immune Complexes

Circulating Immune Complexes (CIC) can be formed upon drug administration due to immunogenicity/ADA formation and presence of soluble (multimeric) ligands. What is the industry best practice in bioanalysis for better understanding of CICs? Is size-specific bioanalysis of CIC possible by size exclusion chromatography (SEC)-LBA? Faster clearance of immune complexes (ICs) is mainly driven by high molecular weight CICs. Do you agree that the clearing property of ADA is a function of the size of the formed CICs? What is the industry experience in biodistribution regarding the clearance pathway of the CICs, impact of IgG subclasses/Fc variants on CIC PK, impact of CIC formation on PK, impact of CIC formation on immunogenicity, and CIC-complement complexes?

Multi-Domain Biotherapeutics: Immunogenicity Assay Strategies

For low risk multi-domain biotherapeutics (MDB) with low ADA incidence, is it necessary to evaluate ADA domain specificity? If so, would implementation of such an assessment only in Phase III be sufficient? If low ADA incidence does not allow impact assessments of domain specific ADA on the clinical outcome, how will the data be used during BLA review? For MDBs that have domains for function (e.g., ADC Ab and toxin), structure (e.g., ADC linker) or PK (e.g. Fc modification), is it sufficient to assess the domain specificity for only the functional domains and not for all domains? When will epitope mapping be needed instead of domain specificity? Will multiplex domain specificity assays be acceptable? Has this assay approach been implemented in drugs that have successfully received market approval?

Definition of Persistent ADA Response and its Clinical Relevance

Should an ADA response be characterized as "persistent" if only the last immunogenicity time point tested ADA positive? What if earlier time points fluctuate between positive and negative? Should a persistent response be defined as consistently positive over a certain period of time, such as 16 weeks? Does "less than 16 weeks before an ADA-negative last sample" truly represent a persistent response? Can persistence coupled with other characteristics (NAb, high titer) be better predictors of impact of immunogenicity? Should the term "persistent" only be used when there is an association with altered exposure, safety or efficacy?

DISCUSSIONS, CONSENSUS AND CONCLUSIONS

Lessons Learned from Late Stage Clinical Studies

There is continued concern that the immunogenicity risks identified in clinical trials will not accurately reflect the real-world clinical experience due to the limited numbers of patients in clinical trials, overly-restrictive eligibility criteria, assays with low sensitivity and drug tolerance, and relatively short-term follow-up after initial dosing in the clinical trial setting. As a result, the clinical relevance of the ADA responses may not be fully evaluated and unexpected immunogenicity-related safety events or reduced efficacy may occur once more patients are assessed for longer periods of time. This concern has been realized in rare examples where newly emergent immunogenicity-related issues have arisen in the commercial setting. For this reason, comprehensive long-term immunogenicity

monitoring is often required and included in pharmacovigilance and risk management plans in the post-approval setting.

The goal of post-approval monitoring of immunogenicity is to detect signals of clinically relevant ADA. These signals should be relatively apparent, such as a negative impact on PK, identification of a subset of patients losing efficacy or PD response over time, hypersensitivity responses occurring after repeat dosing, a safety signal associated with neutralization of an endogenous counterpart, and additional clinical outcomes (acute or chronic) that are suggestive of immunogenicity related mechanisms.

The best strategy to adopt should be science-based, data- and purpose-driven. Literature has no shortage of reports that provide industry experience from protein therapeutics [21,25,77]. Case studies like those described below demonstrate that the immunogenicity section of product inserts contains useful information as well. Finally, regulatory guidelines are invaluable [26,57,78].

Despite the perceived need to continue vigilant post-approval immunogenicity monitoring, many cases have demonstrated that data obtained from late-stage clinical trials have accurately predicted the real-world immunogenicity experience. It was agreed that there are several case studies within the industry experience in which more sensitive bioanalytical assays and re-analysis of pivotal clinical trial data were performed due to a post-marketing request (PMR) yet similar immunogenicity findings were found. Relevant, meaningful and quality clinical trial data could predict patients' immunogenicity in the post-approval setting.

One case regarded an approved intravenous administered enzyme replacement therapy for which the FDA requested the optimization of the NAb assay and the inclusion of a titer-based step, as well as re-optimization of the drug specific IgE assay in the post-marketing setting. Following re-optimization of the assays, the sponsor was asked to re-test all the clinical specimens from the pivotal clinical trial with the new assays, as well as investigate a prophylactic immune tolerance regimen in a cohort of patients treated with the biotherapeutic. Low-level sporadic IgE positivity was detected in a newly developed ImmunoCAP assay but these results were not associated with or predictive of hypersensitivity, similar to the results obtained with the original IgE assay in the clinical trials. Likewise, the newly optimized NAb assay demonstrated comparable results to those derived from the original assay, demonstrating no association with efficacy endpoints and similarly sustained PD effect across all NAb quartiles [79]. These results supported the release from the immune tolerance PMR 5 years after drug approval without the need for performing an immune tolerance induction (ITI) regimen.

Conversely, clinically relevant ADA was detected for bacterially derived enzyme replacement therapy where all patients developed a biphasic immune response [80]. As a result of the wide immunogenicity monitoring, information was incorporated into the label instructing physicians to titrate the dosage in a step-wise manner based on tolerability to achieve an effective maintenance dosage and consider increasing the dosage in patients who have not achieved phenylalanine reduction. Extensive immunogenicity assay optimization and re-testing was requested in the post-approval setting, despite the data not being used to inform clinical decision making. In addition, the sponsor received a PMR to perform ITI in an attempt to reduce the overall immunogenicity against the biotherapeutic.

Based on industry experience with assessing risks and benefits of biotherapeutics in the post-market setting, recommendations were provided on how to decide what is justifiable immunogenicity monitoring and what is unwarranted and of limited clinical value. Sufficient ADA data to understand clinical response and impact should be available at the time of submission. Monitoring should be based on scope of available data, clinical signals (safety and efficacy related observations), and the overall clinical impact of the immunogenicity data. In cases where the ADA analysis from the clinical trials demonstrates a lack of clinical impact on PK, PD, and/or safety profiles, continued immunogenicity monitoring may be unwarranted in the post-marketing setting and should be discussed with the Agency.

Clinical impact (triggers), for example, a decrease in efficacy or an emergent safety finding, should be defined by the sponsor on a case-by-case basis depending on the drug mode of action. Establishing clinical triggers for immunogenicity monitoring is more relevant than defining titer-based threshold criteria for long-term testing.

For chronic multi-dose treatments, a time frame of a minimum of 1–2 years of total clinical experience per patient was proposed. One year is typically sufficient for a single dose therapy while 2 years may be needed for multi-dose product depending on dosing intervals. Clinical experience may include the pivotal study period and post-market monitoring. For high immunogenicity risk products, active discussions with the relevant regulatory agencies are encouraged throughout the product life-cycle management. Acute treatments may need special considerations based on the dosing regimen and clinical risk.

If PMR and/or PMC is warranted, to determine when enough post-market immunogenicity monitoring has been completed and the anti-drug immune response has been sufficiency evaluated, an endpoint should be defined in the clinical protocol based on clinical triggers. If, within the 1 to 2F-year monitoring, there is no observed immunogenicity-related clinical signals (loss of efficacy or immunogenicity related safety findings), cessation of immunogenicity monitoring should be considered. PK and PD data are usually not required as part of PMR and/or PMC although that information is valuable for understanding the correlation of immunogenicity and clinical signals.

Further industry and health authority discussions are needed to determine when there is value in continuing to monitor immunogenicity, develop more sensitive bioanalytical assays, and re-analyze immunogenicity data in the post-market setting.

Circulating Immune Complexes

Administration of therapeutic proteins might evoke the formation of immune complexes. The occurrence of an immune response with the formation of ADA which bind the therapeutic protein is one potential mechanism leading to the formation of immune complexes. CICs can influence drug activity, PK and safety. CIC formation can reduce the clearance of the target and/or drug [81], while CIC clearance via FcγR and/or complement can increase drug clearance [82,83]. Clearance of CIC depends on the size and structure (e.g., lattice) [84] and immunoglobulin class within the complex. CIC can also provoke adverse events, like hypersensitivity reactions [80,85,86]. A detailed understanding of the immune complex formation, structure, and PK could improve drug development and interpretation of preclinical/clinical data.

Methodologies for measuring CICs are evolving. SEC and asymmetric flow field flow fractionation (A4F) can detect CICs but provide no actual size information if not coupled to specific detectors such as Multi Angle Light Scattering (MALS). Furthermore, application of these technologies to analyze biological samples is challenging with regard to sensitivity and specificity [87,88]. There is a need to develop methods to better characterize CIC, particularly in biological samples, in terms of their size and to better understand their time of onset, structure, and their impact on PK. Detailed knowledge in this field is of increasing complexity and growing importance with the increasing complexity of newly-designed biotherapeutics. Major bioanalytical challenges include how to monitor and evaluate CIC in the product life cycle. However, progress has been made for accurate size-specific CIC bioanalysis by use of SEC-LBA [89]. These methods are best suited for research purposes to better understand drug characteristics and to design improved therapies with the desired PK and clearance profiles. Decisions regarding clinical monitoring of CIC needs to be based on potential for clinically-relevant outcomes.

Recommendations were given on the industry best practice in bioanalysis for better understanding of CIC and its role in clinical studies. First, a consistent definition of the term CIC is required. The term CIC could describe different complexes such as drug-ADA, drug-target, or complement complexes. In addition, a better understanding of *in vivo* significance of CICs is needed. There is a differentiation between clinical relevance and research relevance (e.g., clinical/safety events vs. PK). Safety and impact-based monitoring was recommended for CIC. If there is a safety signal in non-clinical or early phase studies, then relevant CIC (e.g., complement complexes) measurement could be added to the protocol [80]. The aim would be to understand whether CIC is involved in the mechanism of the clinical/safety events.

So far, commercial generic functional CIC assays based on complement activation have been used, which are not drug- and size-specific but more practical in a clinical setting and likely demonstrate association with some of the safety related signals. Whether drug- and size-specific analysis adds clinical value needs further discussion and investigation. Immunoglobulin isotype (such as IgM) and IgG subclass of ADA may determine the ability to impact drug PK profile and safety more than the CIC size. In addition, very limited data is available on whether CIC size specific analysis is feasible [90]. Given the concentration of drug and ADA in each patient varies over time, the size (based on molar ratio of each component) of the CIC is expected to vary from patient to patient. Size specific analysis is technically challenging, and more investigation is necessary to understand the relevance to CIC clearance. At present, these topics are considered highly relevant for research purposes to understand product attributes but only limited data is available. Further discussions at future WRIB meetings are needed to update and expand the recommendations based on newly acquired industry experience. For now, clinical safety signals matter more than CIC size measurement, because the correlation between the two is unknown.

Multi-Domain Biotherapeutics: Immunogenicity Assay Strategies

With recent advances in technology, biologics became more complex with the emergence of multi-domain modalities. Multi-domain biotherapeutics (MDBs) contain complex structures multiple functional domains, and, sometimes, a multi-step pharmacological MoA. Examples include ADC, Fc-fusion proteins, PEGylated proteins, and bi/multi-specifics [91–93].

Like other biotherapeutic molecules, the PK/PD and exposure-response correlation for MDBs rely heavily on the bioanalytical strategy employed for their characterization. However, the multi-functional domains and dynamic nature of these complex biological entities require additional scientific considerations for selecting an appropriate bioanalytical strategy. MDBs can induce immune response to various domains, which may have different effects on the PK, PD, safety and efficacy of the molecule. Therefore, characterization of ADA domain specificity has become a regulatory expectation for this class of molecule.

The conventional approach to assess domain specificity uses competitive inhibition with domain-containing molecules. Although it is advantageous to employ a single assay, this type of method often has a unique bioanalytical challenge for detection of domain specific ADA subtypes, leading to false negative classification and thus an inaccurate immunogenicity assessment. An alternative approach to overcoming this challenge is to determine various ADA subtypes with multiple assays. However, it is resource intensive to develop multiple assays and challenging to compare results across assays.

Other unique bioanalytical challenges for MDBs include the need to better characterize *in vivo* biotransformation and/or structural instability, generation of multiple domain-specific critical reagents, evaluation of potential assay interference(s) due to multiple circulating targets and/or pre-existing/treatment-emergent ADA against individual functional domains, quantitation of multiple analytes, requirement of high assay sensitivity and alternate analytical platforms owing to the low clinical doses, and assay continuity across different development stages of the program.

Since there is no single bioanalytical strategy that works for all MDBs, the strategy will need to be adapted on a case-by-case basis depending on the MDB structure, its MoA and target biology, and the PK questions that need to be addressed at a specific development stage of the program.

The 2016 and 2018 White Papers in Bioanalysis [15,21] recommended domain specificity for MDBs at the confirmatory assay tier to characterize immune dominant regions. When developing an assay, an approach based on the complexity of therapeutic MoA was recommended. Discussions on the use of streamlined competitive binding assays using components of MDBs versus more demanding assays for individual domains were inconclusive. Sponsors were recommended to consult with the corresponding regulatory agencies. However, two years of additional experience with MDBs has allowed the initial recommendations on the need for domain specificity characterization and the best bioanalytical approaches to be updated in this White Paper.

The need for evaluating ADA domain specificity should be based on a risk assessment that evaluates the clinical and safety risk instead of a business risk. A high risk MDB has domains with high homology to endogenous counterparts whereas low risk MDBs refer to lack of structural homology and low risk of cytokine release syndrome [94]. The best approach is to first monitor total ADA. If total ADA is positive with impact on PK/PD and safety, then domain specificity should be analyzed. ADA domain specificity for the relevant domains as part of MDBs is not considered equivalent to ADA antigenic epitope mapping. B cell epitopes to proteins could be either linear amino acid based and/or 3D-structure based conformational epitopes given the nature of a polyclonal antibody response. Epitope mapping is part of the tier 3 immunogenicity testing as part of ADA characterization if patient ADA showed negative impact on PK, PD, and/or safety.

Characterization of ADA domain specificity is recommended (tier 3 of immunogenicity testing) and should assess responses to all functional domains that play a role in mechanism of action of the drug. In certain cases where characterization information from phase I and II is sufficient to support low clinical risk, further analysis should be discussed with regulatory authorities. However, most often the data from phase I & II studies is insufficient to make a determination.

For MDBs that have domains for function (e.g., ADC Ab and toxin) and for structure (e.g., ADC linker) or PK (e.g., Fc modification), it was discussed whether it is sufficient to assess the domain specificity for the functional domains only. From a regulatory perspective, data are only necessary for the functional domains but additional data may be helpful to sponsors for follow-up molecules. For example, Fc domain specificity characterization is relevant if it has a pharmacological function or is modified/engineered. This can be determined before phase III and sufficient sample volume should be collected to perform retrospective domain analysis if needed. Furthermore,

although only functional domain information is recommended, epitope mapping may be useful to improve a platform or molecule for follow-up generations.

Characterization of ADA domain specificity presents technical challenges, particularly for low titer total ADA samples, and several methodologies exist [95,96]. It was determined that multiplex domain specific assays are acceptable, but experience is very limited. Recently, an approach has been published to overcome technical challenges of the competitive domain specificity method [97]. A review of the different approaches, technical aspects and challenges for MDB assays was recommended as a topic for a future WRIB.

Definition of Persistent ADA Response and its Clinical Relevance

There is an acknowledged lack of relationship between the reported ADA incidence and associated clinical parameters for example, PK, efficacy, safety [15]. Magnitude (titer) and duration (persistence) of the ADA response may correlate better with clinical impact of ADA, rather than the ADA incidence [26,57]. Although methods for objective assessment of the duration of ADA responses have been proposed, colloquial classifications into transient and persistent responses are more prevalent. Industry and regulatory harmonization efforts have provided some basis for the definition of "persistent" ADA response. However, use of these definitions with some modifications is observed in practice. The impact of these modifications on the correlation between duration of ADA response and clinical safety and efficacy was discussed.

The current definition of a persistent ADA response, as outlined by Shankar et al. [98], is a "treatment-induced ADA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ADA-positive samples (irrespective of any negative samples in between) are separated by a period of 16 weeks or longer, or treatment-induced ADA incidence only in the last sampling time point of the treatment study period or at a sampling time point with less than 16 weeks before an ADA-negative last sample".

Since there is no FDA guidance for the nomenclature and the EMA guideline references the Shankar paper, the possibility of employing an alternative definition for "persistence" was discussed. It was discussed that low-titer ADA positive results near the assay cut point that do not show upward trending in titer and/or consistent positivity over time are unlikely to be clinically meaningful and should be distinguished from a clinically meaningful treatment-induced persistent ADA response.

The impact of the following definition was evaluated in a presented case study: "Treatment-emergent ADA positive response with two or more consecutive ADA positive sampling time points, separated by at least 12/16-week period (based on nominal sampling time), with no ADA negative samples in between, regardless of any missing samples". Using this definition, three categories were possible: persistent (potential impact on drug exposure; may or may not translate to impact on efficacy), transient (no impact on drug exposure, efficacy or patient safety), and indeterminate (generally no impact on drug exposure and/or efficacy; no impact on patient safety). An analysis of clinical study data resulted in some patients identified as persistent by the working definition "switch" to the transient or indeterminate category using the modified definition. Preliminary analysis suggested that patients identified as "persistent" by the working definition may have drug concentration profiles that appear to be similar to ADA negative patients. Considering inter-patient variability, patients identified as "ADA persistent" by the modified definition exhibited a noticeable impact on drug concentration profile. Analysis is underway in other studies to assess if these observations are consistent across studies.

This topic was extensively discussed in the 2016 White Paper in Bioanalysis [15]. The recommendations stated:

- "To properly classify the ADA duration data, at least 1 year of immunogenicity data were recommended. A 16- or 12-week duration were both deemed acceptable to characterize the patient as having an ADA as persistent response and an adequate sampling schedule should be selected"
- "The duration of follow-up testing for ADA-positive patients should be data driven and should not default to continuous monitoring until subjects become baseline negative"

These variations led to discussion to bring consistency to the definition of "persistent". It was recommended that the existing white paper defining transient versus persistent responses is sufficient [98]. This information is rarely placed on labels but if it is, the sponsor definition should be stated. If the sponsor deviates from the working definition for example, using the 12 weeks criteria instead of the 16 weeks, it is recommended that this be clearly stated in the dossier. However, there are challenges to interpretation of low-level positive responses near to the assay cut point, potential false positives particularly if no clinical relevance is associated. Therefore, discussions on the

term "persistent" may be required with other stakeholders for example, physicians, because of the potential for misinterpretations with clinical relevance.

RECOMMENDATIONS

Below is a summary of the recommendations made during the 14th WRIB:

Lessons Learned from Late Stage Clinical Studies

- **Clinically relevant ADA signals** should be relatively apparent in the form of negatively affected PK after repeat dosing, a subset of patients losing efficacy or PD response over time, occurrence of hypersensitivity responses, or a safety signal associated with neutralization of an endogenous counterpart
- The best **strategy for post-approval monitoring of immunogenicity** should be science-based and data-driven. Published data, case studies, labels, and guidelines should be used to inform decisions
- Sufficient **immunogenicity data** to understand clinical impact should be collected and available at the time of approval
- The need for immunogenicity monitoring should be based on safety signals and clinical impact
- **High ADA titers** alone may have little clinical relevance and monitoring until samples become negative may be challenging and of limited clinical value
- Clinical triggers for post-marketing immunogenicity monitoring should be defined by the sponsor on a case-by-case basis depending on the mode of action of the biotherapeutic. Establishing clinical triggers for immunogenicity monitoring is more relevant than defining a titer threshold
- For chronic treatments, a time frame of a minimum of 1–2 years of total clinical experience per patient was proposed, depending on the immunogenicity risk categorization of the drug. Clinical experience may include the pivotal study period and post-marketing monitoring, if warranted
- For **high risk products**, active discussions with the relevant regulatory agencies are encouraged throughout the product lifecycle. Acute treatments may need special considerations based on the regimen and clinical risk
- An endpoint should be defined based on clinical triggers to determine when enough post-marketing immunogenicity monitoring has been collected and the anti-drug immune response has been sufficiently assessed
- **PK and PD data** are valuable for understanding the correlation of immunogenicity and clinical consequence but not required

Circulating Immune Complexes

- A consistent definition of the term CIC needs to be established for a better understanding of the *in vivo* significance of CICs
- It is important to continue with the collective effort from industry to understand the **immune complex formation**, structures, and impact on PK to improve drug development and interpretation of preclinical/clinical data
- It was agreed that additional studies are needed to better characterize CIC in terms of their size and to better understand their formation, structure, PK, and clinical relevance
- **Safety and impact-based monitoring** in patients was recommended for CIC. If a CIC-related clinical signal is identified in non-clinical or early phase studies, then inclusion of CIC measurement in the phase III clinical protocol should be considered
- Functional CIC based on complement activation may be more suited for clinical impact driven analysis, rather than size-based CIC analysis

Multi-Domain Biotherapeutics: Immunogenicity Assay Strategies

- It was agreed that **characterization of ADA domain specificity** has become a regulatory expectation for multi-domain biotherapeutics
    - **The conventional approach** to assess domain specificity uses competitive inhibition with domain-containing molecules
    - **An alternative approach** to overcoming the challenge of detection of low level ADA subtypes is to determine various ADA subtypes with multiple domain detection assays
    - **Domain specificity** is not considered equivalent to epitope mapping

- As such there is no single bioanalytical strategy that works for all MDBs; rather the strategy will need to be adapted on a case-by-case basis depending on the MDB structure and its MoA, target biology, and the PK question that needs to be addressed at a specific development stage of the program
- The need for evaluating ADA domain specificity should be based on a **risk assessment** that evaluates the clinical and safety risk instead of a business risk. The best approach is to monitor total ADA, and if ADA is positive with impact on PK/PD and safety, then domain specificity can be analyzed
- Characterization of ADA domain specificity is recommended (tier 3 of immunogenicity testing), especially for **low risk molecules**, to help elucidate ADA domain binding specificity and its impact if there is a clinical signal. If characterization information from phase I and II indicates low clinical risk, this immunogenicity data may be sufficient to streamline Phase III monitoring
- Data are only necessary for the functional domains, but evaluation of other domains, including epitope mapping, may be useful for the design of a novel drug with lower immunogenicity. **Multiplex domain specific assays** are acceptable, but experience is very limited due to technical challenges

Definition of Persistent ADA Response and its Clinical Relevance

- It was confirmed that magnitude (titer) and duration (persistence) of the ADA response may correlate better with clinical impact of ADA than only considering incidence alone
- Consensus was achieved that the existing Shankar white paper definition of transient versus persistent responses is still adequate [98]
- If information on ADA persistence is included on the label, **the sponsor-specific definition of persistence should be provided in the label**, particularly if the sponsor deviates from the standard definition (e.g., using the 12 weeks criteria instead of the 16 weeks)

