Ligand Binding Assays: Summary and Consensus from the Bioanalytical Workshop (CC V)

Binodh DeSilva

Executive Director Immunochemistry and Biomarker Development

Bristol-Myers Squibb

Contributors

- Lauren Stevenson
- Sherri Dudal
- Boris Gorovitz
- Russell Weiner

Validation Parameters: Selectivity

- Selectivity section includes: (a) Interferences from substances physiochemically similar to the analyte, and (b) Matrix effects
  - For LBAs:
    - (a) Interferences from substances physiochemically similar to the analyte = Specificity of the reagents
    - (b) Matrix effects= selectivity of the assay
  - Suggest clarification and separation of specificity and selectivity assessments
    - Agreed

Validation Parameters: Selectivity

- Specificity
  - Cross-reactivity
    - Noting that small molecule concomitant meds are not physiochemically similar to large molecule therapeutics
  - Other potentially interfering substances (e.g. soluble target, ADA)
- Selectivity
  - Matrix effects

Validation Parameters: Selectivity (Specificity)

- Interferences from substances physiochemically similar to the analyte
  - "When possible, the LBA should be compared with a validated reference method (such as LC-MS) using incurred samples and predetermined criteria to assess the accuracy of the method" (lines 522-524)
    - Implication that orthogonal methods for large molecule bioanalysis are somehow superior to LBAs is erroneous
    - For a typical proprietary large molecule PK assay, there is no absolute reference method other than the chosen validated method for the analyte

Validation Parameters: Selectivity (Specificity)

- Interferences from substances physiochemically similar to the analyte(cont. from previous slide)
- Consensus:
  - Intent was to address interference issues
  - Sentence may be removed, or if retained, modified to indicate that if an interferent is identified, an orthogonal method (free from the interference) may be considered.
  - This may be addressed in method development where it would guide final method selection.

Validation Parameters: Selectivity (Matrix Effects)

- The calibration curve in biological fluids should be compared with calibrators in buffer to detect matrix effects…" (lines 530-531)
  - PK assay calibrators are typically in matrix, effects of which are addressed during method development (not a validation parameter)
    - Agreed
- Parallelism of diluted study samples should be evaluated with diluted standards to detect matrix effects (532-533)
  - Parallelism not standard for detection of matrix effects in PK assays as incurred samples typically not available at pre-study validation
  - For PK assays, testing 10 or more individual matrix samples unspiked and spiked at LLOQ (and HQC) levels is appropriate (DeSilva 2003; EMA 2012)
    - Agreed

Parallelism?

- When (and how) to perform parallelism assessments was not discussed
  - Context in draft guidance was in reference to matrix effects only
  - Industry perspective
    - Parallelism not routinely performed for PK LBA
    - Used as an investigation tool
    - See GBC Harmonization Team L2 white paper (AAPS J. Dec 2013)

Validation Parameters: Accuracy, Precision and Recovery

- Approaches suggested for A&P differ from EMEA as well as whitepaper (De Silva et al 2003)
- The following need clarification:
  - Validation samples – 3 concentrations levels required for A&P (line 540)
    - Should also include LLOQ and ULOQ (total of 5 QCs) to validate the full range of the curve
    - Agreed
  - A&P should be measured using a minimum of five determinations per concentration (lines 539-540)
    - Unclear how many A&P runs should be performed (5?) or how many sets of spiked samples should be included in each run
    - Intra-assay precision: should be assessed in one single run with 5 determinations, or multiple runs and each run includes 3 sets per concentration level?
    - Agreed: 6 runs with 3 sets per run would align with current best practice recommendations for both intra- and inter-assay

Validation Parameters: Accuracy, Precision and Recovery

- A&P should be demonstrated for high concentrations diluted into the range of quantitation (lines 554-555)
  - Two assay characteristics are entangled in this directive: Dilutional linearity and A&P
  - Dilutional Linearity should be separated and also include prozone assessment
    - Agreed
- Total Error should be calculated from QCs in A&P assessment not from Standards
  - Agreed
  - Aligned with CC III recommendations (30%; 40% at LLOQ)

