

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

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)



‡ 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:
fabio garofolo@hotmail.com

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

³⁷Norwegian Medicines Agency, Oslo, Norway³⁸GlaxoSmithKline, Rockville, MD, USA³⁹REGENXBIO, Rockville, MD, USA⁴⁰AstraZeneca, Molndal, Sweden⁴¹Merck & Co., Inc., Kenilworth, NJ, USA⁴²Novartis, Cambridge, MA, USA⁴³Merck & Co., Inc., West Point, PA, USA⁴⁴CRISPR Therapeutics, Cambridge, MA, USA⁴⁵Janssen 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

- Introduction
- SECTION 1 – New Insights in Biomarker Assay Validation (BAV)
 - Discussion Topics
 - Discussions, Consensus & Conclusions
 - Recommendations
- SECTION 2 – Current & Effective Strategies for Critical Reagent Management
 - Discussion Topics
 - Discussions, Consensus & Conclusions

- Recommendations
- SECTION 3 – Flow Cytometry Validation in Drug Discovery & Development & CLSI H62
 - Discussion Topics
 - Discussions, Consensus & Conclusions
 - Recommendations
- SECTION 4 – Interpretation of the 2019 FDA Immunogenicity Guidance
 - Discussion Topics
 - Discussions, Consensus & Conclusions
 - Recommendations
- SECTION 5 – *In vivo* & *Ex vivo* Gene Therapy & Vaccine Bioanalytical Challenges
 - Discussion Topics
 - Discussions, Consensus & Conclusions
 - Recommendations
- References

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:	Good Practices where x = clinical, laboratory or manufacturing
GTx:	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:	<i>In vitro</i> 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 operating procedure
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 Piccoli¹, Devangi Mehta², Alessandra Vitaliti³, John Allinson⁴, Shashi Amur⁵, Steve Eck⁶, Cherie Green⁷, Michael Hedrick⁸, Shirley Hopper⁹, Allena Ji¹⁰, Alison Joyce¹¹, Virginia Litwin¹², Kevin Maher⁵, Joel Mathews¹³, Kun Peng⁷, Afshin Safavi¹⁴, Yow-Ming Wang⁵ & Yan Zhang⁸

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 Amaravadi¹⁵, Nisha Palackal²², Sai Thankamony⁸, Chris Beaver¹⁶, Eris Bame², Thomas Emrich¹⁷, Christine Grimaldi¹⁸, Jonathan Haulenbeek⁸, Alison Joyce¹¹, Vellalore Kakkanaiah¹⁹, David Lanham²⁰, Kevin Maher⁵, Andrew Mayer²¹, Paul C Trampont²³ & Laurent Vermet²⁴

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 Vitaliti³, Virginia Litwin¹², Devangi Mehta², Eris Bame², Naveen Dakappagari²⁵, Steve Eck⁶, Catherine Fleener²⁶, Fabio Garofolo²⁷, Cherie Green⁷, Michael Hedrick⁸, Vellalore Kakkanaiah¹⁹, David Lanham²⁰, Kevin Maher⁵, Steven Piccoli¹, Cynthia Rogers²⁸, Shabnam Tangri²⁵, Paul C Trampont²³ & Yuanxin Xu²⁹

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 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 Liang³⁰, Manoj Rajadhyaksha²², Susan Richards¹⁰, Becky Schweighardt³¹, Shobha Purushothama², Daniel Baltrukonis³², Jochen Brumm⁷, Elana Cherry³³, Jason Delcarpini³⁴, Carol Gleason⁸, Susan Kirshner⁵, Robert Kubiak⁶, Luying Pan³⁵, Michael Partridge²², João Pedras-Vasconcelos⁵, Qiang Qu³⁶, Venke Skibeli³⁷, Therese Solstad Saunders³⁷, Roland F Staack¹⁷, Kay Stubenrauch¹⁷, Al Torri²², Daniela Verthelyi⁵ & Haoheng Yan⁵

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 defensible 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 NAb in oncology, rare diseases, or infectious diseases? If we have PK, ADA and a good PD marker, what additional information about clinical relevance do NAb samples 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 led 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 Gorovits¹¹, Rachel Palmer¹⁰, Mark Milton⁴², Brian Long³¹, Bart Corsaro³⁸, Vahid Farrokhi¹¹, Michele Fiscella³⁹, Neil Henderson⁴⁰, Vibha Jawa⁴¹, Jim McNally³⁵, Rocio Murphy⁴³, Cynthia Rogers²⁸, Shabnam Tangri²⁵, Hanspeter Waldner⁴⁴, Yuanxin Xu²⁹ & Tong-Yuan Yang⁴⁵

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 Feysaguet *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/ μ g), 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*

- in 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 (GlaxoSmithKline), Pillutla R (Bristol-Myers Squibb), Kaur S (Genetech), Yu H (Boeinger 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, Schleyen J, Papadimitriou 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 drug 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, Mingozi 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, Mingozi 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/s-earch-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](http://web.archive.org/web/20090228074930/http://www.fda.gov/cdrh/CLIA/CLIAfedregin.html)