Acknowledgements

- The US FDA, Europe EMA, UK MHRA, Health Canada, Norway NoMA, Brazil ANVISA, Japan MHLW, and for supporting this workshop
- Dr. Steven Alley (Seattle Genetics), Dr. Anna Edmison (Health Canada), Dr. Chris Evans (GSK), Dr. Christine Fandozzi (Merck), Dr. Sally Fischer (Genentech), Dr. Fabio Garofolo (BRI), Dr. Christine Grimaldi (Boehringer Ingelheim), Dr. Lindsay King (Pfizer), Dr. Rocio Murphy (Merck), Dr. Hendrik Neubert (Pfizer), Dr. Manoj Rajadhyaksha (Regeneron), Dr. Catherine Soo (Health Canada), Dr. Susan Spitz (Incyte), Dr. Roland F. Staack (Roche), Dr. Scott Summerfield (GSK), Dr. Alessandra Vitaliti (Novartis), Drs. Jan Welink (EMA), Dr. Haoheng Yan (US FDA), Dr. Tong-yuan Yang (Janssen), Dr. Hongbin Yu (Boehringer Ingelheim), Dr. Yan Zhang (BMS) for chairing the workshop and/or the White Paper discussions
- All the workshop attendees and members of the Global Bioanalytical Community who have sent comments and suggestions to the workshop to complete this White Paper
- Future Science Group as a trusted partner

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. Stevenson L, Rocci M, Garofolo F *et al.* 2013 White Paper on Recent Issues in Bioanalysis: "Hybrid" – the best of LBA & LC/MS. *Bioanalysis* 5(23), 2903–2918 (2013).
7. Fluhler E, Hayes R, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 1 – small molecules by LCMS). *Bioanalysis* 6(22), 3039–3049 (2014).
8. Dufield D, Neubert H, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 2 – hybrid LBA/LCMS, ELN & regulatory agencies' input). *Bioanalysis* 6(23), 3237–3249 (2014).
9. Stevenson L, Amaravadi L, Myler H *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 3 – LBA and immunogenicity). *Bioanalysis* 6(24), 3355–3368 (2014).
10. Welink J, Fluhler E, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 1 – small molecules by LCMS). *Bioanalysis* 7(22), 2913–2925 (2015).
11. Ackermann B, Neubert H, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 2 – hybrid LBA/LCMS and input from regulatory agencies). *Bioanalysis* 7(23), 3019–3034 (2015).
12. Amaravadi L, Song A, Myler H *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 3 – LBA, biomarkers and immunogenicity). *Bioanalysis* 7(24), 3107–3124 (2015).
13. Yang E, Welink J, Cape S *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV) (Part 1 – small molecules, peptides and small molecule biomarkers by LCMS). *Bioanalysis* 8(22), 2363–2378 (2016).
14. Song A, Lee A, Garofolo F *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV): (Part 2 – Hybrid LBA/LCMS and input from regulatory agencies). *Bioanalysis* 8(23), 2457–2474 (2016).
15. Richards S, Amaravadi L, Pillutla R *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV): (Part 3 – LBA, biomarkers and immunogenicity). *Bioanalysis* 8(23), 2475–2496 (2016).
16. Welink J, Yang E, Hughes N *et al.* 2017 White Paper on recent issues in bioanalysis: aren't BMV guidance/guidelines 'Scientific'? (Part 1 – LCMS: small molecules, peptides and small molecule biomarkers). *Bioanalysis* 9(22), 1807–1825 (2017).
17. Neubert H, Song A, Lee A *et al.* 2017 White Paper on recent issues in bioanalysis: rise of hybrid LBA/LCMS immunogenicity assays (Part 2: hybrid LBA/LCMS biotherapeutics, biomarkers & immunogenicity assays and regulatory agencies' inputs). *Bioanalysis* 9(23), 1895–1912 (2017).
18. Gupta S, Richards S, Amaravadi L *et al.* 2017 White Paper on recent issues in bioanalysis: a global perspective on immunogenicity guidelines & biomarker assay performance (Part 3 – LBA: immunogenicity, biomarkers and PK assays). *Bioanalysis* 9(24), 1967–1996 (2017).
19. Welink J, Xu Y, Yang E *et al.* 2018 White Paper on Recent Issues in Bioanalysis: 'A global bioanalytical community perspective on last decade of incurred samples reanalysis (ISR)' (Part 1 – small molecule regulated bioanalysis, small molecule biomarkers, peptides & oligonucleotide bioanalysis). *Bioanalysis* 10(22), 1781–1801 (2018).
20. Neubert H, Olah T, Lee A *et al.* 2018 White Paper on Recent Issues in Bioanalysis: focus on immunogenicity assays by hybrid LBA/LCMS and regulatory feedback (Part 2 – PK, PD & ADA assays by hybrid LBA/LCMS & regulatory agencies' inputs on bioanalysis, biomarkers and immunogenicity). *Bioanalysis* 10(23), 1897–1917 (2018).
21. Stevenson L, Richards S, Pillutla R *et al.* 2018 White Paper on Recent Issues in Bioanalysis: focus on flow cytometry, gene therapy, cut points and key clarifications on BAV (Part 3 – LBA/cell-based assays: immunogenicity, biomarkers and PK assays). *Bioanalysis* 10(24), 1973–2001 (2018).
22. Garofolo W, Savoie N. The Decennial Index of the White Papers in Bioanalysis: 'A Decade of Recommendations (2007–2016)'. *Bioanalysis* 9(21), 1681–1704 (2017).
23. Fandozzi C, Evans C, Wilson A *et al.* 2019 White Paper on Recent Issues in Bioanalysis: Chromatographic Assays (Part 1 – Innovation in Small Molecules and Oligonucleotides & Mass Spectrometric Method Development Strategies for Large Molecule Bioanalysis). *Bioanalysis* 11(22), 2029–2048 (2019).
24. Booth B, Stevenson L, Pillutla R *et al.* 2019 White Paper On Recent Issues in Bioanalysis: FDA BMV Guidance, ICH M10 BMV Guideline and Regulatory Inputs (Part 2 – Recommendations on 2018 FDA BMV Guidance, 2019 ICH M10 BMV Draft Guideline and Regulatory Agencies' Input on Bioanalysis, Biomarkers and Immunogenicity). *Bioanalysis* 11(23), 2099–2132 (2019).
25. Piccoli S, Mehta D, Vitaliti A *et al.* 2019 White Paper on Recent Issues in Bioanalysis: FDA Immunogenicity Guidance, Gene Therapy, Critical Reagents, Biomarkers and Flow Cytometry Validation (Part 3 – Recommendations on 2019 FDA Immunogenicity Guidance, Gene Therapy Bioanalytical Challenges, Strategies for Critical Reagent Management, Biomarker Assay Validation, Flow Cytometry Validation & CLSI H62). *Bioanalysis* 11(24), 2207–2244 (2019).
26. US FDA, Center for Drug Evaluation and Research. Guidance for Industry: Immunogenicity Testing of Therapeutic Proteins Products – Developing and Validating Assays for Anti-Drug Antibody Detection (2019). https://www.fda.gov/media/119788/download
27. Shankar G, Devanarayan V, Amaravadi L *et al.* Recommendation for the validation of immunoassay used for detection of host antibodies against bioanalytical products. *J. Pharm. Biomed. Anal.* 48, 1267–1281 (2008).
28. Smith BP, Vandenhende FR, DeSante KA *et al.* Confidence interval criteria for assessment of dose proportionality. *Pharm. Res.* 17(10), 1278–1283 (2000).

29. Hendriks J, Stals C, Versteilen A *et al.* Stability studies of binding and functional anti-vaccine antibodies. *Bioanalysis* 6(10), 1385–1393 (2014).
30. Michaut L, Laurent N, Kentsch K *et al.* Stability of anti-immunotherapeutic antibodies in frozen human serum samples. *Bioanalysis* 6(10), 1395–1407 (2014).
31. CLSI. CLSI Guideline H62: Validation of Assays Performed by Flow Cytometry (2019). https://www.wrib.org/PDFs/Flow_H62_Proposed_Draft-190919.pdf
32. Cossarizza A, Chang H-D, Radbruch A *et al.* Guidelines for the use of flow cytometry and cell sorting in immunological studies (second edition). *European Journal of Immunology* 49(10), 1457–1973 (2019).
33. Perfetto SP, Ambrozak D, Nguyen R *et al.* Quality assurance for polychromatic flow cytometry. *Nat. Protoc.* 1(3), 1522–1530 (2006).
34. Perfetto SP, Ambrozak D, Nguyen R *et al.* Quality assurance for polychromatic flow cytometry using a suite of calibration beads. *Nature Protocols* 7, 2067–2079 (2012).
35. Bustin SA, Benes V, Garson JA *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55(4), 611–622 (2009).
36. CLSI. CLSI.org
37. Gorovits B, Koren E. Immunogenicity of Chimeric Antigen Receptor T-Cell Therapeutics. *BioDrugs* 33(3), 275–284 (2019).
38. Potthoff B, McBlane F, Spindeldreher S *et al.* A cell-based immunogenicity assay to detect antibodies against chimeric antigen receptor expressed by tisagenlecleucel. *J. Imm. Meth.* 476, 112692 (2020). https://doi.org/10.1016/j.jim.2019.112692
39. EMA 2018. Assessment Report: Kymriah (Tisagenleucel) (2018). https://www.ema.europa.eu/en/documents/assessment-report/kymriah-epar-public-assessment-report_en.pdf
40. Lamers CHJ, Klaver Y, Gratama JW *et al.* Treatment of metastatic renal cell carcinoma (mRCC) with CAIX CAR-engineered T-cells-a completed study overview. *Biochem. Soc. Trans.* 44(3), 951–959 (2016).
41. Annika CK *et al.* Comparison of the ELISPOT and cytokine flow cytometry assays for the enumeration of antigen-specific T cells. *J.Immunol. Methods* 283(1-2), 141–153 (2003).
42. Theresa LW *et al.* Enzyme-linked immunospot, cytokine flow cytometry, and tetramers in the detection of T-cell responses to a dendritic cell-based multipeptide vaccine in patients with melanoma. *Clin. Cancer Res.* 9(2), 641–649 (2003).
43. European Medicines Agency. EMA General Principles to Address Virus and Vector Shedding (EMEA/CHMP/ICH/449035/2009) (2009). https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-10.pdf
44. US FDA, Center for biologics evaluation and research. Guidance for industry: Design and Analysis of Shedding Studies for Virus or Bacteria-Based Gene Therapy and Oncolytic Products (2015). www.fda.gov/media/89036/download
45. Pharmaceuticals and Medical Devices Agency. PMDA Japan Guidance on ensuring the quality and safety of products for gene therapy (2019). https://www.pmda.go.jp/files/000235607.pdf
46. Rangarajan S, Walsh L, Lester W *et al.* AAV5-Factor VIII gene transfer in severe hemophilia A. *N. Engl. J. Med.* 377(26), 2519–2530 (2017).
47. Zen Z, Espinoza Y, Bleu T, Sommer JM, Wright JF. Infectious titer assay for adeno-associated virus vectors with sensitivity sufficient to detect single infectious events. *Hum. Gene Ther.* 15(7), 709–715 (2004).
48. Bustin S, Huggett J. qPCR primer design revisited. *Biomol. Detect. Quantif.* 14, 19–28 (2017).
49. Huggett JF *et al.* The Digital MIQE Guidelines: minimum information for publication of quantitative digital PCR experiments. *Clin. Chem.* 59(6), 892–902 (2013).
50. Wang D, Tai PWL, Gao G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discovery* 18, 358–378 (2019).
51. Gorovits B, Fiscella M, Havert M *et al.* Recommendations for the Development of Anti-Viral Vector Neutralizing Antibody Assays. *AAPS J.* 22(2), 24 (2020).
52. US FDA, Center for Biologics Evaluation and Research. Guidance for Industry: Human Gene therapy for Rare Diseases (2020). https://www.fda.gov/media/113807/download
53. Janetzki S, Panageas KS, Ben-Porat L *et al.* Results and harmonization guidelines from two large-scale international Elispot proficiency panels conducted by the Cancer Vaccine Consortium (CVC/SVI). *Cancer Immunol. Immunother.* 57(3), 303–315 (2008).
54. Janetzki S, Price L, Schroeder H *et al.* Guidelines for the automated evaluation of Elispot assays. *Nat. Protoc.* 10(7), 1098–1115 (2015).
55. Immudex. Proficiency Panels:How Well Do You Perform T-cell Immune Monitoring? https://www.immudex.com/proficiency-panels.aspx
56. ICH. ICH Q2R1: Validation of Analytical Procedures: Text and Methodology (2005). https://database.ich.org/sites/default/files/Q2%28R1%29%20Guideline.pdf

57. European Medicines Agency. EMA Guideline on Immunogenicity Assessment of Therapeutic Proteins.EMEA/CHMP/BMWP/14327/2006 Rev 1 (2017). https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-immunogenicity-assessment-therapeutic-proteins-revision-1_en.pdf
58. Calcedo R, Morizono H, Wang L *et al.* Adeno-associated virus antibody profiles in newborns, children, and adolescents. *Clin. Vaccine Immunol.* 18, 1586–1588 (2011).
59. Mingozzi F, Chen Y, Edmonson SC *et al.* Prevalence and pharmacological modulation of humoral immunity to AAV vectors in gene transfer to synovial tissue. *Gene Ther.* 20, 417–424 (2013).
60. Calcedo R, Vandenberghe LH, Gao G *et al.* Worldwide epidemiology of neutralizing to adenoassociated viruses. *J. Infect. Dis.* 199, 381–390 (2009).
61. Finco D, Baltrukonis D, Clemens-Egan A *et al.* Comparison of competitive ligand-binding assay and bioassay formats for the measurement of neutralizing antibodies to protein therapeutics. *J. Pharm. Biomed. Analysis* 54(2), 351–358 (2011).
62. Wu Y, Akhgar A, Li JJ *et al.* Selection of a ligand-binding neutralizing antibody assay for benralizumab: comparison with an antibody-dependent cell-mediated cytotoxicity (ADCC) cell-based assay. *AAPS J.* 20(3), 49 (2018).
63. Wu B, Chung S, Jiang XR *et al.* Strategies to determine assay format for the assessment of neutralizing antibody responses to biotherapeutics. *AAPS J.* 18(6), 1335–1350 (2016).
64. USP-NF <1106.1> Immunogenicity Assays – Design and Validation of Assays to Detect Anti-Drug Neutralizing Antibody (2015).
65. US FDA, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Guidance for Industry: Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product (2016). https://www.fda.gov/media/88622/download
66. US FDA, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Guidance for Industry: Scientific Considerations in Demonstrating Biosimilarity to a Reference Product (2015). https://www.fda.gov/media/82647/download
67. US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. Guidance for Industry, Bioanalytical Method Validation (2018). https://www.fda.gov/files/drugs/published/Bioanalytical-Method-Validation-Guidance-for-Industry.pdf
68. US FDA, Center for Drug Evaluation and Research. Draft Guidance for Industry: Clinical Immunogenicity Considerations for Biosimilar and Interchangeable Insulin Products (2019). https://www.fda.gov/media/133014/download
69. Devanarayan V, Tovey MG. Cut point and performance characteristics for anti-drug antibody assays. In: *Detection and Quantification of Antibodies to Biopharmaceuticals.* Michael G (Ed.). Tovey, Wiley (2011).
70. Kurki P. Compatibility of immunogenicity guidance by the EMA and the US FDA. *Bioanalysis* 11(17), 1619–1629 (2019).
71. Kubiak RJ, Zhang J, Ren P, Yang H, Roskos LK. Excessive outlier removal may result in cut points that are not suitable for immunogenicity assessments. *J. Imm. Meth.* 463, 105–111 (2018).
72. Zhang J, Arends RHGP, Kubiak RJ, Rosko LK, Liang M, Lee N, Chen C-K, Yang H. A new method for identification of outliers in immunogenicity assay cut point data. *J. Imm. Meth.* 484–485, 112817 (2020).
73. Kubiak RJ, Zhang J, Zhang J, Zhu Y *et al.* Correlation of screening and confirmatory results in tiered Immunogenicity testing by solution-phase bridging assays. *J. Pharm. Biomed. Anal.* 74, 235–245 (2013).
74. Kubiak RJ, Arends RGHP, Lee N, Liang M, Zhang J, Roskos LK. Confirmatory cut point has limited ability to make accurate classifications in Immunogenicity assays. *Bioanalysis* 12(4), 246–256 (2020).
75. Tovey MG, Lallemand C. Immunogenicity and other problems associated with the use of biopharmaceuticals. *Ther. Adv. Drug Safety* 2(3), 113–128 (2011).
76. Jawa V, Terry F, Gokemeijer J *et al.* T-cell dependent immunogenicity of protein therapeutics pre-clinical assessment and mitigation – updated consensus and review. *Front. Immunol.* 11, 1301 (2020).
77. Rosenberg AS, Sauna ZE. Immunogenicity assessment during the development of protein therapeutic. *JPP* 70(5), 584–594 (2018).
78. US FDA, Center for Biologics Evaluation and Research. Guidance for Industry: Human Gene Therapy for Rare Diseases (2020). https://www.fda.gov/regulatory-information/search-fda-guidance-documents/human-gene-therapy-rare-diseases
79. Andrew CM *et al.* Antibodies that neutralize cellular uptake of elosulfase alfa are not associated with reduced efficacy or pharmacodynamic effect in individuals with Morquio A syndrome. *J. Immunol. Meth.* 440, 41–51 (2017).
80. Gupta S *et al.* Association of immune response with efficacy and safety outcomes in adults with phenylketonuria administered pegvaliase in phase 3 clinical trials. *EBioMedicine* 37, 366–373 (2018).
81. Bruce NR, Moo JC. Antibodies as carrier proteins. *Pharm. Res.* 15(11), 1652–1656 (1998).
82. Hogarth PM. Fc receptors are major mediators of antibody based inflammation in autoimmunity. *Curr. Opin. Immunol.* 14(6), 798–802 (2002).
83. Chauhan AK, Moore TL. T cell activation by terminal complex of complement and immune complexes. *J. Biol. Chem.* 286(44), 38627–38637 (2011).

84. Rojas JR, Taylor RP, Cunningham MR *et al.* Formation, distribution, and elimination of infliximab and anti-infliximab immune complexes in cynomolgus monkeys. *J. Pharm. Exp. Ther.* 313(2), 578–585 (2005).

85. van Schouwenburg PA *et al.* Adalimumab elicits a restricted anti-idiotypic antibody response in autoimmune patients resulting in functional neutralisation. *Ann. Rheum. Dis.* 72(1), 104–109 (2013).

86. Krishna M, Steven GN. Immunogenicity to biotherapeutics – the role of anti-drug immune complexes. *Front. Immunol.* 7, 21 (2016).

87. Hoffmann E *et al.* Generation, characterization, and quantitative bioanalysis of drug/anti-drug antibody immune complexes to facilitate dedicated *in vivo* studies. *Pharm. Res.* 36(9), 129 (2019).

88. van Schie KA *et al.* Restricted immune activation and internalisation of anti-idiotype complexes between drug and antidrug antibodies. *Ann. Rheum. Dis.* 77(10), 1471–1479 (2018).

89. Regenass-Lechner F *et al.* Immunogenicity, inflammation, and lipid accumulation in Cynomolgus monkeys infused with a Lipidated Tetranectin-ApoA-I fusion protein. *Toxicol. Sci.* 150(2), 378–389 (2016).

90. Tada M, Takuo S, Akiko I-W. Development and characterization of an anti-rituximab monoclonal antibody panel. *MAbs* 10(3), 370–379 (2018).

91. Labrijn AF *et al.* Bispecific antibodies: a mechanistic review of the pipeline. *Nat. Rev. Drug Discov.* 18(8), 585–608 (2019).

92. Hoofring SA, Lopez R, Hock MB *et al.* Immunogenicity testing strategy and bioanalytical assays for antibody-drug conjugates. *Bioanalysis* 5(9), 1041–1055 (2013).

93. Stubenrauch K, Kunzel C, Vogel R, Tuerck D, Schick E, Heinrich J. Epitope characterization of the ADA response directed against a targeted immunocytokine. *J. Pharm. Biomed. Anal.* 114, 296–304 (2015).

94. Gorovits B, Peng K, Kromminga A. Current Considerations on Characterization of Immune Response to Multi-Domain Biotherapeutics. *BioDrugs* 34(1), 39–54 (2020).

95. Gorovits B *et al.* Recommendations for the characterization of immunogenicity response to multiple domain biotherapeutics. *J. Immunol. Methods* 408, 1–12 (2014).

96. Broders O, Wessels U, Zadak M, Beckmann R, Stubenrauch K. Novel bioanalytical method for the characterization of the immune response directed against a bispecific F(ab) fragment. *Bioanalysis* 12(8), 509–517 (2020).

97. Vainshtein I, Sun B, Roskos LK, Liang M. A novel approach to assess domain specificity of anti-drug antibodies to moxetumomab pasudotox, an immunotoxin with two functional domains. *J. Immunol. Methods* 477, 112688 (2020).

98. Shankar G, Arkin S, Cocea L *et al.* Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. *AAPS J.* 16(4), 658–673 (2014).