Validation Parameters: Calibration Curve

- Method validation experiments should include a minimum of 6 runs conducted over several days with at least 6 concentrations (including LLOQ, low, medium and high)….(lines 575-577)
  - Clarification needed: although LLOQ and ULOQ are standard calibrator concentrations, the other levels mentioned seem to refer to QC levels
  - Low, mid, high QC concentrations should not coincide with calibrator concentrations
- Consensus
  - The 6 concentrations span the quantitative assay range, LLOQ-ULOQ (does not include anchor points outside assay range although anchor points may be included for curve fitting purposes)
  - Generally speaking, the number of standard calibrator points need to be adequate to support the curve fit model

Validation Parameters: Calibration Curve

- LLOQ and ULOQ
  - Should be demonstrated independently as QCs in A&P, not defined solely by standard curve performance
    - Agreed
- Total Error
  - Belongs in A&P and should be assessed with QCs
    - Agreed

Sample Analysis

- All study samples from a single subject should be analyzed in a single run (Line 730)
  - For PK LBAs, this would require a large variety of sample dilutions to be incorporated into a single run which can increase likelihood of dilution errors and compromise data quality
  - Assay A&P demonstrates suitability of testing samples from a single subject across multiple runs
- Consensus
  - Understood that it is not always possible, practical or scientifically preferable
  - Do when practical and scientifically appropriate

Sample Analysis: Standard Curve Range

- The calibration (standard) curve should cover the expected study sample concentration range (lines 703-704)
  - Calibration curves for LBAs are frequently narrow (1-3 logs) and cannot cover study sample range
  - Instead, sample dilution is employed and dilutional linearity is validated pre-study
- Consensus:
  - statement will be removed from the guidance

Sample Analysis: Placement of Quality Controls

- High QC should be placed at high end of the range of the expected study sample concentrations (lines 618-619)
  - For LBAs, QCs should be placed in reference to the assay range, not study sample concentration range, as samples are diluted into range
    - Agreed

Sample Analysis: Placement of Quality Controls

- If the study sample concentrations are clustered in a narrow range of the standard curve, additional QCs should be added in the sample range (Line 725)
  - QCs in LBAs are already generally relatively “close” to each other
  - If sample results cluster around mid-range of the assay, it simply indicates that sample dilutions were well chosen
  - May be a consideration for some biomarker assays and small molecule assays – recommend removing from LBA section
- Incomplete Consensus
  - LBA industry position unchanged (see above)
  - General agreement that most LBAs (typical assay ranges of 2 logs) would not require additional QCs
  - Some agreement that bulk of sample results should be bracketed by 2 QCs – no agreement on “how close” the bracketing QCs should be (for assays with larger dynamic ranges)
  - For diluted samples, which comprise the bulk of samples analyzed by LBA, acknowledged that adjustments in dilution factors can help ensure that sample results are derived from entire curve range
  - Suggestion – sharing of data sets that speak to the utility of additional QCs would be informative to help determine when/if such practice would be warranted

Biomarker Analysis

Focus on Key Issues

- Non-FDA approved biomarker assays
  - In-house developed, commercial kits, research use only (RUO), lab developed tests (LDTs)
- Definitive Quantitative
- Relative Quantitative
  - Dealing with QCs
- FDA approved Dx kits

Guiding Principle

- When biomarker data will be used to support EFFICACY (not safety) or labeled dosing instructions, it is critical to ensure the integrity of the data and therefore the assay should be fully validated
- Industry recommendation:
  - When data will be used to support pivotal determination of effectiveness or label dosing instructions, utilize a fit-for-purpose approach for BOTH assay validation and for setting assay acceptance criteria

Non-FDA Approved Biomarker Assays

- It is recognized that obtaining appropriate biomarker reference standard or blank matrix may not always be possible. Additionally, reference standard may not be identical to the endogenous biomarker. As a consequence, developing and validating the biomarker assay CANNOT always be done to meet the standards of a PK assay (refer to the Lee 2006 biomarker white paper)

Non-FDA Approved Biomarker Assays

Pharmaceutical Research, Volume 23, No. 2, February 2006 (© 2006)
DOI: 10.1007/s11095-005-9045-3

Research Paper

Fit-for-Purpose Method Development and Validation for Successful Biomarker Measurement

Jean W. Lee,[1,16,17] Viswanath Devanarayan,[2] Yu Chen Barrett,[3] Russell Weiner,[3] John Allinson,[4] Scott Fountain,[5] Stephen Keller,[6] Ira Weinryb,[7] Marie Green,[8] Larry Duan,[9] James A. Rogers,[10] Robert Millham,[10] Peter J. O'Brien,[11] Jeff Sailstad,[12] Masood Khan,[13] Chad Ray,[14] and John A. Wagner[15]