White Paper

For reprint orders, please contact: reprints@future-science.com

Bioanalysis

2019 White Paper on Recent Issues in Bioanalysis: FDA Immunogenicity Guidance, Gene Therapy, Critical Reagents, Biomarkers and Flow Cytometry Validation (Part 3 – Recommendations on 2019 FDA Immunogenicity Guidance, Gene Therapy Bioanalytical Challenges, Strategies for Critical Reagent Management, Biomarker Assay Validation, Flow Cytometry Validation & CLSI H62)

Steven Piccoli[‡,#,1], Devangi Mehta[‡,#,2], Alessandra Vitaliti[‡,#,3], John Allinson[‡,4], Shashi Amur[‡,5], Steve Eck[‡,#,6], Cherie Green[‡,#,7], Michael Hedrick[‡,#,8], Shirley Hopper[‡,9], Allena Ji[‡,10], Alison Joyce[‡,11], Virginia Litwin[‡,#,12], Kevin Maher[‡,5], Joel Mathews[‡,13], Kun Peng[‡,7], Afshin Safavi[‡,14], Yow-Ming Wang[‡,5], Yan Zhang[‡,8], Lakshmi Amaravadi[§,15], Nisha Palackal[§,22], Sai Thankamony[§,8], Chris Beaver[§,16], Eris Bame[§,#,2], Thomas Emrich[§,17], Christine Grimaldi[§,18], Jonathan Haulenbeek[§,8], Alison Joyce[§,11], Vellalore Kakkanaiah[§,#,19], David Lanham[§,#,20], Kevin Maher[§,#,5], Andrew Mayer[§,21], Paul C Trampont[§,#,23], Laurent Vermet[§,24], Naveen Dakappagari[#,25], Catherine Fleener[#,26], Fabio Garofolo[*,#,27], Cynthia Rogers[#,¥,28], Shabnam Tangri[#,¥,25], Yuanxin Xu[#,¥,29], Meina Liang[¶,30], Manoj Rajadhyaksha[¶,22], Susan Richards[¶,10], Becky Schweighardt[¶,31], Shobha Purushothama[¶,2], Daniel Baltrukonis[¶,32], Jochen Brumm[¶,7], Elana Cherry[¶,33], Jason Delcarpini[¶,34], Carol Gleason[¶,8], Susan Kirshner[¶,5], Robert Kubiak[¶,6], Luying Pan[¶,35], Michael Partridge[¶,22], João Pedras-Vasconcelos[¶,5], Qiang Qu[¶,36], Venke Skibeli[¶,37], Therese Solstad Saunders[¶,37], Roland F Staack[¶,17], Kay Stubenrauch[¶,17], Al Torri[¶,22], Daniela Verthelyi[¶,5], Haoheng Yan[¶,5], Boris Gorovits[¥,11], Rachel Palmer[¥,10], Mark Milton[¥,42], Brian Long[¥,31], Bart Corsaro[¥,38], Vahid Farrokhi[¥,11], Michele Fiscella[¥,39], Neil Henderson[¥,40], Vibha Jawa[¥,41], Jim McNally[¥,35], Rocio Murphy[¥,43], Hanspeter Waldner[¥,44] & Tong-Yuan Yang[¥,45]

[1]GlaxoSmithKline, Collegeville, PA, USA
[2]Biogen, Cambridge, MA, USA
[3]Novartis Pharma, Basel, Switzerland
[4]Immunologix Laboratories, Tampa, FL, USA
[5]US FDA, Silver Spring, MD, USA
[6]AstraZeneca, Gaithersburg, MD, USA
[7]Genentech, South San Francisco, CA, USA
[8]Bristol-Myers Squibb, Princeton, NJ, USA
[9]UK MHRA-NIBSC, London, UK
[10]Sanofi, Framingham, MA, USA
[11]Pfizer, Andover, MA, USA
[12]Caprion Biosciences, Montreal, QC, Canada
[13]Ionis Pharmaceutical, Carlsbad, CA, USA
[14]BioAgilytix, Durham, NC, USA
[15]Shire/Takeda, Lexington, MA, USA
[16]Syneos Health, Princeton, NJ, USA
[17]Roche Pharma Research & Early Development, Roche Innovation Center, Munich, Germany
[18]Boehringer Ingelheim, Ridgefield, CT, USA
[19]PPD, Richmond, VA, USA
[20]Eurofins Bioanalysis Services, Abingdon, UK
[21]GlaxoSmithKline, Collegeville, PA, USA
[22]Regeneron Pharmaceuticals, Tarrytown, NY, USA
[23]Covance Central Laboratory, Greenfield, IN, USA
[24]Sanofi, Montpellier, France
[25]Navigate BioPharma Services, Inc., a Novartis subsidiary, San Diego, CA, USA
[26]Pfizer, San Diego, CA, USA
[27]Angelini Pharma, Pomezia, RM, Italy
[28]bluebird bio, Cambridge, MA, USA
[29]Alnylam Pharmaceuticals, Cambridge, MA, USA
[30]AstraZeneca, South San Francisco, CA, USA
[31]BioMarin, Novato, CA, USA
[32]Pfizer, Groton, CT, USA
[33]Health Canada, Ottawa, ON, Canada
[34]Celldex Therapeutics, Needham, MA, USA
[35]Shire/Takeda, Cambridge, MA, USA
[36]EMD Serono/Merck KGaA, Billerica, MA, USA

[‡]SECTION 1 – New Insights in Biomarker Assay Validation (BAV) (authors are presented in alphabetical order of their last name, with the exception of the first three authors who were session chairs, working dinner facilitators, and/or notetakers).
[§]SECTION 2 – Current & Effective Strategies for Critical Reagent Management (authors are presented in alphabetical order of their last name, with the exception of the first four authors who were session chairs, working dinner facilitators, major contributors and/or notetakers).
[#]SECTION 3 – Flow Cytometry Validation in Drug Discovery & Development & CLSI H62 (authors are presented in alphabetical order of their last name, with the exception of the first four authors who were session chairs, working dinner facilitators, major contributors and/or notetakers).
[¶]SECTION 4 – Interpretation of the 2019 FDA Immunogenicity Guidance (authors are presented in alphabetical order of their last name, with the exception of the first five authors who were session chairs, working dinner facilitators, and/or notetakers).
[¥]SECTION 5 – *In vivo* and *Ex vivo* Gene Therapy and Vaccine Bioanalytical Challenges (authors are presented in alphabetical order of their last name, with the exception of the first four authors who were session chairs, working dinner facilitators, major contributors and/or notetakers).

Disclaimer: The views expressed in this article are those of the authors and do not reflect official policy of the US FDA, Europe EMA, France ANSM, UK MHRA, Brazil ANVISA, Health Canada, Japan MHLW, and Norway NoMA. No official endorsement by the FDA, EMA, ANSM, Health Canada, MHRA, ANVISA, MHLW or NoMA is intended or should be inferred.

*Author for correspondence: fabiogarofolo@hotmail.com

newlands press

[37]Norwegian Medicines Agency, Oslo, Norway
[38]GlaxoSmithKline, Rockville, MD, USA
[39]REGENXBIO, Rockville, MD, USA
[40]AstraZeneca, Molndal, Sweden
[41]Merck & Co., Inc., Kenilworth, NJ, USA
[42]Novartis, Cambridge, MA, USA
[43]Merck & Co., Inc., West Point, PA, USA
[44]CRISPR Therapeutics, Cambridge, MA, USA
[45]Janssen R&D, Spring House, PA, USA

The 2019 13th Workshop on Recent Issues in Bioanalysis (WRIB) took place in New Orleans, LA, USA on April 1–5, 2019 with an attendance of over 1000 representatives from pharmaceutical/biopharmaceutical companies, biotechnology companies, contract research organizations and regulatory agencies worldwide. WRIB was once again a 5-day, week-long event – a full immersion week of bioanalysis, biomarkers, immunogenicity and gene therapy. As usual, it was specifically designed to facilitate sharing, reviewing, discussing and agreeing on approaches to address the most current issues of interest including both small- and large-molecule bioanalysis involving LCMS, hybrid LBA/LCMS, LBA cell-based/flow cytometry assays and qPCR approaches. This 2019 White Paper encompasses recommendations emerging from the extensive discussions held during the workshop and is aimed to provide the bioanalytical community with key information and practical solutions on topics and issues addressed, in an effort to enable advances in scientific excellence, improved quality and better regulatory compliance. Due to its length, the 2019 edition of this comprehensive White Paper has been divided into three parts for editorial reasons. This publication (Part 3) covers New Insights in Biomarker Assay Validation, Current & Effective Strategies for Critical Reagent Management, Flow Cytometry Validation in Drug Discovery & Development & CLSI H62, Interpretation of the 2019 FDA Immunogenicity Guidance and Gene Therapy Bioanalytical Challenges. Part 1 (Innovation in Small Molecules and Oligonucleotides & Mass Spectrometry Method Development Strategies for Large Molecule Bioanalysis) and Part 2 (Recommendations on the 2018 FDA BMV Guidance, 2019 ICH M10 BMV Draft Guideline and regulatory agencies' input on bioanalysis, biomarkers, immunogenicity and gene therapy) are published in volume 11 of *Bioanalysis*, issues 22 and 23 (2019), respectively.

First draft submitted: 16 October 2019; Accepted for publication: 30 October 2019; Published online: 10 December 2019

Key terms

Biomarker: A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. Molecular, histologic, radiographic, or physiologic characteristics are types of biomarkers. A biomarker is not an assessment of how an individual feels, functions, or survives. Categories of biomarkers include: susceptibility/risk biomarker, diagnostic biomarker, monitoring biomarker, prognostic biomarker, predictive biomarker, pharmacodynamic/response biomarker, and safety biomarker [26].

Qualification: A conclusion based on a formal regulatory process, that within the stated context of use, a medical product development tool can be relied upon to have a specific interpretation and application in medical product development and regulatory review [26].

Immunogenicity: The ability of a substance, including a biotherapeutic, to elicit an immune response *in vivo* which results in an induction of anti-drug antibodies, antigen specific T cells – among others. For the purposes of this article, ADA assays are equivalent to immunogenicity assays.

Cut point: The cut point of the assay is the level of response of the assay that defines the sample response as positive or negative [53].

Context of use: A statement that fully and clearly describes the way the medical product development tool is to be used and the medical product development-related purpose of the use [26].

Index

Acronyms	
AAV:	Adeno-associated virus
ADA:	Anti-drug antibody
ASO:	Antisense oligonucleotide
ASR:	Analyte specific reagents
BA:	Bioavailability
BAV:	Biomarker assay validation
BE:	Bioequivalence
BEAD:	Biotin-drug extraction and acid dissociation
BLA:	Biologics license application
BMV:	Bioanalytical method validation
CAR-T:	Chimeric antigen receptor T cell
CCP:	Confirmatory cut point
CDx:	Companion diagnostics
CLIA:	Clinical laboratory improvement amendments
CLSI:	Clinical Laboratory Standards Institute
COU:	Context of use
CPF:	Concentration, purity and functionality (new acronym in the context of reagent characterization)
CRISPR:	Clustered regularly interspaced short palindromic repeats
CRO:	Contract Research Organization
CSF:	Cerebrospinal fluid
DBS:	Dried blood spots
ECD:	Extracellular domain
ELISPot:	Enzyme-linked immunospot
F/P Ratio:	Fluorochrome to protein ratio
FFP:	Fit-for-purpose
FMO:	Fluorescence minus one
FPR:	False-positive rate
GCLP:	Good Clinical Laboratory Practices
GCP:	Good Clinical Practice
GLP:	Good Laboratory Practice
GMP:	Good Manufacturing Practice

GxP: GTx:	Good Practiceswhere x = clinical, laboratory or manufacturing Gene therapeutics
HDR:	Homologous directed repair
HRMS:	High-resolution mass spectrometry
HSV:	Herpes simplex virus
IA:	Immunoaffinity
IDE:	Investigational device exemption
IHC:	Immunohistochemistry
IND:	Investigational new drug
Indel:	Insertion/deletion
IQR:	Inter-quartile range
ISR:	Incurred sample reproducibility
IVDR:	*In vitro* diagnostic medical device
LBA:	Ligand-binding assay
LCM:	Life cycle management
LCMS:	Liquid chromatography mass spectrometry
LLOQ:	Lower limit of quantitation
LM:	Large molecule
LTS:	Long-term stability
mAb:	Monoclonal antibody
MESF:	Molecules of equivalent soluble fluorochrome
MFI:	Mean fluorescence intensity
MIQE:	Minimum information for publication of quantitative real-time PCR experiments
MRD:	Minimum required dilution
MS:	Mass spectrometry
NAb:	Neutralizing antibody
NHEJ:	Non-homologous end-joining
NIST:	National Institute of Standards and Technology
pAb:	Polyclonal antibody
PBMC:	Peripheral blood mononuclear cells
PBS:	Phosphate-buffered saline
PC:	Positive control (used in an immunogenicity assay)
PD:	Pharmacodynamics
PK:	Pharmacokinetics
QA:	Quality assurance
QC:	Quality control
QP:	Qualification plan
qPCR:	Quantitative PCR
RCL:	Replication competent virus
RNP:	Ribonucleoprotein
RT:	Reverse transcriptase
RUO:	Research use only
S/N:	Signal-to-noise
SCP:	Screening cut point
sgRNA:	Single guide RNA
SIL-IS:	Stable isotope label internal standard

SOP:	Standard operatingprocedure
tAb:	Total antibody
ULOQ:	Upper limit of quantitation
WRIB:	Workshop on Recent Issues in Bioanalysis

Introduction

The 13th edition of the Workshop on Recent Issues in Bioanalysis (13th WRIB) was held in New Orleans, LA, USA, on 1–5 April 2019 with an attendance of over 1000 representatives from pharmaceutical/biopharmaceutical companies, biotechnology companies, contract research organizations, and regulatory agencies worldwide. The workshop included three sequential main workshop days, six additional full-day training sessions that together spanned an entire week in order to allow exhaustive and thorough coverage of all major issues in bioanalysis, biomarkers, immunogenicity and gene therapy.

As in previous years, this year's WRIB continued to gather a wide diversity of international industry opinion leaders and regulatory authority experts working on both small and large molecules to facilitate sharing and discussions focused on improving quality, increasing regulatory compliance and achieving scientific excellence on bioanalytical issues.

The active contributing chairs included Dr Christine Fandozzi (Merck & Co., Inc.), Dr Christopher Evans (GlaxoSmithKline), Dr Brian Booth (US FDA), Dr Renuka Pillutla (Bristol-Myers Squibb), Dr Fabio Garofolo (Angelini Pharma), Dr Becky Schweighardt (BioMarin), Dr Meina Liang (AstraZeneca), and Dr Lauren Stevenson (Biogen).

The participation of regulatory agency representatives continued to grow at WRIB [1–21] including the below:

- **Regulated Bioanalysis:** Dr Sean Kassim (US FDA), Dr Sam Haidar (US FDA), Dr Seongeun (Julia) Cho (US FDA), Dr John Kadavil (US FDA), Dr Arindam Dasgupta (US FDA), Dr Brian Booth (US FDA), Dr Sriram Subramaniam (US FDA), Dr Theingi Thway (US FDA), Dr Nilufar Tampal (US FDA), Dr Jan Welink (EU EMA), Dr Olivier Le Blaye (France ANSM), Mr Stephen Vinter (UK MHRA), Ms Emma Whale (UK MHRA), Dr Anna Edmison (Health Canada), Dr Catherine Soo (Health Canada), Mr Gustavo Mendes Lima Santos (Brazil ANVISA), Ms Thais Correa Rocha (Brazil ANVISA);
- **Biomarkers:** Dr Yow-Ming Wang (US FDA), Dr Abbas Bandukwala (US FDA), Dr Kevin Maher (US FDA), Dr Shashi Amur (US FDA), Dr Shirley Hopper (UK MHRA), Dr Yoshiro Saito (Japan MHLW-NIHS);
- **Immunogenicity:** Dr João Pedras-Vasconcelos (US FDA), Dr Haoheng Yan (US FDA), Dr Susan Kirshner (US FDA; remote), Dr Daniela Verthelyi (US FDA; remote), Dr Elana Cherry (Health Canada), Dr Akiko Ishii-Watabe (Japan MHLW-NIHS), Dr Venke Skibeli (Norway NoMA), Dr Therese Solstad Saunders (Norway NoMA);
- **Gene Therapy:** Dr Nirjal Bhattarai (US FDA), Dr Heba Degheidy (US FDA).

The 13th WRIB was designed to cover a wide range of topics in bioanalysis, biomarkers, immunogenicity and a special full-day session dedicated to gene therapy bioanalytical challenges. Moreover, the 13th WRIB included daily working dinners and lectures from both industry experts and regulatory representatives, which culminated in open panel discussions amongst the presenters, regulators and attendees in order to reach consensus on items presented in this White Paper.

While the 13th WRIB continued its traditional emphasis on method development challenges and novel solutions in bioanalysis, it also included an in-depth focus on the recently released ICH M10 BMV Draft Guideline [22]. Three full sessions, two working dinners and three open forums were dedicated to cover the hot topics of the ICH M10 BMV Draft Guideline, and to actively interact with the regulators' expert panel and work together as a Global Bioanalytical Community with the goal to provide official comments on the ICH M10 BMV draft guideline. Harmonized topics among US FDA, EU EMA, Health Canada, Japan MHLW and Brazil ANVISA regulations, unresolved issues and on-going industry/regulator discussions were thoroughly evaluated to support the regulatory recommendations of the ICH M10 BMV Draft Guideline which, when finalized, will supersede the regional guidance of the participating health authorities.

The three sessions and open forums on ICH M10 activities at WRIB were organized and coordinated by Dr Brian Booth (US FDA, ICH M10 EWG Regulatory Chair), Dr Jan Welink (EU EMA), Dr Anna Edmison (Health

Canada), Dr Akiko Ishii-Watabe (MHLW, ICH M10 EWG Rapporteur), Dr Yoshiro Saito (MHLW), and Ms Thais Correa Rocha (ANVISA) and with input and active participation of numerous industry/regulator opinion leaders.

A total of 48 recent issues ('hot' topics) were addressed and distilled into a series of relevant recommendations. Presented in the current White Paper is the background on each issue, exchanges, consensus and the resulting recommendations on these 48 topics.

Due to its length, the 2019 edition of this comprehensive White Paper has been divided into three parts for editorial reasons. This publication covers Part 3 recommendations.

Part 1 – *Bioanalysis* Volume 11, Issue 22 (November 2019)

Innovation in Small Molecules and Oligonucleotides:

- Novel Therapeutic Modalities (two topics);
- Innovation in Small Molecules (three topics);
- Small Molecule Biomarkers by LCMS (one topic);
- Oligonucleotides (one topic).

Mass Spectrometry Method Development Strategies for Large Molecules Bioanalysis:

- Innovation in Hybrid LBA/LCMS Assays (five topics);
- Biomarker Assays (three topics).

Part 2 – *Bioanalysis* Volume 11, Issue 23 (December 2019)

Implementation of 2018 FDA BMV Guidance (ten topics):

- Industry/Regulators' Feedback on ICH M10 BMV Draft Guidelines (14 topics);
- Input from Regulatory Agencies on Bioanalysis & BMV;
- Input from Regulatory Agencies on Immunogenicity & Biomarkers.

Part 3 – *Bioanalysis* Volume 11, Issue 24 (December 2019)

New Insights in Biomarker Assay Validation (BAV):

- Fit-for-Purpose and Context of Use (one topic);
- Free Assays (one topic);
- BAV Guidelines (three topics).

Current & Effective Strategies for Critical Reagent Management:

- Characterization and Stability (three topics);
- Life Cycle Management (one topic);
- Flow Cytometry (one topic);
- Challenges and Approaches (one topic).

Flow Cytometry Validation in Drug Discovery & Development & CLSI H62:

- Regulatory Expectations and Validation (two topics);
- Challenges and Approaches (three topics);
- Data Analysis (two topics).

Interpretation of the 2019 FDA Immunogenicity Guidance:

- Drug Tolerance (one topic);
- Critical Reagent and Positive Control Characterization (one topic);
- LCM and Clinical Relevance of ADA (one topic);
- Challenges and Approaches (three topics);

- Cut Points and the FDA Immunogenicity Guidance (one topic).

Gene Therapy Bioanalytical Challenges:

- Approaches to Gene Therapy Bioanalysis (four topics);
- Vaccines (one topic);
- Challenges (two topics);
- Immunogenicity (one topic).

SECTION 1 – New Insights in Biomarker Assay Validation (BAV)

Steven Piccoli[1], Devangi Mehta[2], Alessandra Vitaliti[3], John Allinson[4], Shashi Amur[5], Steve Eck[6], Cherie Green[7], Michael Hedrick[8], Shirley Hopper[9], Allena Ji[10], Alison Joyce[11], Virginia Litwin[12], Kevin Maher[5], Joel Mathews[13], Kun Peng[7], Afshin Safavi[14], Yow-Ming Wang[5] & Yan Zhang[8]

Authors in section 1 are presented in alphabetical order of their last name, with the exception of the first three authors who were session chairs, working dinner facilitators, and/or notetakers. Author affiliations can be found at the beginning of the article.

Discussion Topics & Consolidated Questions Collected from the Global Bioanalytical Community

The following paragraphs report the consolidated questions collected from the Global Bioanalytical Community. Four discussion topics were extracted from these questions and considered as the most relevant 'hot topics'. They were reviewed by internationally recognized opinion leaders before being submitted for discussion during the 13th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Fit for Purpose & Context of Use

Fit for Purpose & Context of Use

Is BAV always FFP based on the COU? Can 'full' validation be prescribed for biomarker assays, regardless of whether it is an exploratory marker for drug development or a qualified biomarker, or must all biomarker assays be FFP to meet COU? What is 'full' validation – is it constant or does it change depending on the COU? What does 'full' validation mean for biomarkers? Is there agreement that FFP biomarker assay validation is not the 'easy way out', or is it scientifically driven by each COU and may require meeting criteria that are more (or less) rigorous than for a PK assay? How are the FFP criteria determined for allowable assay variability (total error, imprecision, and bias)? What studies are performed to evaluate intra-subject and inter-subject variability? If biomarker assay performance requirements should be driven by scientific rationale, is there agreement that given the FFP nature driven by COU, would it be difficult to capture singular requirements in a guidance document?

BAV Guidelines

Accuracy

Considering that "*accuracy is one of the utmost fundamental requirements for validation of any assay including fit-for-purpose biomarker assays*" [21], is there agreement that most biomarker assays have relative accuracy? What requests have come from the regulatory agencies? The 2018 White Paper in Bioanalysis Part 3 stated: "*as flow cytometry assays lack the availability of reference standards and the data generally fall into the category of quasi-quantitative, it is thus not possible to validate accuracy in the traditional manner*" [21]. What can be done to satisfy the request to assess accuracy? How can absolute quantitation be brought to quasi-quantitative techniques? What is done in industry laboratories for both exploratory and regulated biomarkers?

Parallelism

Is parallelism the key experiment to demonstrate that the method is FFP to measure the endogenous analyte? Is there alignment that spike/recovery experiments of recombinant material in the matrix are not a reflection of the ability of the assay to measure endogenous analyte? What has been added to the revised C-Path White Paper [23] regarding parallelism? How does the parallelism evaluation have a direct impact on the determination of MRD and sensitivity? How is the MRD calculated based on parallelism?

BAV & Regulations

What are the commonalities and differences between the approaches to BAV of FDA, EMA, MHLW and consensus White Papers [12,15,18,21,23]? How can the industry support and encourage regulatory harmonization? Should industry drive the conversation to a single set of scientific practices which will satisfy all? Should all biomarkers be treated in a technologically agnostic fashion for BAV? How are new technologies to be integrated into existing practices and regulations? Do we need better (or at least some) definitions (i.e., regulatory clarification) for the differences in BAV and data generation for confirmatory, clinical (CLIA/CAP/CE), exploratory, primary and secondary endpoints? In which clinical phase should a newly developed biomarker test be sent to a CLIA lab instead of conducted in a GCLP lab? In what situation should a qualified target biomarker test used to assess disease stage or inclusion/exclusion criteria for a clinical trial enrollment be sent to a CLIA lab? When is it appropriate to use a correction factor for BAV? Has the opinion on this evolved over time and in light of new White Papers and regulations [23–25]? What is the current thinking on ISR for biomarker assays based on the 2018 FDA BMV Guidance [25]? What are the current practices on assessing long-term endogenous QC stability in light of new White Papers and regulations?

Discussions, Consensus & Conclusions

Fit for Purpose & Context of Use

Fit for Purpose & Context of Use

Discussions on this topic began by obtaining consensus on the FFP and COU nomenclature and its application in BAV. In the realm of biomarker assays, the FFP approach is the equivalent to analytically validating and characterizing the assay for the intended COU. Context of use is defined as "*a statement that fully and clearly describes the way the medical product development tool is to be used and the medical product development-related purpose of the use*" [26]. In other words, the COU defines the 'P' or 'purpose' in FFP, specifically how the biomarker data will inform the scientific question and decision-making for the study, drug program, or patient. If the scientist does not have a clear understanding of the intended COU, the assay cannot be appropriately validated for its intended purpose. Importantly, the FFP approach to biomarker assay validation should not be viewed as the 'easy way out', rather it is the scientifically-driven approach to assay validation. The FFP approach to biomarkers is often iterative, where the biomarker assay and validation may need to be refined as one gains new knowledge about the biomarker or the COU evolves.