- Industry Recommendations
  - Definitive Quantitative: When appropriate reference standard AND blank matrix are available the assay should be validated to meet the same standards as a PK assay

Non-FDA Approved Biomarker Assays

- Relative Quantitative: When appropriate reference standard AND blank matrix are NOT available
  - The assay should be validated and assay acceptance criteria should be set in a fit-for-purpose manner and DO NOT need to meet ALL of the same standards as a PK assay. However, every effort should be made to evaluate critical assay parameters (eg. standard curve performance, stability, specificity, parallelism, etc.)
  - QC samples may not be able to be prepared with known concentrations. Therefore, QC ranges may be prepared by pooling L/M/H biomarker containing matrix and used as analytical QCs for assay run acceptance

Approved Diagnostic (Dx) Kits

- It is recognized that Dx kits are likely not approved for our intended use and may lack critical information
- Industry Position
  - Supplement the lab's validation data by conducting additional assay verification experiments in a fit-for-purpose manner. This can include evaluations of stability, specificity, drug interference, parallelism, standard curve performance, etc.
  - Assay Verification- Proof of analytical performance of critical assay characteristics in your lab (evaluate the fit-for-purpose assay performance)

Approved Dx Kits

- Industry Position (continued)
  - Dx assays with sparse calibration standards should include additional verification experiments using additional standards to evaluate the calibration range and standard curve performance.
  - The goal of conducting these additional experiments is to generate a sufficient body of evidence that support using the Dx assay as intended by the Dx manufacture WITHOUT modification.

LBA: New technology/platform

LBA: New technology/platform

- Line 887-889: Technology /platforms used for large molecule evaluation
  - A technology /platform is a specific method that may involve a new instrument for measuring a large molecule analyte.
  - A NEW technology/platform refers to a non-LBA or LBA method for measuring large molecules that has not yet been used for regulatory submission.
  - Technology /platforms that have been used for regulatory submission are considered suitable for general use and no longer NEW.

Large molecule analytes

- LBA measurement of large molecules involves the recognition of a protein epitope. In this case, the analyte may or may not be bound to another protein.
- Immunoassays are based on affinity interactions and new methods may not be based on this interaction for analyte measurement so differences in absolute measurements may occur. Ex. Changes in Ab capture reagent can lead to different absolute analyte measurements.
- Therefore, the analyte species measured may not be the same between large molecule methods.
- The key is to be able to correctly interpret PK/TK data used in regulatory studies.

Cross Validation

- Line 127-129 : A reference is made to using LC-MS/MS as a reference method in the cross validation section.
  - When two different methods are used such as LC-MS/MS and ELISA, they may not be measuring the same analyte species
  - the former involves quantification based on similar physico-chemical properties and the latter involves affinity binding.
  - Ex. ‘Free’ ELISA method versus ‘total’ analyte method such as is the case with LC-MS/MS.

Industry perspective:

- Change terminology to bridging between methods in a program or study, using spikes and / or real samples if available

Cross validation of a new method

- Line 119-131: Cross-validation is deemed necessary to compare and validate a change in platform. Is this always relevant?
  - Is cross-validation really necessary if it is validated on its own?
  - Should a cross-validation be used only to compare a change in method during development in the same species and matrix? Is this always comparable?
  - If a new technology is deemed significantly different and measures a specific form of analyte, or functions differently in terms of binding kinetics or other eventual differences, but validated on its own, it can be implemented for bioanalysis.
- Industry perspective:
  - If you start with a new technology and use it for the following studies in a program, this is acceptable -agreed
  - If there is a new technology introduced after a previous one, a cross-validation would be required -agreed

Analytical run definition

- Line 548-552: reference to single analytical run. A single analytical run has not been defined and leaves it open for implementation of platforms and new technologies.
  - A run could be defined as a continuous series of samples consisting of a standard curve, sets of QCs and unknown samples and can involve multiple solid supports.
  - LBAs tend to limit a run to a solid support ie plate or CD and this is not always warranted especially if inter-solid support differences are deemed to be low via QC checks.
  - Opinion on a single analytical run? Length?
- Industry perspective:
  - There is no consensus on the batch size –left open.
  - Analytical run begins with a set of calibrators and has QCs on each support that is used during the run. --agreed