Given that each individual COU drives the FFP BAV, the acceptance criteria and performance expectations for any assay cannot be prescribed *a priori*. Rather, the analytical error and biological variability in the measurements should be determined and related to the desired clinical validation (i.e., COU) to set appropriate analytical validation acceptance criteria [23]. Critical BAV parameters for a quantitative or relative quantitative FFP assay typically include (relative) accuracy, precision, analytical measurement range, parallelism, specificity, selectivity, and sample stability. Assessments should be based on the endogenous analyte. In general, these BAV parameters are agnostic of the biomarker assay technology, but how each parameter is defined and assessed may vary based on the technology platform and should be scientifically justified. Also, there may be additional validation parameters necessary based on the specific technology. The concept and proper implementation of FFP has been thoroughly summarized by Lee *et al.* [27] and expanded upon in the C-Path White Paper [23] for single-plex ligand and immuno-binding assays, mass spectrometry, and enzyme-based assays. Regardless of whether the biomarker is exploratory (e.g., utilized for internal decision-making) or a regulatory endpoint, the FFP and COU concepts can be universally applied to ensure a biomarker assay validation that is scientifically defensible.

BAV Guidelines

Accuracy

Due to the general lack of certified reference material, there are few biomarker assays that are considered absolute quantitative. Thus, it is well-recognized by the biomarker community that in such cases analytical accuracy can only be described by relative accuracy or bias. Additionally, some quasi-quantitative assays (e.g., flow cytometry, IHC) do not utilize a calibration curve, but report a continuous numerical response, and relative accuracy/bias is rather a reflection of the specificity of the method. It was agreed that finding suitable and meaningful alternative approaches to establishing accuracy in such cases poses a tremendous challenge. Accepted alternative approaches for establishing accuracy may include proficiency testing with survey material, comparison to a reference methodology,

or verification with specimens obtained from patients with a diagnosis confirmed by orthogonal methods. While these approaches are often applied in clinical/diagnostic laboratories, they may not be possible when validating novel biomarker methods. Critically, regulatory agencies will expect that accuracy is scientifically addressed during method validation. Hence, relative accuracy (bias) of the assay should be understood to properly set acceptance criteria based on the COU.

Parallelism

The purpose of parallelism is to assess the relationship between the sample-dilution and standard-calibrator response curves and determine if the calibrator material and surrogate matrix are suitable for quantifying the endogenous analyte. While sometimes sourcing the appropriate samples with the endogenous biomarker can pose challenges, parallelism is an essential experiment required to appropriately develop and characterize all ligand-binding (antibody-dependent) biomarker assays. Parallelism assessments inform on multiple parameters including surrogate matrix selection, MRD optimization, selectivity, and estimation of assay sensitivity of the endogenous analyte in intended matrix [28]. Parallelism assessments cannot be replaced by spike/recovery experiments with recombinant material. If parallelism cannot be performed pre-study, it should be performed in-study as soon as samples become available to understand how the endogenous analyte behaves in the assay and whether additional assay optimization is required to support the COU. The approach for quantitative assessment of parallelism discussed in the 2014 articles and White Papers commonly used by the industry is deemed still valid and highly recommended [9,12,23,28].

BAV & Regulations

In the regulatory arena, there continues to be open discussions on the need for BAV guidance, however to date there is no commonality in approaches between the various regulatory bodies as global standards have not yet been developed. Currently, only FDA's BMV [25] discusses BAV, but it is not comprehensive, and other regulatory agencies have not issued BAV-specific guidance. In addition, FDA's BMV focuses primarily on the assay requirements needed for PK assessment, which may not be applicable to the biomarker assays. ISR, which is required for PK assays, may not be appropriate for biomarker assays as endogenous QCs (incurred samples) from the relevant sample matrices can be used to monitor assay performance during sample analysis. Given the breadth and complexity of biomarkers and the technology platforms utilized, a single guidance is unlikely to cover all potential scenarios. Thus, the C-Path White Paper [23] is the first formal attempt to drive harmonization, particularly around the concept of a COU-driven biomarker assay validation and alignment on the parameters that should be assessed in BAV, even though the White Paper addresses analytical validation of assays for the purpose of biomarker qualification. It is acknowledged that each technology will have its own considerations, but the core principles can be applied to all, although how those parameters are executed may vary between technologies. Consequently, there is no *a priori* guidance for BAV acceptance criteria, as this should be defined by the intended COU and by determining the total allowable error to discriminate a drug induced or clinically relevant difference in the biomarker. Discussion indicated that there is the general expectation that FDA will publish a BAV-specific guidance, likely leveraging the C-Path White Paper as a resource, and other agencies should take a similar approach to harmonize their guidance documents. It was agreed that BAV guidance needs to be continually harmonized between industry and regulators in a data-driven manner to evolve with best scientific practices. As part of the approach to continuous scientific discourse on BAV, industry and regulators should continue to share data at WRIB and other conferences to understand what works and what needs improvement for the next step in harmonization.

As biomarkers are substantially shaping drug development practices, there was significant discussion on aligning best regulatory practices for biomarker testing in therapeutic clinical development. GLP regulations were developed to assure the quality of data generated for toxicology and safety pharmacology studies in animals and so does not apply to most of the exploratory work done in animal pharmacology laboratories. Analysis of human specimens does not fall under GLP regulations either and should be conducted following the concept of GCLP [29], which will add an element of quality to work performed on clinical specimens.

In the USA, CLIA regulations were established in 1988 [80,81]. The CLIA regulations requires that clinical laboratories are certified by their state and well as the Centers for Medicare & Medicaid Services (CMS). The purpose of CLIA regulations is to ensure the quality of assay work performed in "*any facility which performs laboratory testing on specimens derived from humans for the purpose of providing information for the diagnosis, prevention, treatment of*

disease, or impairment of, or assessment of health." CLIA regulations do not apply to animal safety studies but do apply to all clinical laboratory testing (e.g., diagnostic testing to screen for or monitor specific diseases or conditions). They may not even apply to clinical biomarkers when the tests are for research use in drug development. There are many differences between GLP, GCLP and CLIA regulations [30], causing challenges when deciding which regulations to apply to a study.

Biomarkers employed in drug development typically fall into two main categories: 1) internal decision-making and 2) patient care decision-making. In clinical trials, the terminology for primary, secondary, and exploratory biomarker endpoints have no distinction in regulatory compliance – they are instead definitions of what is critical to the sponsor in the clinical study design. In this sense internal decision-making biomarkers are all exploratory. Biomarker analyses for primary and secondary endpoints are regulated by GxP guidelines and subject to BAV review, with primary and secondary endpoints generally anticipated to be held to more stringent standards than exploratory endpoints. In the US, if biomarkers are intended to be used for individual patient treatment/medical decisions (i.e., reporting a patient-specific test result as opposed to aggregate data reporting), biomarkers likely must be tested under CLIA. When such results from US clinical trials are used to make individual patient-treatment decisions, IDE regulations may apply. This includes biomarkers used for clinical trial enrollment criteria or individual dose selection and an appropriate development course should be reviewed with the regulatory agency.

Recommendations

Below is a summary of the recommendations made during the 13th WRIB:

1. **FFP BAV** analytically validates and characterizes the assay for the intended **COU**:
   - The analytical error and biological variability in the measurements should be determined and related to the desired COU to set appropriate analytical validation acceptance based on the endogenous analyte: (relative) accuracy, precision, analytical measurement range, parallelism, specificity, selectivity, and stability.
2. The majority of biomarker assays are not absolutely quantitative; thus, analytical accuracy can only be described by **relative accuracy**, unless a certified reference standard is used:
   - Relative accuracy (bias) of the appropriate sample should be understood to properly set acceptance criteria based on the COU;
   - For quasi-quantitative and qualitative assays (flow cytometry, IHC, etc.), the continuous numerical response and relative accuracy/bias is a reflection of the specificity of the method, which may be addressed by alternative approaches based on scientific rationale.
3. **Parallelism** is an essential experiment to appropriately characterize the endogenous analyte in all ligand-binding (antibody dependent) biomarker assays:
   - Parallelism cannot be replaced by spike/recovery experiments with recombinant material;
   - If parallelism cannot be performed pre-study, it should be performed in-study as soon as samples become available to understand if the assay is appropriate for the COU or requires additional optimization.
4. **C-Path White Paper** was the first formal attempt to drive harmonization on **BAV** as global regulatory standards for BAV do not exist:
   - While each technology will have its own considerations, the core/general principles outlined above can be applied [23], but how those parameters are validated may vary between technologies;
   - BAV needs to be continually harmonized between industry and regulators in a data-driven manner. In the US, if individual patient treatment/medical decisions are intended, biomarkers must be tested under CLIA, and IDE regulations may apply. This includes biomarkers used for clinical trial enrollment criteria and patient stratification, but the approach should be reviewed with regulatory agencies;
   - BAV for primary and secondary biomarker endpoints have no regulatory definition but are subjected to GxP guidelines – they are defined by what is critical to the sponsor.

SECTION 2 – Current & Effective Strategies for Critical Reagent Management

Lakshmi Amaravadi[15], Nisha Palackal[22], Sai Thankamony[8], Chris Beaver[16], Eris Bame[2], Thomas Emrich[17], Christine Grimaldi[18], Jonathan Haulenbeek[8], Alison Joyce[11], Vellalore Kakkanaiah[19], David Lanham[20], Kevin Maher[5], Andrew Mayer[21], Paul C Trampont[23] & Laurent Vermet[24]

Authors in Section 2 are presented in alphabetical order of their last name, with the exception of the first four authors who were session chairs, working dinner facilitators, major contributors and/or notetakers. Author affiliations can be found at the beginning of the article.

Discussion Topics & Consolidated Questions Collected From The Global Bioanalytical Community

The following paragraphs report the consolidated questions collected from the Global Bioanalytical Community. Six discussion topics were extracted from these questions and considered as the most relevant 'hot topics'. They were reviewed by internationally recognized opinion leaders before being submitted for discussion during the 13th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Characterization & Stability

Long-Term Stability

How extensively should stability of critical reagents used for LBA be characterized? Is it justified to use historical knowledge and experience of reagent performance to put these in high risk vs low risk categories? What characterization methods should be used to test stability? Is the performance in a functional assay always practical, and is it enough? Drug stability is extensively tested at the expected storage temperatures for pre-clinical and clinical use. Can this stability experience be extrapolated for modified drug proteins, for example, biotin or ruthenium conjugated drugs? What new reagent modalities should be considered as stability risks that may warrant more frequent testing? For example, do engineered proteins (non-traditional mAb, but 'mAb-like' proteins) have higher risks for instability? Should accelerated stability studies be carried out at increased temperatures for critical reagents? Is this applicable to long-term storage at lower temperatures?

Re-Testing

How (in-assay performance vs orthogonal testing) and at what interval should critical reagents be re-tested? What is the simplest and most robust way to purify reagents for long-term use? How are the expiration or re-test dates for labeling determined?

Characterization

What are the recommended minimum best practices for reagent characterization? Is there agreement that a new lot of critical reagent is acceptable as long as there is acceptable assay performance and performance that is continuously monitored (i.e., if established accuracy and precision is met, additional characterization is not necessary)? Are there recommendations for the use of mass spectrometry approaches to characterize critical reagents?

Life Cycle Management

Life Cycle Management of Critical Reagents

What are best practices for generating assay specific reagents? What is the best timing in the overall development lifecycle? How is it best to screen for the desired reagent properties during the antibody generation phase? What is necessary for characterization of reagents to set a baseline for lot-to-lot reproducibility? How is orthogonal characterization data used in the life-cycle of the assay? What are good scientific practices vs regulatory requirements/expectations? What types of bridging studies are performed to ensure consistency in assay performance when transitioning from one reagent to another?

Flow Cytometry

Critical Reagents in Flow Cytometry

Use of commercial reagents is common in flow cytometry, however it is often difficult to get QC/QA data behind the very lightweight certificates of analysis. What information should we seek from vendors as i) the minimum requirement and ii) the gold standard? What in-house efforts to assess reagents from commercial suppliers are considered appropriate over and above empirical analysis using a QC sample test? What are the best practices for switching to a new reagent (e.g., switching the fluorochrome of a reagent using the same clone, a new clone or a new source) in the middle of a study due to unavailability of the original reagent from the vendor, or when encountering any other method performance issues? What is the best industry practice for characterization of cell types: cell subsets; heterogeneity of blood dendritic cells depending on their origins; differentiation of hematopoietic stem cells? What actions should be taken if a new lot of critical reagent shows a consistent but different result to the original reagent lot – among others, a reagent used to measure a MFI endpoint is consistently 30% higher in

signal response when tested using QC materials, but the reagent shows specificity and the proportions of cells are comparable to the initial lot?

Challenges & Approaches

Challenges & Approaches Using Critical Reagents

What are minimum requirements for the description of critical reagents (i.e., expression system, glycosylation state and sequence)? These are critical to epitope recognition, yet very little if any information is typically given directly in methods. What are minimum recommendations for characterizing reagents and generation of certificates of analysis? What is the value, or lack thereof, for acceptance criteria of critical reagents? Are there universal buffers or excipients one can use to extend the stability of reagent antibodies? Are there recommendations on formulation buffers for conjugated reagents? Are there technical challenges in generating mAbs and pAbs? What percentage of antibodies used as positive controls are polyclonals versus monoclonals? What types of issues are seen when switching from polyclonals to monoclonals? What types of issues are seen when switching from hybridoma-derived to recombinant engineered monoclonals? What can be learned from industrial inventory management or lean management? What should a critical reagent maintenance program contain? What are the unique issues related to reagents in PK assays: consider new generation of therapeutic modalities – among others, bi- and tri-specific antibodies, gene therapy modalities, RNA and ASO therapies?

Discussions, Consensus & Conclusions

Characterization & Stability

Well characterized critical reagents are essential building blocks for high performing bioanalytical assays that utilize reagents, irrespective of assay platforms. Due to recent advances in novel platforms for antibody-, cell- and gene-based therapies, the reagents used for these biologic modalities as well as the assays themselves are increasing in complexity. Thus, a tailored approach for critical reagent life cycle management is important for assay ruggedness and robustness.

A sustainable and reliable supply of reagents is a key element to attain acceptable assay performance over the lifetime of a drug program, which may be many years to a decade or more. Thus, during the assay life-cycle multiple reagent lots may be utilized, where inconsistencies and minor differences in these protein reagents can impact assay performance. To mitigate lot-to-lot variation, large lots may be produced, characterized and stored preferentially in single use aliquots to extend the shelf-life of the reagent and preserve function, but this carries a potential long-term stability risk. When available and suitable to a particular application, reagents that are labeled as ASR should be considered. ASR reagents are manufactured under GMP regulations and may provide materials with less inter-lot variability than similar reagents labeled as RUO, which may be manufactured under less stringent conditions. Published White Papers in recent years have recommendations for expiry ranges for purified monoclonal and polyclonal antibody reagents (labeled or unlabeled), Fc fusion proteins, recombinant proteins (e.g., targets) and commercial reagents and these recommendations work well in most cases [31,32]. However, new biotherapeutic modalities have necessitated the use of novel reagent types and antibody or target-coupled beads where little to no historical data or guidance exists to substantiate expiry recommendations. LTS testing for novel and certain traditional reagent types may be warranted until analytical (biophysical properties) and functional (specificity, assay performance) data show a novel reagent is stable and fit for use in the assay over the long term. In addition, a preventative approach may also be implemented to produce robust reagents from their inception by using optimized coupling procedures, storage buffers [33] or alternative reagent formats [34].

With the use of chromatography methods, purity of critical reagents should remain constant from the time reagents are first generated and throughout the life cycle of the LBA. While maintaining a high level of reagent monomer content is important for assay robustness, other biophysical parameters can impact assay performance. It has been observed that even highly monomeric conjugated reagents can have impaired functionality, despite their high level of purity content by size exclusion chromatography. This may be due to the inability of standard purity assessments to detect insoluble aggregates, subvisible and visible particles. In addition, purity does not provide information on the potential level of unfolding of the reagent that could impact functionality prior to aggregation. Methods such as static or dynamic light scattering as well as full spectrum fluorescence may help to further understand additional biophysical changes in critical reagents. Especially for labeled assay reagents, this may also be due to the limited capability of standard purity methods to resolve differences in the labeling position and labeling grade. Conjugation of novel multi-specific molecules poses an additional challenge in terms

of stability and handling. Buffer exchange of conjugated reagents into formulation buffers with cryoprotectants has addressed previously encountered performance issues for some assays, while in other cases has led to impaired assay performance over time. It was agreed that an empirical approach should be employed to test stability of conjugated reagents under multiple cryoprotectants and formulation buffers to ensure long term reagent ruggedness and robustness. Similarly, optimization of desalting conditions, as part of the reagent conjugation process and handling, can be a critical step for obtaining optimal reagent performance in LBA.

Long-Term Stability

The extent of characterization of stability of critical reagents used for bioanalytical assays is driven by the context of use of the reagent. Critical reagents should be identified in the method and may be considered high or low risk depending on the type of molecule and any modifications or labeling constructs. Historical data and trending records can provide insight into the extent of stability testing required for a particular reagent. Reagents that may require more frequent retesting include soluble receptors, receptor ECDs, nanobodies, bispecifics, affimers, and aptamers.

Recommended characterization methods included biophysical characterization, for example, size exclusion chromatography, and potency/functionality by orthogonal assays such as surface plasmon resonance to understand binding characteristics. Accelerated stability/stress tests can be used to rule out poor quality reagents at the outset or for storage buffer optimization. However, it is not recommended to extrapolate this approach to long-term stability unless a robust stress model that allows extrapolation has been established.

Drug stability is extensively tested for the expected storage temperatures during clinical use. Consensus was that it is not recommended to extrapolate this stability experience for modified drug proteins (e.g., biotin or ruthenium conjugated drug proteins) because the labeling can alter the isoelectric point of the molecule, and based on that the formulation may change, resulting in a drug-derived reagent with an altered performance in the assay.

Re-Testing

There was an overwhelming agreement among attendees that 'expiry' and 're-test' dates should be treated independently and differently. The expiration of a compound or reagent should be based on stability data demonstrating that the functionality is impaired after a certain amount of time under certain conditions. Therefore, the use of an expiration date should be limited to cases where there is adequate data to demonstrate a loss of stability. The expiration dates for reagents labeled in-house, or outsourced for labeling, should be determined based on their performance in assays.

Consensus was that it is preferable to use the term re-evaluation or retesting; this process being driven with an appropriate SOP. When determining the retest frequency, there is no one rule for all critical reagents. The frequency will depend on the type of reagent (e.g., mAbs: every 2–10 years, more labile proteins/peptides: as often as 6 months or based on perceived risk) [31,35].

Characterization

Best practices for reagent characterization recommend a minimum of a **C**oncentration assessment, a **P**urity determination, and the determination of **F**unctionality by the bioanalytical assay and by an orthogonal method (e.g., one that utilizes a different assay principle from the bioanalytical assay; CPF). The acronym CPF was created by the expert panel to make it easier to remember the 'must assess' parameters. It is also important to understand incorporation ratios, presence of unlabeled protein, aggregates, and how they impact the assays. It was agreed that even if not considered as the first choice for characterization, the bioanalytical assay is the best indicator of reagent functionality.

Multiple tools and techniques should be considered for the characterization of critical reagents based on need/context. Some suggestions included size exclusion chromatography, Octet, Biacore, SDS-Page, Nanodrop, and LCMS.

It was agreed that there is no universal buffer, like PBS, for reagents, given that each critical reagent's unique characteristics should be considered when determining appropriate buffer solutions. For labeled drugs, formulation buffer is a good choice for storage after labeling; adding sugars and anti-microbial agent (e.g., azide or ProClin™) is also recommended.

Life Cycle Management

Life Cycle Management of Critical Reagents

The generation, sourcing and life cycle management of high-quality critical reagents are fundamental for the development and validation of robust and rugged analytical methods for long-term biotherapeutic support throughout the drug development process. Ideally, high quality, reproducible and sustainable reagents should be generated early in the biotherapeutic development lifecycle. The lack of such reagents can result in delays to method development and validation (e.g., require assay re-optimization, re-validation and cross validation as well as potentially affect the translatability of data across studies or phases of development). While adequate reagent characterization is important, the initial focus should be on generating reagents that meet the long-term needs of any given project. Important reagent characteristics may guarantee the desired functionality, including affinity and specificity when appropriate, as poor reagent selection will result in assays that are potentially unsuitable and unsustainable. Best practices include a lean process that is fit-for-purpose, avoids waste, fully understands the needs of the customer/assay scientist, and has transparency. Reagents may be tested using a crude version of the assay to obtain information about the reagent characteristics that are most important for optimal assay performance. This will then inform how reagents are managed. It is best practice to generate monoclonal cell-line derived reagents as early as possible in the process and then consider moving to a recombinant antibody early in the lifecycle, particularly if the clone is a poor producer. It is also recommended to have the sequence available and to ensure long-term supply for clinical programs which are lengthy in nature.

It was agreed that orthogonal characterization data may help to better understand important aspects of the reagents that are critical for assay performance, especially when the reagent will be used over the long-term. It can also be used as needed for troubleshooting or to bridge different lots of reagents.

Regarding lot-to-lot bridging for PK assays, consensus was that bridging comparison across the entire assay range should be performed. For ADA assays, it is necessary to evaluate the assay performance around the cut point, sensitivity, and drug tolerance.

Flow Cytometry

Critical Reagents in Flow Cytometry

The increasing importance of biomarker data in modern pharmaceutical drug development has seen a growing use of flow cytometry as a key platform technology; especially in the fields of immuno-oncology and cell-based therapies. Therefore, the control of critical reagents used in flow cytometry is important, especially when an assay is employed across several phases of a drug program, or in complex multi-site clinical programs.

Quality of reagents selected for these assays is critical to assay performance. As stated above, ASR (analyte specific reagents) should be considered when available. Use of commercial reagents is common in flow cytometry, however it is often difficult to get elaborate QC/QA data behind the limited content of the certificates of analysis. For characterization information on commercial flow cytometry reagents, it is recommended to minimally obtain concentration, clonality, F/P ratio and immunogen information from the vendor. If this information is not available from the vendor, it is recommended to assess these in-house by evaluating specificity, characterization of F/P ratio and titration for consistent and optimal performance. The use of FMO gating control and compensation matrices also provide valuable performance data for a reagent or cocktail of reagents.

Though not ideal, there was consensus that bridging assays and a partial validation can be utilized when there is a need to switch a critical reagent in the middle of the study due to unavailability of the original reagent or for any other method performance issues (e.g., switching to a new reagent: new clone, or new supplier, or changing the fluorochrome of a reagent on the same clone). It is recommended to compare the performance of both lots of reagents side-by-side in the assay, using three to six samples. It would also be insightful to compare the F/P ratios between both lots as such could drive a consistent but different result. Lastly, if MFI is the read out, then MESF bead normalization or other similar approach is highly recommended.