Analytical run definition

- Line 626-627 & 722-724: also supports longer single analytical runs
  - Newer technologies may allow for longer runs with interspersed QC sets so a rule of 5% of the number of unknown samples is useful for sample analysis.
- Industry perspective:
  - A rule of using a sets of QCs that equals 5% of the total unknown samples fits with the longer the ‘96well’ runs that may be implemented with newer technology or different formats such as 384well plates…etc. -agreed

Calibration curve: multiple analytes

- Line 572-573: A calibration curve should be generated for each analyte in the sample. Further clarification is required as different cases could exist.
  - Industry perspective: a standard curve must be run for each analyte for combination studies and each method validated individually. Lack of interference should be verified between each analyte. agreed
  - For technologies that involve multiplexing, each analyte must be validated separately and mixed in the same batch.-agreed
  - Will add multiplexing to biomarkers section
- Line 740-741: Samples involving multiple analytes should not be rejected based on the data from one analyte failing the acceptance criteria. --useful for future technologies
  - Add: The failed analyte can be re-run under the same conditions and the other analytes that previously passed, not considered.
  - agreed

Automation with technology

- Line 647-648: re-injection
  - Remove reinjection reproducibility sentence
- Line 731-732: Carry-over
  - Where semi-automated or automated platforms are used, carry-over should be eliminated during method development (ie by applying stringent washing conditions).
  - If disposable tips are used, this is not applicable.
  - Leave as is and carry-over evaluation will apply only if applicable

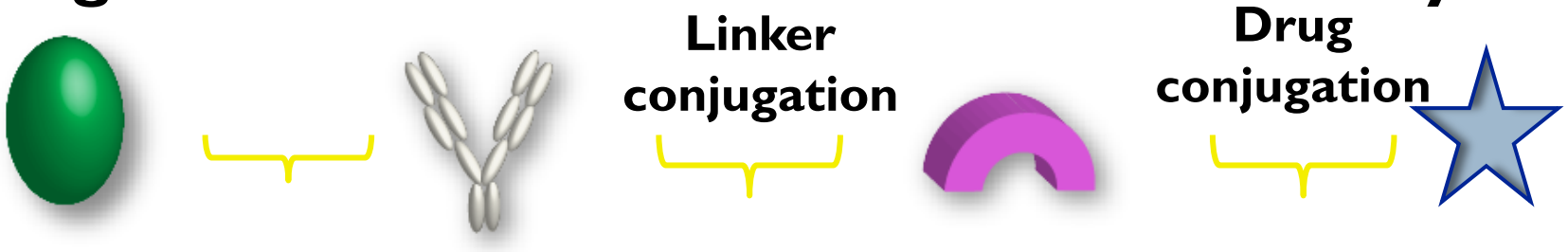
Duplicate analysis in LBAs

- Unknown samples are usually run in duplicate and some technologies /platforms or advances in LBAs can allow for singlet analysis.
  - If the precision of the QCs during method validation are within an acceptable range, the unknown samples can be run in singlets.
  - This acceptance range can be evaluated during method validation statistically where all QCs are run in duplicate.
  - As LC-MS/MS techniques run samples in singlets with an accepted precision of 15%, the required precision can be the same or debated for LBAs. The L3 team submitted a survey to the LBA community and generally, an acceptance range of 5-10% was comfortable. Why not 15%?
  - During sample analysis, the QC samples should be run in duplicate to assure this acceptance range. ISR describes the reproducibility of measuring samples.
- Agreement on the concept of running singlet analysis for LBAs if the method allows this and the number of replicates can be driven by the data during validation. This can be based on inter-plate precision and scientific rational.
- Two sets of QCs run as singlets in addition to a calibrator curve at the beginning of the run.

What is an Antibody Drug Conjugate?