Challenges & Approaches

Challenges & Approaches Using Critical Reagents

In order to provide reliable results, critical reagents need to be carefully selected and characterized. This is especially true for binding proteins (e.g., receptors) or target surrogates like anti-idiotypic antibodies. For the analysis of therapeutic proteins, in addition to characterizing the assay reagents themselves, a careful characterization of the

form of the analyte that is detected by LBA-based immunoassays is needed (e.g., whether it is total, target-binding competent or active drug that is detected). Minimally, it is important to understand the immunogen, expression system, glycosylation, and sequence data.

The value of acceptance criteria was discussed and it was agreed that criteria would vary based on a number of factors. It was concluded that scientists must first understand the relationship between the reagent characteristics and their performance in the assay before setting criteria. It may therefore take a long time before such criteria can be set, but trends may become obvious as the critical reagent knowledge database is built.

Many critical reagents are produced by tagging/labeling molecules, for example, with biotin, ruthenium, or other label molecules for use in immunoassays (e.g., ELISA, or ECLIA assays for PK, immunogenicity and biomarker analysis). Control over the production of critical reagents, particularly across multiple lots over many years is crucial to ensuring the consistency of performance throughout the assay lifecycle during a drug development program. Unexpected and unwanted changes over time can affect assay performance requiring time to troubleshoot or revalidate bioanalytical methods and possibly call into question the validity of generated data. It is also important to note that some assays may need to be utilized to support drug products in the post-marketing setting over a long period of time, for example, ADA assays to support patient safety monitoring.

The use of pAbs versus mAbs for positive controls in ADA assays was also discussed. It should be recognized that neither the mAbs positive control nor the pAbs generated for assay development and deployment truly represent the patient ADA response. Polyclonal Abs are generally faster to generate via immunization compared with mAb generation that requires both immunization in mice and hybridoma generation. However, as a system suitability control, mAbs are ideal as the assay can be maintained with consistent performance over a long period of time. Polyclonal Abs may be helpful in characterizing the assay in early stages; a panel of mAbs and pAbs during development can be useful. For consistent assay results, preparation of a large pAb lot from the same purification batch is recommended. However, it is recommended to switch to utilizing mAbs for long-term maintenance of the assay to avoid the issues related to switching between lots of pAbs and limited ability to bridge assays and data. With the development of advanced technologies such as phage display and engineered antibodies, it is also becoming possible to quickly generate mAbs early on in drug development with faster timelines.

Recommendations

Below is a summary of the recommendations made during the 13th WRIB:

1. The extent of characterization of the stability of critical reagents used for bioanalytical assays is driven by the context of use of the reagent and the assay;
2. Historical data, experience and trending charts can provide insight into the stability required for a particular reagent;
3. Characterization and testing in the bioanalytical assay can rule out poor quality reagents at the outset. Accelerated stability tests could give an indication of the long-term stability of the reagent and potential issues with freeze–thaw if included as part of stability testing. However, it is not recommended to extrapolate this approach for long-term stability unless a robust predictive stress model that allows extrapolation has been established;
4. It is not recommended to extrapolate drug stability onto modified drug proteins;
5. The concepts of 'expiry' and 'retest' dates should be treated independently and differently. The expiration of a compound or reagent should be based on stability data;
6. The retest frequency will depend on the type of reagent and executed based on an established SOP;
7. Reagent characterization should include, as a minimum, **C**oncentration assessment, **P**urity determination, and the determination of **F**unctionality by an orthogonal method (**CPF**) that is not the same assay principle as the bioanalytical assay;
8. There is no universal buffer for reagents. PBS should not be considered by default;
9. When generating assay specific reagents by cell culture and cell lines, assess monoclonality as early as possible in the process and then as a safe-guard move to a recombinant antibody early in the lifecycle;
10. It is recommended to have the reagent sequence available to ensure long-term supply for a lengthy clinical program;
11. For PK assays, reagent bridging comparison should evaluate the entire assay range. For ADA assay reagents, it is necessary to evaluate performance around the cut point, sensitivity, and drug tolerance;

12. When selecting critical reagents for flow cytometry, ASRs should be considered when appropriate and available as these are produced under GMP regulations and may provide greater consistency across lots than RUO products, which may not be manufactured under GMP;
13. For commercial flow cytometry reagents, it is recommended to minimally obtain concentration, clonality, F/P ratio and immunogen information from the vendor;
14. For flow cytometry reagents:
    - If characterization needs to be done in-house, it is recommended to assess specificity, characterize F/P ratio, titrate for performance and use FMO gating control and compensation matrix assessments;
    - The best practice when changing any aspect of a reagent in the middle of the study includes performing a partial validation and bridging between assays;
    - If a new lot of critical reagent shows a consistent but different result to the original reagent lot, a comparison of assay performance using three to six samples and both lots is recommended. If MFI is the read-out, then MESF bead normalization is recommended.

SECTION 3 – Flow Cytometry Validation in Drug Discovery & Development & CLSI H62

Alessandra Vitaliti[3], Virginia Litwin[12], Devangi Mehta[2], Eris Bame[2], Naveen Dakappagari[25], Steve Eck[6], Catherine Fleener[26], Fabio Garofolo[27], Cherie Green[7], Michael Hedrick[8], Vellalore Kakkanaiah[19], David Lanham[20], Kevin Maher[5], Steven Piccoli[1], Cynthia Rogers[28], Shabnam Tangri[25], Paul C Trampont[23] & Yuanxin Xu[29]

Authors in Section 3 are presented in alphabetical order of their last name, with the exception of the first four authors who were session chairs, working dinner facilitators, major contributors and/or notetakers. Author affiliations can be found at the beginning of the article.

Discussion Topics & Consolidated Questions Collected from the Global Bioanalytical Community

The following paragraphs report the consolidated questions collected from the Global Bioanalytical Community. Six discussion topics were extracted from these questions and considered as the most relevant 'hot topics'. They were reviewed by internationally recognized opinion leaders before being submitted for discussion during the 13th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Regulatory Expectations & Validation

Flow Cytometry Biomarker Assay Validation & Regulatory Expectations

What are the key parameters for developing and validating flow cytometry-based biomarker assays? What are the minimal standards for exploratory, secondary and primary endpoint validation? What is the strategy for endpoints that transition from exploratory to higher endpoints such as enrollment criteria or label enabling? Can we use early clinical trial data as sources for validation, especially in cases of rare or difficult to obtain patient samples? What are the current and recommended practices?

Challenges & Approaches

Reagent Qualification

What are the best practices for reagent qualification? How many samples and runs are recommended? What are the acceptance criteria?

Absolute Counts

What approaches have been used? What is done for lyse/wash assays or bulk lyse assays? What requests have come from the Regulatory Agencies? What is the regulators' perspective on best practice for absolute counts by flow cytometry? Is it the same for exploratory vs secondary/primary objective?

Assay Performance Monitoring

What QC materials have been used? How many per run? What are the QC acceptance criteria? What is done if a run fails? Are the same requirements needed for an exploratory endpoint vs a secondary endpoint vs a primary endpoint vs enrollment criteria? How is assay variability monitored, especially for new complex exploratory methods (e.g., by QC)? What about replicates as for other technologies?

Data Analysis

Mining Multiplex Data

Can existing data be 'mined' in order to report new populations (e.g., $CD3^+CD4^+$ was validated together but now $CD3^+$ alone is required)? How would this be validated? A part of the discussion is intended to consider FFP principles as it applies to the application of new prospective analysis of existing validation 'raw' data. What are the current and recommended practices?

Big Data

Flow assays are getting more complex. This drives up the number of reportable results, complexity of gating, and complexity of analysis. How is this being handled? Is there a current best practice? What does the future look like?

Discussions, Consensus & Conclusions

Regulatory Expectations & Validation

Flow Cytometry Biomarker Assay Validation & Regulatory Expectations

Flow cytometry is currently the primary technology for multiplex single cell analysis. It is used in biomarker discovery and drug development, in assessment of drug-target engagement, pharmacodynamics, cellular pharmacokinetics, safety and in the assessment of efficacy biomarkers, that can be applied anywhere from exploratory to critical decision-making endpoints. Newer technologies such as mass cytometry and spectral cytometry are allowing for higher-dimensional evaluations. These technologies will undoubtedly generate data which will assist in elucidating biological and pathological pathways as well as new biomarkers.

There are currently no regulatory guidance documents specific to flow cytometry assay validation, however there are a number of industry White Papers that address best practices and recommendations for developing and validating flow cytometry methods [21,36–45]. Recently, the CLSI has developed a new draft guideline, H62: Validation of Assays Performed by Flow Cytometry [24]. This provides, for the first time in the flow cytometry field, validation guidelines for the international community. Its authors include representatives from the biotech industry, clinical laboratories, FDA, NIST, and reagent/instrument manufacturers. H62 is intended to be broadly applicable, providing best practices for basic researchers as well as guidance for approaches to assay development and key validation parameters associated with clinical diagnostic and drug development assays (primary, secondary and exploratory endpoints). The final CLSI H62 guideline is anticipated in 2020.

The development of a robust analytical method and understanding the underlying biology are key steps to designing a FFP validation based on the COU. Once analytical assay validation is completed and considered acceptable, clinical validation to establish the correlation between the biomarker and the outcome of interest can be assessed [46].

Given that it is sometimes difficult to obtain appropriate validation samples (e.g., in the case of rare diseases or patient-specific samples), it was agreed that early clinical trial samples (obtained for use with the proper informed consent and SOPs) can be used as sources of relevant biologic material for validation when not available during the assay development and validation phases. The process should be pre-defined in a validation plan and may require consultation with regulators depending on the COU. Provided the assay is well-established, a supplemental validation can be performed upon receipt of early clinical study samples representing the true disease state. Otherwise, creativity in generating validation samples that mimic the disease state, using disease cell lines, or genetically manipulated biological specimens that do or do not carry the markers of interest can offer efficient solutions for investigation with exploratory endpoints. There was focused discussion on validation strategies for methods performed by flow cytometry which transition from exploratory to higher endpoints (primary and secondary, or as patient enrollment criteria and label-enabling), which resulted in the same conclusions and recommendations described above in the BAV & Regulations section. For flow cytometric methods utilized for individual patient-treatment decisions, validation and sample analysis should follow CLIA regulations. For clinical trials where results are used to make individual patient-treatment decisions, IDE regulations may apply. Ultimately, there was consensus that the COU of the data should drive the FFP validation, employing best scientific judgment and consultation with regulators early and often as the COU evolves.

Challenges & Approaches

The inclusion of a flow cytometric method in a clinical trial presents a myriad of technical and operational challenges. Some key considerations in developing a robust flow cytometry biomarker assay were discussed and

recommendations for reagent qualification and monitoring, best practices for absolute counts, and appropriate assay controls are reviewed below. In addition, emerging approaches on novel gating strategies were discussed.

Reagent Qualification

New reagents (antibodies, cells for QC, and other critical reagents) should be qualified before being implemented on a study and should also be continuously monitored by functional assessment to ensure long-term stability and performance. Qualification and characterization steps should be driven by the intended use of the reagent. Qualification of a new reagent starts with titration in the appropriate reaction volume prior to evaluating assay performance. Change in reagent lot will usually require a reagent cross testing process prior to implementation. A crossover approach can be conducted bridging a minimum of three samples which will preferentially include normal healthy, patient or QC samples, depending on the relative nature of the assay. If a bridging assessment does not meet the acceptance criteria, a full qualification for the new reagent is required; ensure assay performance criteria with the new reagent meets the study needs, that is, receptor density (MFI), proportion of subset populations and so on.

Absolute Counts

The recommended best practice for absolute counts by flow cytometry is to use a 'lyse/no wash' assay approach on a single instrument platform. This eliminates the potential loss of cells during wash steps and is considered to be the most robust approach. For instruments that do not precisely measure acquired k sample volume, validated counting beads can be added to create a bead count to volume ratio to facilitate precise counting. If the assay is a 'lyse/wash' assay, a secondary 'buddy' tube for each sample is recommended to enumerate parent population using a 'lyse/no wash' procedure. The absolute count of the parent population from the 'buddy' tubes can then be applied for calculating the daughter subsets from the 'lyse/wash' assay and monitoring of percent relationships between the buddy and testing tubes can confirm that results are not impacted by selective loss/gain of cells from the testing tube processing. While the use of a single platform is preferred, dual platform approaches that take advantage of the flow cytometer's ability to accurately measure percent composition of cell populations whose absolute count per unit volume can be accurately measured by other instruments (most notably automated hematology analyzers) can also be suitable. Though in some COU it may be sufficient, the use of hemocytometer counts based on trypan blue exclusion is known to underestimate viability and is considered a less robust approach.

Assay Performance Monitoring

There are no standardized approaches to implementing QCs to monitor flow cytometry assay performance. The overall approach to assay performance monitoring should be scientifically-driven and related to the COU. The following approaches for monitoring assay performance were recommended as industry best practices. Critically, any QC samples used to control the flow cytometry method and critical reagents should be relevant to the cell population of interest. Stabilized whole blood is the most convenient and common QC reagent for immunophenotyping in whole blood or bone marrow. For functional studies, where live cells capable of responding to a functional pathway interrogated in the method are required, cryopreserved PBMCs are commonly used. Lastly, for patient-specific cell populations, for example in studies related to CAR-T cell therapy, lymphoma, or leukemia assessment, the use of 'spiked' QCs is a valid approach. In instances where appropriate QCs are not used, it is necessary to control the flow cytometry method and critical reagents via alternate means (e.g., using replicates, or monitoring internal populations). Implementation of QCs is recommended but may not be feasible for every run, and often it may be acceptable that QCs are implemented periodically to track that the assay is performing consistently. QCs should be well characterized for suitability with the assay criteria established before implementation in sample analysis. Overall, the process should be defined, scientifically justified, and the rationale should be documented.

Data Analysis

Mining Multiplex Data

Data from most flow cytometers are generated in the flow cytometry standard (.fcs) format [47], a structure which allows for reading of the data file from a variety of software packages. Owing to the specification of the .fcs files, files can be re-analyzed, or mined, for the evaluation of additional reportable results which were not included in the initial validation. For example, if a validated method reported $CD8^+$ T cells ($CD3^+$, $CD8^+$) and NK cells ($CD3^-$, CD56 and/or CD16), the data files could later be mined for the presence of NKT cells ($CD3^+$, $CD56^+$)

or $CD8^+$ NK cells ($CD3^-$, CD56 and/or CD16, $CD8^+$). The validation of these additional reportable results is sometimes called an 'electronic validation' as no new pre-analytical sample processing is involved. The group felt that this was an acceptable approach provided that the samples used during the initial validation were relevant to the study population. The electronic validation should follow the same workflow as the initial validation and include a validation plan supporting the intended analysis. The COU as well as the intended population should be considered to determine if the original validation data set is appropriate for data mining. Furthermore, care must be taken to ensure consistency between the common reportable results generated in the initial and electronic validations. If the initial gate and the new gating have overlapping reportable results, a plan on how to address this case needs to be defined and documented in advanced. It is critical to establish analytical confidence in the lower limit of quantitation for newly mined phenotypes based on acceptance criteria defined in the original validation or new COU.

Big Data

Another area of considerable advancement in the field of single cell analysis is in automated data analysis processes which will become more important with the higher-dimensional technologies. There are both supervised and unsupervised automated data analysis approaches available. As a field, it was discussed that more bioinformatic approaches should be integrated to advance new opportunities in biomarker discovery. It was agreed that moving towards automated analysis is suggested for the field, but these approaches still need to be scientifically and technically validated with well-established instruments and assays. Importantly, automated approaches will require documented explanation of the algorithm to ensure appropriate assessment of the biological and analytical variations. Full qualification of the bioinformatic software (security controls, if cloud based) must be performed prior to reporting of results as analytically validated.

Recommendations

Below is a summary of the recommendations made during the 13th WRIB:

1. **Flow Cytometry BAV** should include:
   - A state-of-the-art panel design, comprehensive method feasibility and validation plan are needed to ensure assay robustness. This should include antibody clone evaluation and appropriate antibody: fluorophore pairing. Where applicable, in addition to apparently healthy donors assay performance evaluation should include disease state samples in order to evaluate expected expression levels and potential interference;
   - Early clinical trial samples may be used as sources of relevant biologic material for validation when there are no available pre-study samples in the case of rare diseases or patient-specific samples. This clinical verification activity is typically performed as an 'in life' study after issuing an interim validation report illustrating analytical validity;
   - Where applicable, biological variability assessed using at least two baseline samples, separated by an appropriate time window, to evaluate intra-subject biological variability.
2. When flow cytometry assays are used in the US for individual patient-treatment decisions, the validation and testing should be performed following the CLIA associated regulations. Use of non-FDA cleared or approved devices in a US clinical trial may be subject to IDE regulations;
3. **Reagent qualification**:
   - Titrate the reagent first;
   - Bridge a minimum of three samples which may include normal healthy, patient or QC samples, depending on the relative nature of the assay. A full re-qualification is needed if the initial bridging shows a difference;
   - Acceptance criteria should be the same as the assay, capturing what is important to the study – that is, receptor density (MFI), proportion of subset populations – among others.
4. Best practice for **absolute counts:**
   - Use a 'lyse/no wash' assay on a single platform method (e.g., utilizing counting beads is preferable to a two platform method that calculates the absolute count as the product of the lymphocyte count from a hematology instrument and the percent value from the flow cytometer);
   - If the assay is a lyse/wash format, a secondary 'buddy' tube for each sample is recommended to enumerate parent population using a lyse/no wash procedure. The absolute count from the 'buddy' tubes parent population can then be applied to calculations for the daughter subsets in the 'lyse/wash' assay;

- If the use of a single platform approach is not feasible, a dual platform can be utilized.

5. Best practice for **assay monitoring**:
   - QC samples used should be relevant to the cell population of interest. In general, stabilized whole blood is the most convenient and common QC reagent for immunophenotyping in whole blood or bone marrow;
   - For functional studies, cryopreserved PBMCs are usually the best choice;
   - For patient-specific cell populations in studies related to CAR-T cell therapy, lymphoma, or leukemia, 'spiked' QCs are a logical approach;
   - In instances where appropriate QCs are not used, it is necessary to control the flow cytometry method and critical reagents via alternate means (e.g., using replicates), with emphasis on having *a priori* defined acceptance criteria;
   - Use of QCs is not always needed for every run; QCs may be implemented periodically;
   - The QC process should be defined, scientifically justified, and documented.
6. **Data mining** of existing data for validation of new cell populations is acceptable:
   - The samples used during the initial validation need to be relevant to the study population and a new validation plan for the new population needs to be defined in advance;
   - The COU as well as the intended population should be considered to determine if the original validation data set is appropriate for data mining;
   - Care must be taken if the initial gate and the new gating strategy have overlapping reportable results, and a plan on how to address this case needs to be defined and documented in advanced.
7. It was agreed that moving towards an automated analysis for **big data** is the goal. These approaches need to be scientifically and technically validated with well-established instruments, software and assays. Automated approaches require a documented explanation of the algorithms to ensure appropriate assessment of the biological and analytical variations. Timelines for full software and/or system qualification must be taken into consideration.

SECTION 4 – Interpretation of the 2019 FDA Immunogenicity Guidance

Meina Liang[30], Manoj Rajadhyaksha[22], Susan Richards[10], Becky Schweighardt[31], Shobha Purushothama[2], Daniel Baltrukonis[32], Jochen Brumm[7], Elana Cherry[33], Jason Delcarpini[34], Carol Gleason[8], Susan Kirshner[5], Robert Kubiak[6], Luying Pan[35], Michael Partridge[22], João Pedras-Vasconcelos[5], Qiang Qu[36], Venke Skibeli[37], Therese Solstad Saunders[37], Roland F Staack[17], Kay Stubenrauch[17], Al Torri[22], Daniela Verthelyi[5] & Haoheng Yan[5]

Authors in Section 4 are presented in alphabetical order of their last name, with the exception of the first five authors who were session chairs, working dinner facilitators, and/or notetakers. Author affiliations can be found at the beginning of the article.

Discussion Topics & Consolidated Questions Collected from the Global Bioanalytical Community

The following paragraphs report the consolidated questions collected from the Global Bioanalytical Community. Seven discussion topics were extracted from these questions and considered as the most relevant 'hot topics'. They were reviewed by internationally recognized opinion leaders before being submitted for discussion during the 13th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Drug Tolerance

Expectations for Approaches Regarding Drug Tolerance & Assay Specificity

What is the FDA expectation regarding drug tolerance for ADA assays – for example, steady state C_{trough}, C_{max} or at the sampling time points of the expected drug concentration for the anticipated high dose level? Is it the same expectation for NAb? How many validation runs are necessary for drug tolerance testing (both screening and confirmation formats)? How are drug tolerance results reported (PC sensitivity in the presence of drug at different concentrations vs level of drug tolerated in the presence of PC at different concentrations, mean or median if multiple runs)? Should drug tolerance for modalities other than antibodies be considered in molarity over mass units/ml, considering the potential order of magnitude size and molecular weight differences? How many companies are using methods to improve drug tolerance (e.g., BEAD)? Does the improved drug tolerance reveal data that is more clinically relevant? Is drug-tolerant immunogenicity testing in ocular fluids required? Is limiting outlier exclusion to only pre-existing reactivity/antibodies defendable assuming appropriate assay sensitivity is maintained?

Should we exclude outliers from consideration in the calculation of false-positive rates during validation? ADA data often does not correlate to efficacy. Does the reduction of the assay sensitivity target to 100 ng/ml make sense? Can FDA confirm the definition of 'false positive'? Is it correct to consider that the false positive is a screened positive and confirmed negative sample? If excess drug used in the confirmatory assay is not able to suppress the screening response to the assay baseline (i.e., close to the screening cut point), does it mean that the response is not specific? Should assessment of cross-reactivity be quantitative or can it be qualitative? Is cross-reactivity intended primarily for ADA recognition of endogenous proteins or do we need to consider pre-existing ADA binding to drug from previous exposure to a similar drug? What type of justification is acceptable when removing data points as outliers (statistical or biologic)?

Critical Reagent & Positive Control Characterization

Expectations for Approaches Regarding Positive Controls

Can FDA clarify their expectations for the characterization of PC? What is the FDA recommendation on the impact of PC and maintenance of assay performance for long-term immunogenicity monitoring? What are the expectations on evaluation of impact of PC on ADA assay sensitivity and performance for multiple epitope recognition? What is the influence of purification vs affinity purification of anti-serum PC on cut point determination? What is the impact of surrogate PC on determining sensitivity and clinical relevance? Some health authorities still request stability data for ADA assays. Do we agree that PC stability data is not needed and is not relevant? What is the requirement for recovery of spiked low-positive PC samples? How is it ensured that the 1% failure of low-positive PC does not change over time? How should this be assessed during validation to meet FDA expectations? How can discrepancies be correlated between the 1% failure LPC with assay robustness?

Assay Life Cycle Management & Clinical Relevance of ADA

Expectations regarding clinical relevance of ADA

The FDA immunogenicity guidance [48] requests follow up of samples until the responses have reverted to baseline for higher risk molecules. This can be challenging for practical, logistical and analytical reasons. Could we meet regulatory expectations by using a pre-determined titer level rather than absolute baseline? Once the clinical relevance of immunogenicity has been determined within clinical studies, what is the value of improving assay sensitivity or drug tolerance in the post-marketing setting? Does the utility of post-marketing assay improvements change if no clinical decisions are made based upon the assay results? If no discernible clinical impact of immunogenicity was found in clinical trials, should immunogenicity monitoring be continued in the post-marketing setting? If so, for how long?

Practical Challenges & Potential Solutions

Matrix Effects

Should hemolysis and lipemic tests be performed in ADA assay validation as in PK assay validation? What standardized matrix is recommended for testing bilirubin (icterus)? For the matrix interference evaluation due to hemoglobin (hemolysis) and lipids (lipemia), is it acceptable to use standardized matrix already employed in bioanalysis (i.e., 5% hemolyzed blood in plasma and more than 300 mg/dl triglycerides in subject natural plasma as per lipemic index)? Is diluted serum an acceptable surrogate matrix for aqueous humor and other rare matrix types (e.g., ophthalmology/ocular samples, but can be generalized to other rare matrix types like CSF)?

Pre-Existing Antibodies

When is a population considered to have a high prevalence of pre-existing antibodies (i.e., at what percentage of the population)? What is an appropriate negative control for populations with a high prevalence of pre-existing positives? Considering there is often a high degree of correlation between screening assay signal and confirmatory assay signal, when using competitive inhibition in bridging assays, is there any value in having a confirmation assay in populations with high prevalence of pre-existing Abs? Would it be more efficient to move directly to characterization assays (e.g., titer, neutralization)? The finalized immunogenicity guidance [48] suggests that a titer that is two dilution steps greater than the pre-treatment titer may be used to characterize a response as treatment boosted. Is this a statistically meaningful criterion, and are other methods being used throughout industry? In general, is there a minimum threshold Tier 2 percent inhibition that can be attributed to being a real pre-existing

antibody response (e.g., 50%)? Should we re-think the 5% false-positive concept for low risk human mAbs? Can we use a 1% false-positive rate?

Neutralizing Antibody Assays

Since integration of PK/PD/ADA data for low risk molecules is an alternative approach to interpreting neutralizing activity, can members provide examples of how it is being used? Do we need NAbs in oncology, rare diseases, or infectious diseases? If we have PK, ADA and a good PD marker, what additional information about clinical relevance do NAbs provide? Can we bank NAb samples and test as a post-marketing commitment?

Cut Points

Expectations Regarding Cut Points

If we concede that signal in the presence of drug should be similar to the SCP, is it correct to expect that a valid confirmatory response must be proportional to and can be calculated from the screening response? Do responses in the presence of excess drug that are significantly higher than the screening cut point indicate that the CCP is no longer applicable and that neither screening nor confirmatory assays can generate accurate ADA classifications, assuming the drug concentration used in the confirmatory assay is sufficient to completely suppress a high titer ADA? In terms of cut point assessment (SCP and CCP), what are the approaches being taken (method and statistical analyses) to ensure that biological/subject-sample variability is being considered? What are the considerations for re-setting study specific cut points after pre-study validation? The guidance recommends the use of a lower 90% CI on the cut point (95th percentile) in place of other approaches, as described in Shen *et al.* [49]. The suggestion is to base the cut point calculation using the average value for each sample rather than the individual values. This reduces the overall standard deviation, resulting in a lower bound below that expected from analytical variability. Would it be more appropriate to base the calculation on an estimate of SD that incorporates analytical variability? When is a dynamic cut point (instrument or analyst specific cut point) inevitable and acceptable? If response in the presence of excess drug corresponds to the level of non-specific binding in each sample, can it be used to monitor assay performance and integrity of critical reagents? Does using the signal in the presence of drug only to calculate %inhibition values deprive the researcher of a valuable piece of information? Guidance [48] recommends using 90% and 80% confidence intervals to ensure that the screening and confirmatory tiers generate at least 5% and 1% false positives, respectively. This means that the 2–11% false-positive rate mentioned in the 2018 White Paper in Bioanalysis Part 3 [21] is no longer tenable and the actual acceptable range detected must be higher. What is the maximum percentage of false positive ADA classifications that can be tolerated in the final data without confounding correlations between ADA and PK/PD and safety? What is the best approach to change the confirmatory assay cut point from 99.9 to 99% for a project that has been in the pipeline for many years? The guidance [48] suggests verifying that an established cut point factor is appropriate for a new patient population. As long as the positive rate is within the acceptable range of positives (e.g., 2–11%), is there added value in this exercise? What are the expectations for setting the cut point of the titration assay? Do we need to use screening cut point statistical requirements? How can we guarantee a robust cut point of the titration assay over time? Is setting a titer cut point that is at a minimum signal level in the linear range of the titer curve approach acceptable? If we use study specific cut points in the ADA assays for a drug in the same indication, does this imply population differences in the studies? Can these data still be used in an integrated manner to assess overall ADA incidence?

Discussions, Consensus & Conclusions

Drug Tolerance

Expectations for Approaches Regarding Drug Tolerance & Specificity

Tiered testing strategies are typically used for ADA testing. A common challenge for ADA testing is the development of an ADA detection method with adequate sensitivity in the presence of the drug. It is well known that the drug can interfere with the detection of ADA, resulting in false negatives. This is a particular concern for therapeutic monoclonal antibodies and other drugs that have high sustained circulating drug levels.

The consensus was that immunogenicity assays should demonstrate drug tolerance consistent with the drug levels at the time points of ADA sample collection. Generally, sampling at C_{trough} is sufficient; however, novel modalities may necessitate earlier sampling in part due to lack of prior knowledge or concerns about early safety events. To demonstrate drug tolerance during assay validation, one to three runs by one or more analysts are recommended. There is no mandatory format for reporting results; both mean and median values are acceptable. Frequently, data

are reported in a grid format for various PC concentrations, each in the presence of various drug concentrations in either molar ratios or mass units. When determining assay sensitivity, the target of 100 ng/ml ADA in the presence of expected concentrations of drug in the patient is recommended.

In terms of the methods to improve assay drug tolerance, consensus was that acid treatment is the most common approach. However, the acid dissociation procedure does not always reduce drug interference to the desired levels, potentially due to re-association of drug with ADA upon neutralization. In addition, acid dissociation procedures can also lead to loss of low affinity antibodies. Several alternative approaches have been developed to reduce drug interference from ADA detection (e.g., ACE [50], SPEAD [51], PandA [52]). Sponsors should validate the use of any new approaches to improve drug tolerance. It is a regulatory expectation that methods developed are specific for ADA and matrix interferences are evaluated and mitigated; for example, reduce any interference from target and rheumatoid factor [48,53,54]. Current industry practice is to assess immunogenicity systemically using a drug tolerant assay regardless of the route of administration (e.g., intravenous, subcutaneous, intra-ocular).

ADA recognition of endogenous proteins, also called cross-reactivity, may pose a safety risk. Hence, the clinical relevance of this reactivity should be explored in the context of the risk assessment. Cross-reactivity assessment is generally qualitative in nature, although cross-reactive antibody titers may be evaluated in some cases. Induction of antibodies that cross-react with other related therapies, for example, anti-PEG antibodies, may also have clinical implications. Studies should be performed to address both concerns.

Critical Reagent & Positive Control Characterization

Expectations for Approaches Regarding Positive Controls

Positive controls play an important role in ADA assay development and validation. These are surrogate ADAs routinely used to ensure that assay specificity, sensitivity and drug tolerance meet the study's needs. Purified pAbs from hyper immunized animals and mAbs have been commonly used as positive controls. For multi-domain therapeutics such as bispecific antibodies, fusion proteins and pegylated proteins, evaluation of ADA specificity against different product domains is recommended [55]. Antibody engineering using phage display technology has been used to generate positive controls aimed at specific regions or epitopes of biotherapeutic products. Because it may be important to understand the specificity of the immune response against the different domains, the assay or assays must be able to detect ADA to the different domains; this may require one positive control that is reactive to all domains (generally a pAb) or multiple positive controls. In certain situations, characterization of ADA isotypes might help elucidate the mechanism of unwanted immune responses. It is emphasized that positive controls are surrogates and the assay's ability to detect ADA in study samples is not contingent on the positive controls. Therefore, specificity, sensitivity and drug tolerance determined using positive controls only provide an approximation of the true assay performance.

For characterization of PCs, the majority view was that during assay validation only short-term stability, for example, ambient temperature and freeze–thaw stability, are useful but there was sufficient experience and general acceptance that performing long-term stability of the positive control, is often not necessary [56–58]. However extended stability still needs to be considered for critical reagents.

When setting the LPC, a 1% failure rate is recommended but not mandatory. The goal is to ensure assay LPC is set up appropriately to capture changes in assay performance over time.

Assay Life Cycle Management & Clinical Relevance of ADA

Expectations regarding clinical relevance of ADA

The emergence of ADA may have negative clinical consequences on the treatment outcomes of biologics. For this reason, sponsors are expected to fully characterize and understand the immunogenicity profile of new biologic therapeutics.

The ongoing regulatory push to establish more sensitive ADA assays has had the effect of increasing the FPR and resulted in a concern that increasing assay sensitivity by increasing FPR is also causing a reduction of assay specificity that further confounds the immunogenicity dataset. Taking all these factors into account, it makes it more challenging to determine the clinical relevance of ADA. The regulators recommend that the analysis of clinical impact includes not only data on binding antibodies, but also titer and NAbs.

For high risk molecules the FDA immunogenicity guidance [48] requests follow up of subjects until ADA responses have reverted to baseline. For low risk molecules or molecules where the immunogenicity profile is well established,

the extent of the follow-up can be decided using a risk-based approach. This can be challenging for practical, logistical and analytical reasons but consensus confirms that this is the method that is often being followed. Patient compliance and undue burden for an extremely sick population should also be taken into consideration before implementing long-term follow-up. Alternative approaches such as titer reduction to a level known not to have clinical consequences may be an option. Other strategies or approaches are possible but should be discussed with regulators. Regardless of the selected approach, it is key that appropriate informed consent is obtained for this follow-up testing.

The necessity of immunogenicity monitoring in the post-marketing setting was discussed, especially when ADA showed no clinical relevance during the clinical trials. Regulators may request this when there are concerns that the type or affinity of ADA changes over time and the trials performed were too short or insufficiently powered to properly assess the longevity of the ADA response and their putative impact on patient safety. When requests are made to improve the ADA assay sensitivity or drug tolerance in post-marketing requirements and post-marketing commitments studies, generally it is accompanied by a request to re-test the samples from the clinical trials. In addition, improved assays may be required when testing for additional indications.

Practical Challenges & Potential Solutions

Matrix Effects

Cases were discussed where regulators have requested that the impact of hemolysis and lipemic matrices should be investigated in ADA assay validation. Industry perspective was that the current body of data indicates lack of interference from lipemic and hemolyzed samples. However, health authorities have seen data where hemolysis and lipemia can impact assay performance. Current expectation is that these evaluations continue to be performed as part of validation until enough data are collected to strongly demonstrate no impact. For rare matrices (e.g., ophthalmology/ocular samples, CSF), surrogate matrix may be used with appropriate rationale.

Pre-Existing Antibodies

The presence of pre-existing antibodies to a biotherapeutic may elevate ADA responses in treatment-naïve populations which can confound assay cut point calculations and increase the risk of false negative results in-study. Depending on the prevalence of pre-existing antibodies in the treatment-naïve population, different strategies can be used to mitigate their impact on assay cut points. A variety of practical approaches to dealing with a high prevalence of pre-existing antibodies have been put forward by investigators [59]. For example, one approach is to screen individuals and select the negative population for the negative control pool.

Usually, there is a high correlation between screening and confirmatory assay signals, in populations with a high prevalence of pre-existing antibodies. In some situations, sponsors have titrated ADA immediately after screening without the use of a confirmation step [59]. A strategy like this was generally deemed acceptable, as long as it allows for the detection of treatment emergent ADA and does not confound the assessment of clinical impact of ADA.

The finalized immunogenicity guidance [48] suggests that a titer that is at least two dilution steps greater than the pre-treatment titer may be used to characterize a response as treatment boosted. Other approaches have been used as well [60]. It is recommended to report the numeric titer value not the log titer. One can consider orthogonal methods (e.g., immuno depletion) to distinguish true pre-existing antibodies from non-specific binding [61].

If sponsors are considering restricting outlier exclusion to pre-existing reactivity/antibodies, or any new ways of analyzing the data, a clear justification should be provided. In addition, sponsors can request to present their new strategies to health authorities at regulatory meetings within the context of their product-specific program.

Neutralizing Antibody Assays

The regulatory expectation for neutralizing antibody assays to detect NAb as part of the tiered bioanalytical approach to support immunogenicity assessment of pivotal clinical studies was discussed. It was concluded that development of neutralizing antibody assays is of concern to regulators as NAb can inform the safety and efficacy of the program. The regulatory expectation for neutralization assays is not indication-driven and is independent of product class. The goal is to correlate the induction of neutralizing antibody responses with clinical outcomes and to include neutralizing antibody rates as part of product labelling.

In some instances, when there are very sensitive PD biomarkers, sponsors can consider their use as an alternative approach to assessing neutralizing activity. Regulators confirmed that they are open to discussing this option when a scientific justification is included in the overall data package.

Cut Points

Expectations Regarding Cut Points

A critical parameter of immunogenicity assays is the setting of an appropriate cut point. This impacts reported assay sensitivity and identifies a clinical sample as positive or negative for ADA using a tiered analysis. There is general alignment within industry on how to establish cut points using statistically-based approaches. WRIB has played an important role as a forum for such immunogenicity discussions and through published White Papers that have influenced industry practice [12,15,18,21].

A tiered approach for testing of samples for the presence of ADA was envisioned to ensure low frequency of false negative and false-positive classifications in the final data set obtained after two consecutive tiers. Current regulatory expectations recommend that cut points be based on a false-positive rate of 5% and 1% in the screen and confirmatory tiers, respectively [48].

While selection of suitable false-positive rates for each tier has been extensively discussed, relatively little attention has been dedicated to understanding conditions required for the confirmatory tier to accurately eliminate false positives generated in the screening tier. Two scenarios with the confirmation assay were discussed: the common observation that signal in the screening assay is highly correlated with percent inhibition in the confirmation assay [62]; and the relatively rare examples when samples with high screening signal show sufficient inhibition (at or above the confirmatory cut point) to be confirmed as positive but the signal in the presence of drug is still significantly above the screening cut point. For the first scenario, in most instances the two tiers are highly correlated, which has let to proposals that the confirmation assay is similar to a screening assay with a 1% FPR, and there may be little benefit in performing the second tier. However, the FDA disagrees with screening using a 1% FPR because of the potential increase in false negatives. In the rare instances where the second scenario is observed, an investigation may be warranted to understand the magnitude of the drug-specific raw signal (true ADA), and to examine what serum components may be generating the non-inhibitable (non-specific) assay signal.

If during the confirmation step, the responses in the presence of drug are significantly higher than that without, it may indicate that the confirmatory cut point, determined in a naïve population, is no longer applicable and that neither screening nor confirmatory assays can generate accurate ADA classifications. A specific cut point in patient populations should be considered. Orthogonal methods and sample dilution may also be needed to understand the assay signal.

Assay development should include a determination of the appropriate amount of unlabeled drug spiked into the samples for the competition. Additionally, verification that the screening assay has sufficient specificity and does not detect non-specific binding need to be demonstrated. Therefore, the desired high sensitivity and specificity of ADA detection may be accomplished in the screening tier. In addition, the confirmatory assay in its current format can also be applied for monitoring of reagent integrity and assay performance. Regulators stated that alternative approaches will be considered if properly justified.

Pre-Study & In-Study Assay Cut Points

In-study cut points are derived from samples collected in a clinical study prior to treatment, as opposed to cut points derived from commercial samples or prior studies in other populations. Because differences in patient population, serum collection, or storage conditions can lead to differences in the distribution of the pre-treatment scores, in-study cut points may be required to adequately assess immunogenicity in the study population. This leads to a cut point lifecycle during the development program for a novel drug, where the trade-off between operational simplicity (not changing the cut point when conducting a new clinical trial) and adhering to the targeted false-positive rates (and hence changing the cut point) needs to be carefully considered. However, consensus was reached that if a suitable screening false-positive rate is observed with in-study baseline samples, then a study specific cut point is not required. Similarly, if a cut point is determined for one population, and the screening false-positive rate in another subsequently tested population is within the false-positive range, there is generally no need to reassess a new cut point. It is, however, also recommended to carefully assess the distribution of the scores and identify changes in mean and variance that can indicate the need for a new cut point. The assessment of the distribution of the scores is more powerful than a criterion based on observed false-positive rates, and as such can offer advantages in situations where few in-study samples are available.

In practice, cut point determination can be influenced by many factors such as sample size, data distributions and transformation, populations, pre-existing reactivity, and methods used for determining analytical and biological outliers. Some voiced that current industry practices often lead to excessive removal of the inherent biological

variability and could result in low cut points which only represent the analytical variability of the method [63]. This has resulted in debate within the industry on whether these ultra-low cut points are appropriate or could lead to over reporting of positive ADAs that may not be clinically relevant.

Assay cut points have historically been determined by the analysis of a panel of ~50 individual samples, by different analysts, over different days, to account for both biological and analytical variations. Cut point values that more accurately reflect the heterogeneity usually observed in the target population by giving appropriate weight to the biological factors that are usually the major contributors to assay variability may be generated using baseline samples from the patient population. In addition, the statistical approach used to remove outliers, especially when dealing with heterogeneous diseased populations, can also affect the false-positive rates observed for the therapeutic being tested. One of the approaches is using box plots with 3 IQR rather than with 1.5 IQR to assess the SCP and CCP since 3 IQR is less likely to remove all biological variability from the cut point data set. Regulators stated that they will consider this approach, if properly justified. The use of a dynamic cut point (i.e., instrument or analyst specific cut point) is strongly discouraged.

For a project that has been in the pipeline for many years where the CCP was calculated using 0.1% FPR, regulators recommended that sponsors provide the calculation for both the original and recalculated cut points. If a separate titration assay cut point is needed, using the cut point at the 99.9% confidence level based on the screening assay dataset, The robustness of the cut point of the titration assay should be maintained by using an appropriate study-specific cut point if needed.

Recommendations

Below is a summary of the recommendations made during the 13th WRIB:

1. The assay should demonstrate adequate drug tolerance at the time points selected in the trial. Sampling at C_{trough} is sufficient for modalities like monoclonal antibodies, However, novel modalities may necessitate earlier sampling due to lack of prior knowledge or early safety events. Establishing ADA methods that are drug-tolerant to the presence of drug C_{trough} is generally sufficient;
2. The target of 100 ng/ml for ADA assay sensitivity is recommended, not mandatory;
3. One to three runs with one or more analysts are recommended for drug tolerance testing;
4. There is no mandatory format for reporting drug tolerance results; both mean (and standard deviations) and median (and range) values are acceptable. Most often, data is reported as a table showing different levels of ADA positive controls and drug expressed either as molar ratios or as concentrations in mass units;
5. When generating critical reagents consider aspects of the assay LCM in choice of PC;
6. During ADA assay validation, short-term (e.g., ambient temperature) and freeze–thaw stability of the PC reagent are useful but performing long-term sample stability on PC or patient-derived antibody samples is often not considered necessary;
7. For multi-domain biologics, the assay must be able to detect ADA against each domain. This may require using one polyclonal positive control with reactivity to all domains or multiple positive controls with specificity for each domain;
8. When setting the LPC, a 1% failure rate is recommended, but not mandatory. The goal is to ensure that assay LPC will appropriately capture changes in assay performance over time;
9. Ensure informed consent allows for appropriate follow up of ADA samples as needed;
10. Regulators may request immunogenicity monitoring in the post-marketing setting because there is concern that trials may be too short or have too few patients to accurately assess development of ADA and to observe clinical relevance of ADA;
11. Regulators expect the impact of hemolysis and lipemia on assay performance to be evaluated;
12. For rare matrices (e.g., ophthalmology/ocular samples, CSF), a surrogate matrix may be used with appropriate rationale;
13. One of the approaches to generate a negative control pool for a population with a high prevalence of pre-existing antibodies would be to screen individuals and select the negative population;
14. Sponsors who would like to use outlier exclusion for only pre-existing reactivity, assuming appropriate assay sensitivity is maintained, should contact regulators for a discussion of rationale;
15. Sponsors should consider using orthogonal methods (e.g. immuno depletion) or further sample dilution to distinguish true pre-existing antibodies from matrix interferences;

16. Neutralization assays may help to correlate ADA and clinical outcomes and are included in labelling. The development of neutralizing antibody assays for biologics is expected by regulators. Using alternate approaches like PK/PD integration may be an acceptable approach with suitable justification but should be discussed with regulators;
17. Frequent confirmatory responses where signal in the presence of excess drug is much higher than the screening cut point should trigger an investigation of assay performance and critical reagent integrity;
18. Consensus was reached that if a screening FPR suitable with baseline samples is attained, then a study specific cut point is not required provided suitable outlier analysis was performed. Similarly, if a cut point is determined for one population, and the false-positive rate in another subsequently tested population is within the screening FPR, there is no need to reassess a new cut point. Suitable statistical justification should be provided in the validation report;
19. Using baseline samples from the disease population can generate cut point values that more accurately reflect the heterogeneity usually observed with clinical study populations by giving biological variability greater weight relative to analytical variability, which is typically low;
20. The possibility was discussed of using box plots with 3 IQR rather than with 1.5 IQR to assess the SCP and CCP, since 3 IQR is less likely to remove all biological variability from the cut point data set. However, regulators request that suitable justification be provided. Sponsors may be requested to provide the analysis using both approaches for comparison;
21. The use of a dynamic cut point (i.e., instrument or analyst specific cut point) is strongly discouraged by regulators;
22. To change the confirmatory assay cut point from 99.9% to 99% for a project that has been in the pipeline for many years, regulators recommended that sponsors provide clinical sample analysis data using both the original and recalculated cut points;
23. The expectation for setting the cut point of the titration assay is 99 or 99.9%;
24. When using the FDA guidance recommended lower 90% bound on the screening cut point (95th percentile) and 80–90% lower bound on the confirmatory assay cut point (99th percentile), the variance estimate should include all sample replicates to ensure all possible sources of variability, and not the average response for each sample.

SECTION 5 – *In vivo* & *Ex vivo* Gene Therapy & Vaccine Bioanalytical Challenges

Boris Gorovits[11], Rachel Palmer[10], Mark Milton[42], Brian Long[31], Bart Corsaro[38], Vahid Farrokhi[11], Michele Fiscella[39], Neil Henderson[40], Vibha Jawa[41], Jim McNally[35], Rocio Murphy[43], Cynthia Rogers[28], Shabnam Tangri[25], Hanspeter Waldner[44], Yuanxin Xu[29] & Tong-Yuan Yang[45]

Authors in Section 5 are presented in alphabetical order of their last name, with the exception of the first four authors who were session chairs, working dinner facilitators, major contributors and/or notetakers. Author affiliations can be found at the beginning of the article.

Discussion Topics & Consolidated Questions Collected from the Global Bioanalytical Community

The following paragraphs report the consolidated questions collected from the Global Bioanalytical Community. Eight discussion topics were extracted from these questions and considered as the most relevant 'hot topics'. They were reviewed by internationally recognized opinion leaders before being submitted for discussion during the 13th WRIB. The background on each issue, discussions, consensus and conclusions are in the next section and a summary of the key recommendations is provided in the final section of this manuscript.

Approaches to Gene Therapy Bioanalysis

Safety Assessment & Bioanalysis

Based on the development phase of the therapeutic, biodistribution/shedding and immunogenicity have become integral parts of safety assessments in nonclinical and clinical development of gene therapies. There are many samples that can be collected and many potential analytes that can be generated. Consequently, the cost of bioanalytical support for a gene therapy may be high. How should sponsors determine the right amount of analysis to perform, achieving a balance between ensuring safety and minimizing the costs of drug development?

qPCR Validation

Building on the 2018 White Paper Part 1 recommendations on qPCR validation [19], what are the critical assay validation parameters? Are LBA assay validation criteria suitable for accuracy and precision, %CV and %Bias, respectively? How is a stability program for qPCR assays conducted? For CAR-T programs, is it necessary to assess storage stability of whole blood containing CAR-T cells? Is it necessary to assess stability on extracted gDNA from study samples? Is ISR for qPCR performed in the same way as LBA, LCMS and hybrid LBA/LCMS? How do we conduct selectivity? For a CAR-T program, do we spike CAR-T cells from a normal donor into diseased whole blood if we can procure the diseased whole blood? What is the target amplification size limit (bp), shortest to longest? What is the optimal primer length for specific amplification?

Assessment of Shedding & Infectivity Assays

Does the infectivity assay need to be quantitative when coupled with quantitative qPCR results? Does infectivity data change the follow up plan? Under what circumstances would environmental monitoring be required? What are practical implications of viral shedding results for non-pathogenic vectors like AAV? Is infectivity data ever required? Is data on shedding required for retinal gene therapy? Can matrix selection be focused to key matrices like tears and not assessed in other standard matrix types? For qPCR assays, are sensitivity requirements the same in pre-clinical and clinical assays? Is clinical diagnostic guidance on qPCR assay validation sufficient to demonstrate fit for purpose for viral shedding? Are there any parameters that would be unique for a gene therapy vector vs detection of virus in infectious disease?

ELISpot

Is ELISpot required for a gene therapy product? Does it matter where the gene therapy product is administered (e.g., CNS, eye)? What is the relevance of circulating activated T cells in those cases? ELISpot is a very challenging assay to perform especially for larger multicenter studies. Can ELISpot results be normalized (intra and inter-subject)? What are ELISpot assay expectations and performance for cell recovery and viability; critical steps in standardizing an ELISpot assay; overcoming ELISpot assay variability; controlling assay performance over time or across laboratories; using cryopreserve high-quality PBMCs for ensuring optimal performance in functional ELISpot assays?

Vaccines

Vaccines

What are the recommendations for efficient bridging to newer technologies (e.g., moving from a standard ELISA to a multiplex assay) for vaccine clinical assays? What is the role of assay controls and proficiency panels in QC trending and assay life cycle maintenance? How can it be determined when an assay is out of trend? What controls are needed in an ADA compared to a vaccine LBA and how often should they be run? What data is needed when bridging to a new critical reagent? What is the degree of assay qualification/validation appropriate at each phase of vaccine clinical testing?

Challenges

CRISPR Genome Editing

What special considerations might be needed for a gene editing therapeutic versus a traditional small molecule or large molecule biologic: gene editing confirmation/cell therapy characterization; safety/toxicology; PK/PD; persistence/long term follow up? What is regarded by the community as the optimal therapeutic window of CRISPR – for example, %number of cells with on-target edits? How will the use of different delivery systems of RNP complex influence bioanalytical requirements? Do regulators expect total RNP complex, active RNP complex or both for gene editing therapeutic exposure/biodistribution data? Taking biopsies in preclinical studies is viable for bioanalytical testing; would tissue biopsies be required for clinical studies?

Biodistribution

What are the regulatory expectations for the qPCR assays, both viral capsid and transgene expression assays for biodistribution studies? What are the main challenges in using hybrid LBA/LCMS for transgene products? What are the choices of protein or peptide immunoprecipitation in transgene products? What correlations need

to be established between hybrid LBA/LCMS with other techniques (qPCR and/or flow cytometry) to ensure reliable data in detection of transgene and transgene expression and why? Do regulators expect platform based comparative value of transgene and transgene expression (qPCR, flow cytometry and hybrid LBA/LCMS)? What are the challenges in performing analytical assay compatibility for measuring transgene and transgene expression in different matrices/disease indications? Are there currently available controls/calibrators for maintaining/comparing transgene and transgene expression assays?

Immunogenicity

Immunogenicity

What is the value of anti-capsid total antibody and neutralizing antibody assays? Are both binding and NAb data needed for both virus and transgene? Are functional assays to assess vector activity more useful than conventional binding and NAb assays? What are the expectations for the ADA, NAb assay sensitivity limits? What is the current application of ADA, NAb and cellular immunity (ELISpot) methods as part of inclusion criteria? What is the industry and regulators' agreement on the relevance of pre-existing immunity criterion for pre to post immune response and the current strategies to mitigate risk? What are the mechanisms conferring pre-existing AAV immunity? Can the 2019 FDA Immunogenicity Assays Guidance for biotherapeutics [48] be adapted for gene therapies? What are the areas that are divergent? Is there an agency position on prophylactic immune tolerance regimens that would prevent the formation of antibodies and allow for sequential dosing?

Discussions, Consensus & Conclusions

Approaches to Gene Therapy Bioanalysis

Bioanalysis & Safety Assessment

New gene therapy modalities are gaining significant attention in addressing an unmet medical need. The therapeutic objective is to treat a genetic disease/condition which is often caused by a single gene defect by administering a single or limited number of treatments. The aim is to achieve successful expression of a functional version of a single protein either universally or in a targeted tissue. With encouraging results from preclinical studies and the emergence of gene therapy in the clinical setting, there is a significant need for innovative bioanalytical techniques, for example for the measurement of the transgene products in both preclinical and clinical stages. Various technical challenges such as efficient protein extraction, measurement specificity, sensitivity, accuracy and assay translatability between preclinical and clinical settings need to be addressed.

Biodistribution/shedding and immunogenicity have become integral parts of safety assessments in nonclinical and clinical development of gene therapies. There are many samples that can be collected and many potential analytes that can be assayed. Consequently, the cost of bioanalytical support for a gene therapy will be highly expensive. It is clear that cost cannot be a reason to compromise on safety. It is important to ensure the relevance of all testing using a risk-based approach that has been agreed upon with health authorities. For instance, there may be significant challenges in obtaining positive controls for cellular immune response assays (e.g., ELISpot). Although proper due diligence must be done to attempt to find an appropriate positive control reagent, it can be costly to keep looking indefinitely. Regulators are open to alternatives if the proposal is scientifically valid but are concerned that signals could be missed in early trials, hence it is important to obtain an aligned position between the sponsor and the regulators. Banking of samples collected during the conduct of clinical trials is also recommended if it is unclear what questions may need to be addressed early in the development of a therapeutic.

qPCR Validation

CAR-T therapies present an example of *ex vivo* gene therapeutics. CAR-Ts have been shown to be highly effective for the treatment of some hematological malignancies where high response rates have been observed. Several CAR-T therapies have received marketing approval and many more are in clinical trials. Persistence of CAR-T cells in a patient's circulation plays a critical role in long-term efficacy. Robust methods are needed to monitor circulating CAR-T cells to establish the PK/PD and safety relationship in clinical settings.

Because of its high sensitivity, qPCR is the most commonly used methodology for monitoring the fate of CAR-T cells in a patient's circulation, and given its ultra-high sensitivity, it is especially useful for monitoring low quantities of CAR-T cells as part of long-term studies. To prepare treatment appropriate standards and QCs, it is recommended, when possible, to spike CAR-T cells into diseased whole blood. The optimal primer length for the specific amplification should be sufficient to detect the CAR-T inserted transgene.

With limited regulatory guidance and industry White Papers for qPCR method development and validation, considerations for method development and validation strategies to support regulated bioanalysis for CAR-T therapies in clinical studies continued to be 'hot topics' . Discussions built upon the 2018 White Paper recommendations on qPCR validation [19]. In the absence of regulatory authority guidance, it was suggested to follow scientifically-led method development and validation strategies, with support from the MIQE guidance [64], which focuses on consistency of qPCR performance using a ligand-binding plus enzymatic function assay (based on primer, probes, polymerase and RT enzymes). Initial qualification of the assay describes what performance characteristics can be achieved. Validation then describes the performance of the assay against pre-defined criteria. Sensitivity or limit of quantitation (LOQ; 50 copies/µg gDNA) and limit of detection (LOD), precision, accuracy, DNA extraction efficiency from tissues, and engineering controls to ensure there is no cross contamination should be evaluated. Because qPCR is used for the detection and quantification of viral load in a diagnostic setting, CLSI also provides useful guidance documents for assay validation in the clinical laboratory [65].

Additionally, it is necessary to assess storage stability of whole blood containing CAR-T cells as well as stability of the extracted gDNA. As is the case for any analyte, stability should be assessed under intended sample storage conditions. Stability should also be performed on gDNA extracted from study samples. Utilization of surrogate markers (normalization genes) for stability testing in the relevant matrix may be acceptable depending on the COU. The necessity of performing ISR for qPCR tests was questioned. Given the low number of study samples and limited quantity from some matrices, ISR may not be relevant or feasible.

Assessment of Shedding & Infectivity Assays

Viral vector gene therapies pose unique safety and bioanalytical challenges that can vary based on the type of viral vector used, the properties of the transgene, as well as the route of administration and target tissue. Since viral therapies carry a risk of shedding and potential environmental exposure, studies are required during clinical development to measure viral load in various secretory (e.g., saliva) and excretory (e.g., urine and feces) matrices. The exception is for *ex vivo* administered lentiviral vectors, for which no infectivity assay is required [66]. The shedding data forms part of the environmental risk assessment and is most important for replication competent viruses. Even when the probability of shedding of the virus is low (e.g., for subretinally administered gene therapies which have limited distribution from the site of administration), viral shedding is still typically assessed. A recent draft guidance [67] on gene therapy for retinal disorders does not list viral shedding as a necessary follow up study suggesting that the regulators may not always require this assessment.

The type of matrices and assays required as well as the timing of sample collection are dependent on the type of viral vector and route of administration, outlined in an FDA guidance [66]. In the case of replication incompetent and non-pathogenic vectors such as AAV, assessment of viral vector shedding is still required with the qPCR-based detection being sufficient for monitoring. With vectors having a higher risk of shedding live infectious virus, like HSV oncolytic viruses, cellular infectivity assays may also be required to understand and, if necessary, mitigate the risk of exposure to non-treated individuals. The selected infectivity assay needs to be quantitative even when coupled with qPCR results.

Infectivity data can change the design of clinical trials (i.e., additional or more frequent safety assessments may be required). Long term follow-up is recommended for RCL, although not every patient may need to be evaluated. Real-time analysis is not required and banked samples can be used.

ELISpot

One of the potential concerns associated with viral vector-based gene therapies is the development of cellular immune responses which may result in loss of efficacy or tissue damage. ELISpot is a method commonly used to detect cellular immune response to specific antigens (e.g., viral vector coat proteins that are presented on the surface of infected cells); much like ADA and NAb assays are used to measure humoral responses. The need to monitor these responses should be determined using a risk-based approach while factoring in the route of administration and the type of the viral vector used.

Unlike ADA or NAb detecting analytical protocols, there is no regulatory guidance on how to develop and validate ELISpot methods although industry White Papers are available [68] to clarify on harmonization of practices and analysis of the quality of results. To add to the challenge, ELISpot assays require a more complex workflow from sample collection to testing, especially for larger multicenter studies. Sample collection procedures should be developed with the knowledge of the availability of certain equipment at the study sites. Multiple pre-dose

samples can be collected in order to generate a more robust baseline value. Results can be normalized relative to pre-study values in order to partially mitigate inter-site differences in sample collection and handling. It should be noted that although it would be ideal for all samples to be collected at all study centers, it is not imperative to do so. The potential sources of variability make the need for standardized approaches even more important. ELISpot harmonization consortium White Paper [68] which outlines assay expectations and performance criteria may be helpful to develop ELISpot assays.

Vaccines

Vaccines

Vaccine serologic assays are one of the bases for licensure of vaccine products and are used to measure immunogenicity and vaccine efficacy endpoints in clinical trials. Vaccine clinical assays can also serve as correlates of protection when shown to be predictive of clinical benefit. They are also required in support of post-licensure regulatory commitments including manufacturing changes and new age indications. For as long as the licensed vaccine remains on the market, regulatory agencies require that the clinical endpoint assays be consistent and maintained in a validated state. In response to this regulatory requirement, a phased approach to assay development and validation is used which assures clinical phase appropriate data and assay consistency.

The early development and optimization of the clinical assay is critical and must be robust and rugged enough to enable the assay to perform consistently and endure through potentially decades of clinical testing. During the assay setup phase, the preliminary assay establishes the assay design and identifies critical reagents and parameters. This step can be challenging due to lack of available samples that represent the intended population. It may be necessary to initially rely on knowledge gained from preclinical assay development to evaluate first in man clinical trials. Then once human vaccinated samples are available, additional assay development can be completed. Assay validation requires that pre-defined acceptance criteria for the assay performance are met. During Phase I and early Phase II studies, only assay qualification is needed, with the evaluation of limit of blank (negative samples), LOD, LOQ, linearity/range, specificity, and precision. In later phase studies, a fully validated method is required.

Lifecycle maintenance of validated vaccine assays is essential to ensure that the assay can support long-term endpoints, concomitant studies, or any additional testing commitments required from the regulatory agencies. Lifecycle management for these assays is resource intensive and entails ongoing assay performance tracking and critical reagent bridging. Availability of sample proficiency panels is important for monitoring long-term assay performance.

Assay standard and quality control performance trending is also critical. Best practices indicate that assay controls should be run on every plate to provide data that can be used for assay system suitability and assay performance trending over the long term. No consensus has yet been achieved on best approaches for evaluating assay trending or determining when an assay is considered out of control.

Due to the potentially long term use of the assays, it is desirable to take advantage of newer technologies that may increase efficiency or improve assay robustness. To take advantage of these technologies, bridging needs to occur between the original assay method and the new one. To successfully bridge technologies, an understanding of the relationship between the assays and confirmation that the assays are equivalent are needed. An example of the extensive evaluation needed for the comparison between single and multiplex assays is given by Feyssaguet *et al.* [69]. Clinical samples with antibody concentrations or titers that span the entire range of response are needed and new assays will need to be validated. New critical reagents need to be bridged to assure consistent performance of the assay; bridging should be performed according to O'Hara *et al.* [31].

Challenges

CRISPR Genome Editing

The CRISPR technology is a novel gene editing method that has the potential to transform healthcare by allowing for the development of gene-based therapeutics through gene editing. The CRISPR/Cas9 RNP complex is composed of a sgRNA and the Cas9 endonuclease (a bacterial protein). The sgRNA binds to a specific 'target' sequence on the DNA and allows Cas9 to create a double strand DNA break at that precise sequence. Several cell endogenous repair pathways are known to influence the outcome of CRISPR/Cas9 DNA breaks and the most active are NHEJ and HDR. NHEJ is described as a 'fast and error-prone' pathway, during which the DNA break is thought to be repaired and rebroken repeatedly by the active CRISPR/Cas9 complex until a 'mis-repair' event creates a

permanent indel that could be leveraged for medical care. Conversely, HDR is extremely 'precise, slow and rare (i.e., low frequency)' but its repair outcome is usually ideal for therapeutic applications. Reducing the potential for 'off-target' interactions between CRISPR/Cas9 and DNA as well as understanding the other risks associated with using such a disruptive technology, is key to developing CRISPR/Cas9 as a therapeutic agent. For current *ex vivo* protocols, the anticipated risk of exposing the subject to the CRISPR/Cas9 machinery is low due to long durations between expansion of desired-cell clone and subsequent administration, by which time the RNP complex is expected to have been degraded. However, shorter *ex vivo* incubation protocols may elevate risk of a subject being exposed to the active/inactive RNP complex, for example, potential for undesired editing *in-vivo* and/or immune system initiation, in the form of anti-cas9 antibodies or T-cell activation [70]. Strategies for direct administration of RNP complex would ultimately require greater demonstration of control to avoid exposure to the subject's immune system and/or non-target cell/DNA sequence.

Developing predictive modelling tools and performing experiments to understand the dose relationship between the amount of RNP complex and frequency of cells edited 'on target' over time is of particular importance in order to determine the level of efficacy and safety of using CRISPR/Cas9 gene editing.

The use of different editing strategies and delivery systems to direct the RNP complex to the nucleus of the target cell may influence bioanalytical requirements and potentially the number of analytical endpoints needed. A comprehensive assessment at both pre-clinical and clinical stages should occur to identify the intended use of the assay (e.g., determine biodistribution, activity of complex, off-target effects). Regulators' current expectations for gene editing therapeutic exposure/biodistribution data include the measurement of both total and active ribonucleoprotein complex and, where feasible, the testing of tissue biopsies in clinical studies. Screening of subjects for pre-existing Cas9-specific immune responses and monitoring immune responses following treatment particularly when gene editing components are present in the drug product is recommended.

Biodistribution

Evaluation of biodistribution is one of the key elements of the characterization of a gene therapy treatment. Typical methods used to evaluate distribution of the viral vector and the expression of the target gene include quantitative PCR and flow cytometry protocols. One needs to note that for a viral capsid vector-based modality, the vector genome detected during biodistribution evaluation may often be near the assay LOD or below. Other methods applied for detection of transgene protein product in tissues include western blot and immunoprecipitation mass spectrometry (IP-LCMS) analysis. For transgene protein analysis using an IP-LCMS platform, protein or/and peptide immunoprecipitation approaches can be conducted for the analyte enrichment prior to the LCMS step of analysis. Choosing among the two depends upon various considerations such as the target protein characteristics, availability of required specific reagents, tissue type and matrices – among others. Assay translatability between the preclinical and clinical settings is also an important factor that should be considered. Peptide IP-LCMS may provide opportunities in developing clinical/preclinical assays that are not possible with hybrid assays using anti-protein antibody reagents. This is particularly relevant when high quality anti-transgene protein antibody reagents are not available. Correlations between protein expression data from an IP-LCMS assay and methods designed to detect gene or mRNA transgene transcript levels (e.g., qPCR and/or flow cytometry) may ensure reliability of the data, however we are not aware whether such requests have been made by regulators. On the other hand, in many cases protein and transgene transcript may have different levels with dissimilar turnover characteristics and may not be correlated [71]. Regulators also don't expect platform based comparative values of transgene and transgene expression (qPCR, flow cytometry and hybrid LBA/LCMS), but a justification to explain potential disagreement may be required.

The current challenges related to development and validation of IP-LCMS assays include stability, reagent controls, requirements for sensitivity and analysis throughput. A new FDA guidance is available to address some of these concerns [72]. Generally, it is proposed to consider reducing the number of non-clinical biodistribution studies due to the limited utility of the data that are generated.

Immunogenicity

Immunogenicity

Gene therapy using viral vectors will require a careful assessment of immune responses to the vector components as well as the transgene protein [73,74]. For viral vectors, monitoring for viral capsid specific innate immune responses and capsid and transgene specific T- and B-cells may be required. Pre-existing immune responses to the viral proteins

should be detected as they may modulate the post-dose response and affect gene delivery and expression. Most often, this is done by evaluating presence of total antibody (tAb). Whether the information on development of tAb and neutralizing antibodies (NAb) against the viral vector and transgene protein is required should be determined on a case by case basis. Patient exclusion from a clinical trial for IV administered gene therapies may be based upon pre-existing Ab titer using binding or NAb assays. Pre-existing immunity in serum/plasma may have less relevance in the context of ocular gene therapies and in these cases is less likely to be assessed as inclusion criteria.

The transgene specific immune response may vary based on the prevalence of the endogenous protein, CRIM status of the patients and the site of transgene expression. The contribution due to the risk factors associated with gene delivery, patient's disease state and pharmacogenomics may also influence overall treatment efficacy. Lastly, the serotype of the viral vectors and delivery to an immune privileged vs systemic site will need to be a part of the overall immune monitoring strategy.

For the oncolytic class of viruses where the gene of interest is intended for the killing of cancer/tumor cells, an understanding of the mechanism of action would be key [75]. Even though oncolytic virus targets tumor cells directly and promotes killing through an activation of innate immune response or by expression of a transgene that can augment adaptive effector response, the viral capsid or transgene specific proteins can be exposed to periphery due to lysis of tumor cells resulting in an induction of an adaptive immune response.

The route of GTx delivery may play an important role in deciding whether there is a need for a detailed assessment of immune response including evaluation of pre-existing antibody in order to assess impact on treatment safety and efficacy. If a significant anti-viral vector antibody response post-dose is anticipated, and if the response prevents successful redosing, immune intervention may be required. Some strategies to modulate viral and transgene specific immune responses would include introduction of regulatory elements, codon optimization and CpG reduction. Additionally, to address re-administration in seropositive subjects, IgG removal, immune modulation and adjustments in dosing may be options [73]. If the presence of anti-transgene protein NAb is detected a possible safety risk of NAb impact similar to what has been observed for CRIM negative patients may need to be evaluated. Cell-based NAb assays are typically viewed as favorable as these provide functional information on NAb impact on cellular uptake of the GTx virus. Sensitivity expectations for these assays are similar to NAb assays developed for protein based biotherapeutics. Existing guidance for evaluation of immune responses against biotherapeutics [48,76–79] may be helpful to develop strategies for immunogenicity risk assessment for gene therapies however interpretation of results may differ, particularly in understanding boosted response as they may be due to the gene therapy or other environmental exposure.

Recommendations

Below is a summary of the recommendations made during the 13th WRIB:

1. It is important to ensure the relevance of all gene therapy bioanalytical testing using a risk-based approach that is discussed with the appropriate regulatory agency;
2. Banking of samples, when feasible, is recommended if it is unclear during early development what questions will need answers over the course of development of the therapeutic;
3. Patients with pre-existing anti-gene therapeutic immunity may be excluded from clinical trials or during treatment. Specific decision may depend on the type of the targeted tissue;
4. For qPCR methods to monitor CAR-T cells, the assay needs to be qualified and assay parameters including optimal primer length for specific detection of transgene should be evaluated;
5. For qPCR assay validation, the MIQE guidance [64] may be helpful in the validation study design. Sensitivity (50 copies/μg), precision and tissue extraction, expectation for LOD or LOQ (copy number), and controls to help ensure there is no cross-contamination should be investigated;
6. The utility of biodistribution studies should be evaluated, and in certain circumstances it may be appropriate to reduce the number of biodistribution studies that yield data with low utility;
7. For CAR-T programs, stability should be assessed for the sample storage conditions applied. Stability is also required on gDNA extracted from study samples;
8. Given the low number of study samples analyzed by PCR, ISR testing may not be relevant or feasible;
9. Viral shedding results may be requested for non-pathogenic vectors like AAV with the specific criteria for whether the test is needed are based on the long-term shedding profile information. Cells that have been *ex*

vivo modified may be excluded from the viral shedding requirements The necessity of viral shedding studies should be discussed with the appropriate regulatory agency;

10. The need for infectivity assays to assess for shedding should be based on the product related risk factors. The selected infectivity assay needs to be quantitative even when coupled with qPCR results;
11. Long term follow-up is required to assess delayed adverse events such as insertional mutagenesis or emergence of replication competent virus after gene therapy with products made using retroviruses. Banked samples can be used;
12. ELISpot may be used to monitor for cellular immunity, if used it should be developed using a risk-based approach factoring in the route of administration;
13. ELISpot results can be normalized (intra- and inter-subject), special accommodation should be made for particular sites taking into account known variabilities in the assay. Not every site may be required to conduct every analysis. Multiple baselines can be used;
14. ELISpot harmonization consortium White Paper [68] which outlines assay expectations and performance criteria may be helpful to develop ELISpot assays;
15. Developing predictive modelling tools and performing experiments to understand the dose relationship between the amount of RNP complex and frequency of cells edited 'on target' over time is of particular importance in order to determine the level of efficacy and safety of both *in vivo* and *ex vivo* gene therapies incorporating CRISPR/Cas9 genome editing;
16. Assessment of both the total and active ribonucleoprotein complex in the final drug product, and data on biodistribution/exposure should be collected;
17. Pre-existing immune responses to the viral proteins should be measured as they may modulate the post-dose response and affect gene delivery and expression. Both total antibody and neutralizing antibody tests have been used to date;
18. Whether the information on development of total (tAb) and neutralizing (NAb) antibodies against viral vector and transgene protein is required should be determined on a case by case basis;
19. Existing guidance for evaluating immune responses against biotherapeutics [48] may be helpful to develop strategies for immunogenicity risk-assessment for gene therapies; however, it should be noted that risks associated with gene therapies may be different from those for protein-based biotherapeutics.

Acknowledgments

- The US FDA, Europe EMA, UK MHRA, Brazil ANVISA, Health Canada, Japan MHLW, France ANSM and Norway NoMA for supporting this workshop;
- Booth B (US FDA), Fandozzi C (Merck & Co., Inc.), Evans C (GlaxoSmthKline), Pillutla R (Bristol-Myers Squibb), Kaur S (Genetech), Yu H (Boeringer Ingelheim), Garofolo F (Angelini Pharma), Schweighardt B (BioMarin), Liang M (AstraZeneca), Stevenson L (Biogen); Buonarati M (Inertek), Vitaliti A (Novartis), Litwin V (Caprion), Mehta D (Biogen), Piccoli S (GlaxoSmithKline), Gorovits B (Pfizer), Palmer R (Sanofi), Amavaradi L (Shire/Takeda), Beaver C (Syneos), Richards S (Sanofi), Rajadhyaksha M (Regeneron) for chairing the workshop and/or the White Paper discussions;
- The Biogen Team (Stevenson L, Purushothama S and Mehta D) for the pre-review of Part 3;
- Natasha Savoie (WRIB) for drafting the first draft of this document.
- All the workshop attendees and members of the Global Bioanalytical Community who have sent comments and suggestions to the workshop to complete this White Paper;
- Future Science Group as a trusted partner.

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 & 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. Stevenson L, Rocci M, Garofolo F *et al.* 2013 White Paper on recent issues in bioanalysis: 'hybrid' – the best of LBA & LCMS. *Bioanalysis* 5(23), 2903–2918 (2013).
7. Fluhler E, Hayes R, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 1 – small molecules by LCMS). *Bioanalysis* 6(22), 3039–3049 (2014).
8. Dufield D, Neubert H, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 2 – hybrid LBA/LCMS, ELN & regulatory agencies' input). *Bioanalysis* 6(23), 3237–3249 (2014).
9. Stevenson L, Amaravadi L, Myler H *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 3 – LBA and immunogenicity). *Bioanalysis* 6(24), 3355–3368 (2014).
10. Welink J, Fluhler E, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 1 – small molecules by LCMS). *Bioanalysis* 7(22), 2913–2925 (2015).
11. Ackermann B, Neubert H, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 2 – hybrid LBA/LCMS and input from regulatory agencies). *Bioanalysis* 7(23), 3019–3034 (2015).
12. Amaravadi L, Song A, Myler H *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 3 – LBA, biomarkers and immunogenicity). *Bioanalysis* 7(24), 3107–3124 (2015).
13. Yang E, Welink J, Cape S *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV) (Part 1 – small molecules, peptides and small molecule biomarkers by LCMS). *Bioanalysis* 8(22), 2363–2378 (2016).
14. Song A, Lee A, Garofolo F *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV): (Part 2 – Hybrid LBA/LCMS and input from regulatory agencies). *Bioanalysis* 8(23), 2457–2474 (2016).
15. Richards S, Amaravadi L, Pillutla R *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV): (Part 3 – LBA, biomarkers and immunogenicity). *Bioanalysis* 8(23), 2475–2496 (2016).
16. Welink J, Yang E, Hughes N *et al.* 2017 White Paper on recent issues in bioanalysis: aren't BMV guidance/guidelines 'scientific'? (Part 1 – LCMS: small molecules, peptides and small molecule biomarkers). *Bioanalysis* 9(22), 1807–1825 (2017).
17. Neubert H, Song A, Lee A *et al.* 2017 White Paper on recent issues in bioanalysis: rise of hybrid LBA/LCMS immunogenicity assays (Part 2: hybrid LBA/LCMS biotherapeutics, biomarkers & immunogenicity assays and regulatory agencies' inputs). *Bioanalysis* 9(23), 1895–1912 (2017).
18. Gupta S, Richards S, Amaravadi L *et al.* 2017 White Paper on recent issues in bioanalysis: a global perspective on immunogenicity guidelines & biomarker assay performance (Part 3 – LBA: immunogenicity, biomarkers and PK assays). *Bioanalysis* 9(24), 1967–1996 (2017).
19. Welink J, Xu Y, Yang E *et al.* 2018 White Paper on Recent Issues in Bioanalysis: 'A global bioanalytical community perspective on last decade of incurred samples reanalysis (ISR)' (Part 1 – small molecule regulated bioanalysis, small molecule biomarkers, peptides & oligonucleotide bioanalysis). *Bioanalysis* 10(22), 1781–1801 (2018).
20. Neubert H, Olah T, Lee A *et al.* 2018 White Paper on Recent Issues in Bioanalysis: focus on immunogenicity assays by hybrid LBA/LCMS and regulatory feedback (Part 2 – PK, PD & ADA assays by hybrid LBA/LCMS & regulatory agencies' inputs on bioanalysis, biomarkers and immunogenicity). *Bioanalysis* 10(23), 1897–1917 (2018).
21. Stevenson L, Richards S, Pillutla R *et al.* 2018 White Paper on Recent Issues in Bioanalysis: focus on flow cytometry, gene therapy, cut points and key clarifications on BAV (Part 3 – LBA/cell-based assays: immunogenicity, biomarkers and PK assays). *Bioanalysis* 10(24), 1973–2001 (2018).
22. ICH M10, Draft Bioanalytical Method Validation (2019). www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M10/M10EWG_Step2_DraftGuideline_2019_0226.pdf
23. Piccoli SP, Sauer JM. Points to Consider Document: Scientific and Regulatory Considerations for the Analytical Validation of Assays Used in the Qualification of Biomarkers in Biological Matrices Critical Path Institute (C-Path) (2019). https://c-path.org/wp-content/uploads/2019/06/EvidConsid-WhitePaper-AnalyticalSectionV20190621.pdf
24. CLSI Guideline H62: Validation of Assays Performed by Flow Cytometry. https://clsi.org/standards-development/documents-for-public-review/
25. US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. Guidance for Industry, Bioanalytical Method Validation. Rockville, MD, USA (2018).

26. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. BEST (Biomarkers, EndpointS, and other Tools) Resource. Silver Spring (MD): Food and Drug Administration (US); Bethesda (MD): National Institutes of Health (US); (2016).
27. 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).
28. Stevenson LF, Purushothama S. Parallelism: considerations for the development, validation and implementation of PK and biomarker ligand-binding assays. *Bioanalysis* 6(2), 185–198 (2014).
29. European Medicines Agency. *Committee for Medicinal Products for Human Use (CHMP). Reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples*. EMA/INS/GCP/532137/2010, London, UK (2012).
30. Swanson BN. Delivery of high-quality biomarker assays. *Dis. Markers* 18(2), 47–56 (2002).
31. O'Hara DM, Theobald V, Egan AC *et al.* Ligand binding assays in the 21st century laboratory: recommendations for characterization and supply of critical reagents. *AAPS J.* 14(2), 316–328 (2012).
32. Staack RF, Stracke JO, Stubenrauch K, Vogel R, Schleypen J, Papadimitrious A. Quality requirements for critical assay reagents used in bioanalysis of therapeutic proteins: what bioanalysts should know about their reagents. *Bioanalysis* 3(5), 523–534 (2011).
33. Kubiak RJ, Lee N, Zhu Y *et al.* Storage conditions of conjugated reagents can impact results of immunogenicity assays. *J. Immunol. Res.* 2016, Article ID 1485615 (2016).
34. Ohtake S, Kita Y, Arakawa T. Interactions of formulation excipients with proteins in solution and in the dried state. *Adv. Drug Deliv. Rev.* 63(13), 1053–1073 (2011).
35. King LE, Farley E, Imazato M *et al.* Ligand binding assay critical reagents and their stability: recommendations and best practices from the Global Bioanalysis Consortium Harmonization Team. *AAPS J.* 16(3), 504–515 (2014).
36. Garofolo W, Savoie N. The Decennial Index of the White Papers in Bioanalysis: 'A Decade of Recommendations (2007–2016)'. *Bioanalysis* 9(21), 1681–1704 (2017).
37. Green CL, Brown L, Stewart JJ, Xu Y, Litwin V, McCloskey TW. Recommendations for the validation of flow cytometric testing during dug development: I instruments. *J. Immunol. Methods* 363(2), 104–119 (2011).
38. O'Hara D, Xu Y, Lianz E, Reddy M, Wu D, Litwin V. Recommendations for the validation of flow cytometric testing during drug development: II assays. *J. Immunol. Methods* 363(2), 120–134 (2011).
39. Wood B, Jevremovic D, Béné MC, Yan M, Jacobs P, Litwin V. Validation of cell-based fluorescence assays: practice guidelines from the ICSH and ICCS – part V – assay performance criteria. *Cytometry B Clin. Cytom.* 84(5), 315–323 (2013).
40. Litwin V, Salkowitz-Bokal J, Steele P. Flow cytometry – new considerations for validating laboratory-developed tests. www.aacc.org/publications/cln/articles/2013/december/flow-cytometry
41. Brown L, Green C, Jones N *et al.* Recommendations for the evaluation of specimen stability for flow cytometric testing during drug development. *J. Immunol. Methods* 418, 1 (2015).
42. Du L, Grover A, Ramanan S, Litwin V. The evolution of guidelines for the validation of flow cytometric methods. *Int. J. Lab Hematol.* 37(Suppl. 1), 3–10 (2015).
43. Stewart JJ, Green CL, Jones N. Role of receptor occupancy assays by flow cytometry in drug development. *Cytometry B Clin. Cytom.* 90(2), 110–116 (2016).
44. Green CL, Stewart JJ, Högerkorp CM. Recommendations for the development and validation of flow cytometry-based receptor occupancy assays. *Cytometry B Clin. Cytom.* 90(2), 141–149 (2016).
45. Cossarizza A, Chang HD, Radbruch A *et al.* Guidelines for the use of flow cytometry and cell sorting in immunological studies. *Eur. J. Immunol.* 47(10), 1584–1797 (2017).
46. US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. *Draft Guidance for industry. Biomarker Qualification: Evidentiary Framework*. Rockville, MD, USA (2018).
47. Spidlen J, Moore W, Parks D *et al.* Data file standard for flow cytometry, version FCS 3.1. *Cytometry A* 77(1), 97–100 (2010).
48. US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. *Immunogenicity Testing of Therapeutic Protein Products – Developing and Validating Assays for Anti-Drug Antibody Detection*. Rockville, MD, USA (2019).
49. Shen M, Dong X, Tsong Y. Statistical evaluation of several methods for cut-point determination of immunogenicity screening assay. *J. Biopharm. Stat.* 25(2), 269–279 (2015).
50. Bourdage JS, Cook CA, Farrington DL, Chain JS, Konrad RJ. An Affinity Capture Elution (ACE) assay for detection of anti-drug antibody to monoclonal antibody therapeutics in the presence of high levels of drug. *J. Immunol. Methods* 327(1–2), 10–17 (2007).
51. Smith HW, Butterfield A, Sun D. Detection of antibodies against therapeutic proteins in the presence of residual therapeutic protein using a solid-phase extraction with acid dissociation (SPEAD) sample treatment prior to ELISA. *Regul. Toxicol. Pharmacol.* 49(3), 230–237 (2007).

52. Zoghbi J, Xu Y, Grabert R, Theobald V, Raichards S. A breakthrough novel method to resolve the drug and target interference problem in immunogenicity assays. *J. Immunol. Methods* 426, 62–69 (2015).

53. Wessels U, Zadak M, Reiser A *et al.* Immunogenicity testing of therapeutic antibodies in ocular fluids after intravitreal injection. *Bioanalysis* 10(11), 803–814 (2018).

54. Wessels U, Schick E, Ritter M, Kowalewsky F, Heinrich J, Stubenrauch K. Novel drug and soluble target tolerant antidrug antibody assay for therapeutic antibodies bearing the P329G mutation. *Bioanalysis* 9(11), 849–859 (2017).

55. Gorovits B, Wakshull E, Pillutla R, Xu Y, Manning MS, Goyal J. Recommendations for the characterization of immunogenicity response to multiple domain biotherapeutics. *J. Immunol. Methods.* 408, 1–12 (2014).

56. Andisik M, DeStefano L, Stefan C *et al.* Effects of blood processing and sample storage on the stability of biotherapeutics and anti-drug antibodies. *Bioanalysis* 7(11), 1325–1335 (2015).

57. Michaut L, Laurent N, Kentsch K, Spindeldreher S, Deckert-Salva F. Stability of anti-immunotherapeutic antibodies in frozen human serum samples. *Bioanalysis* 6(10), 1395–1407 (2014).

58. Pihl S, Michaut L, Hendriks J *et al.* EBF recommendation for stability testing of anti-drug antibodies; lessons learned from anti-vaccine antibody stability studies. *Bioanalysis* 6(10), 1409–1413 (2014).

59. Schneider AK, Vainshtein I, Roskos LK, Chavez C, Sun B, Liang M. An immunoinhibition approach to overcome the impact of pre-existing antibodies on cut point establishment for immunogenicity assessment of moxetumomab pasudotox. *J. Immunol. Methods* 435, 68–76 (2016).

60. General Chapter 1106. Immunogenicity Assays – Design and Validation of Assays to Detect Anti-Drug Neutralizing Antibody, U.S. Pharmacopeia and the National Formulary (Pharmacopeial Forum).40(3) (2014).

61. Gorovits B, Clements-Egan A, Birchler M *et al.* Pre-existing antibody: biotherapeutic modality-based review. *AAPS J.* 18(2), 311–20 (2016).

62. Kubiak RJ, Zhang L, Zhang J *et al.* Correlation of screening and confirmatory results in tiered Immunogenicity testing by solution-phase bridging assays. *J. Pharm. Biomed. Anal.* 74, 235–245 (2013).

63. Kubiak RJ, Zhang J, Ren P, Yang H, Roskos LK. Excessive outlier removal may result in cut points that are not suitable for immunogenicity assessments. *J. Immunol. Methods* 463, 105–111 (2018).

64. Bustin SA, Benes V, Garson JA *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55(4), 611–22 (2009).

65. CLSI. CLSI.org

66. US Department of Health and Human Services, US FDA, Center for biologics evaluation and research. Guidance for industry – design and analysis of shedding studies for virus or bacteria-based gene therapy and oncolytic products. www.fda.gov/media/89036/download

67. US Department of Health and Human Services, US FDA, Center for biologics evaluation and research. Draft guidance for industry – human gene therapy for retinal disorders. www.fda.gov/media/124641/download

68. Janetzki S, Panageas KS, Ben-Porat L *et al.* Results and harmonization guidelines from two large-scale international Elispot proficiency panels conducted by the Cancer Vaccine Consortium (CVC/SVI). *Cancer Immunol. Immunother.* 57(3), 303–315 (2008).

69. Feyssaguet M, Bellanger A, Nozay F *et al.* Comparison between a new multiplex electrochemiluminescence assay and the WHO reference enzyme-linked immunosorbent assay to measure serum antibodies against pneumococcal serotype-specific polysaccharides. *Vaccine* 37(16), 2208–2215 (2019).

70. Charlesworth CT, Desphande PS, Dever DP *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* 25, 249–254 (2019).

71. Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. *Cell* 165(3), 535–550 (2016).

72. US Department of Health and Human Services, US FDA, Center for biologics evaluation and research. *Draft guidance for industry – testing of retroviral vector-based human gene therapy products for replication competent retrovirus during product manufacture and patient follow-up*. www.fda.gov/media/113790/download

73. Colella P, Ronzitti G, Mingozzi F. Emerging issues in AAV-mediated *in vivo* gene therapy. *Mol. Ther. Methods Clin. Dev.* 1(8), 87–104 (2017).

74. Vandamme C, Adjali O, Mingozzi F. Unraveling the complex story of immune responses to AAV vectors trial after trial. *Hum. Gene Ther.* 28(11), 1061–1074 (2017).

75. Russell SJ, Kah-Whye P. Oncolytic virotherapy: a contest between apples and oranges. 25(5), 1107–1116 (2017).

76. ICH Considerations: Oncolytic virus. www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-22.pdf

77. Guidance for industry: design and analysis of shedding studies for virus or bacteria-based gene therapy and oncolytic products. www.fda.gov/regulatory-information/search-fda-guidance-documents/design-and-analysis-shedding-studies-virus-or-bacteria-based-gene-therapy-and-oncolytic-products

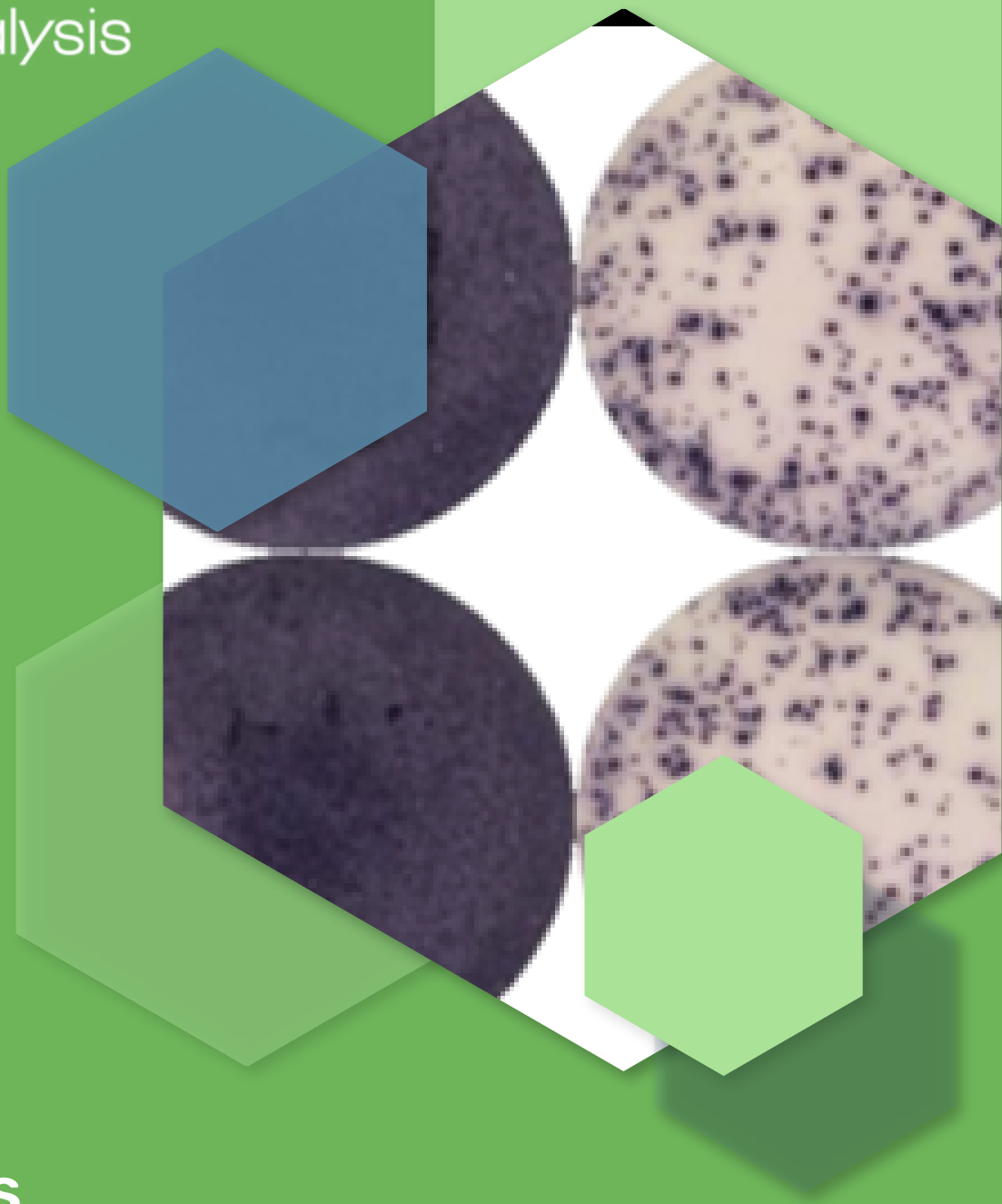
78. Guidance for industry: gene therapy clinical trials – observing subjects for delayed adverse events. www.fda.gov/regulatory-information/search-fda-guidance-documents/gene-therapy-clinical-trials-observing-subjects-delayed-adverse-events

79. ICH Considerations: general principle to address virus and vector shedding 2009: EMEA/CHMP/ICH/449035/2009. www.ema.europa.eu/en/ich-considerations-general-principles-address-virus-vector-shedding

80. Clinical Laboratory Improvement Amendments (CLIA). www.fda.gov/medical-devices/ivd-regulatory-assistance/clinical-laboratory-improvement-amendments-clia

81. CLIA related Federal Register and code of Federal Regulation Announcements. /web.archive.org/web/20090228074930/http://www.fda.gov/cdrh/CLIA/CLIAfedregin.html

Bioanalysis ZONE



Contact us

Editorial department
editor@bioanalysis-zone.com

Business Development and Support
advertising@future-science-group.com

This supplement is brought to you by Bioanalysis Zone in association with:

eurofins | BioPharma Services