Combine selectivity and antitumor activity of a monoclonal Ab with the potency of a small molecular weight drug (cytotoxin)

Targets **Vehicles** **Linkers** **Payloads**



Antibody

- Targeted recognition

Linker

- Stable in plasma
- Labile upon internalization to release drug

Drug

- Highly potent

ADCs: Key considerations

- ADCs are hybrid molecules, in many cases with varying numbers of small molecule drugs (the payload) conjugated to a large protein carrier molecule (typically a mAb).
- ADCs are produced as heterogeneous mixtures of species with a distribution of drug-antibody ratio (DAR) values, depending on the actual sites of drug conjugation to the protein moiety (driven by the conjugation chemistry)
- The heterogeneity of ADC product material is generally well characterized prior to dosing in studies

ADCs: Key considerations (contd)

- However, the heterogeneity of the ADC is known to dynamically change in vivo due to spontaneous, induced deconjugation (e.g. exposure to intra-lysosomal enzymes or low pH)
- Multiple analytes, associated with or derived from the ADC, may be assessed to obtain an appropriate description of ADC exposure in vivo
- The ultimate <clinical study> goal is to determine Exposure-Response Relationships for efficacy and safety signals. Analytes that relate to efficacy may be different from analytes that relate to safety

Analytes Commonly Assessed for ADC Bioanalysis

Analyte type	Analyte(s) Details	Typical Analytical Method(s)
Conjugated Antibody*	Antibody with minimum of DAR equal or greater > 1	LBA
Total Antibody	Conjugated, partially unconjugated and fully unconjugated (DAR equal or > 0)	LBA or Hybrid LC-MS/MS
Antibody-Conjugated Drug*	Total small molecule drug conjugated to antibody	Hybrid LC-MS/MS
Unconjugated Drug	Small molecule drug not conjugated to antibody	LC-MS/MS
Anti-ADC Antibody (anti-therapeutic Antibody (ATA))	Antibodies directed against antibody component of ADC, linker or drug (binding/neutralizing)	LBA

General ADC relevant considerations

- Two most common questions
  - What to measure - refer to Kaur et al 2013, Gorovits et al 2013
  - What are the acceptance criteria to use
    - In general, assays applied for ADCs can be in alignment with the current BMV draft but some special cases exist
    - Regular acceptance criteria may apply although several small molecule (LCMS/MS platform) and large molecule (LBA platform) considerations are presented in the following slides

ADC: LCMS discussion topics
Assay acceptance criteria

- Adjust acceptance criteria for protein LC/MS assays that involve complex sample preparations (e.g. affinity enrichment, protease digestion) to the following:
  - Lines 212-214: Accuracy: The mean value should be within 20% of the nominal value except at LLOQ, where it should not deviate by more than 25%.
  - Lines 221-223: The precision determined at each concentration level should not exceed 20% of the coefficient of variation (CV) except for the LLOQ, where it should not exceed 25% of the CV . Other criteria may be acceptable based on the assay performance data obtained during method validation.
- Industry position
  - specific assay acceptance criteria based on the assay performance data as generated during assay validation may be applied

ADC: LCMS discussion topics
What does Stability mean for ADCs

- Stability of active catabolites released from ADC (e.g. unconjugated drug / linker-drug) should be tested in spiked matrix as is typically done for conventional therapeutics.
- Stability of ADC catabolites may need to be evaluated in the presence of the parental ADC with respect to potential release of catabolites during post-collection storage and handling of samples.
- Such evaluations are relevant to the unconjugated drug and antibody-conjugated drug analytes.
- Due to the very low expected concentration of ADC catabolites relative to the intact ADC, wider acceptance criteria, based on validation data, may be required for corresponding stability tests.

ADC: LCMS discussion topics. Acceptance criteria for stability tests

- Current language, Line 359: Stability sample results should be within 15% of nominal concentrations
- Industry position
  - Decision about acceptable range of variability for analyte stability should be scientifically justified, based on assay validation data and with consideration of how the data will be used. As an example, flexibility may be required for complex modalities with intrinsic instability and / or LC-MS assays that involve complex sample preparations (e.g. affinity enrichment, protease digestion)

ADC: LBA discussion topics. Reference material

- Reference material representing parent therapeutic should be used to prepare assay standards and QC solutions. E.g. Antibody Drug Conjugate reference material should be used to prepare standard and QC solutions for Total Antibody, Conjugated Antibody and Antibody Conjugated Drug analytical tests. Currently there is no definition of Reference standard in the LBA section
- Industry position
  - Reference standard (substance): A well-characterized, traceable, batch of material of known purity and concentration, accompanied by a Certificate of Analysis or similar documentation. Reference Standard is required for validation of a bioanalytical procedure. Ideally, the reference standard should be as similar as possible to the measured analyte (e.g. with respect to drug antibody ratio distribution and unconjugated drug content for ADCs)

ADC: LBA discussion topics. Acceptance criteria for ADC LBA assays

- Current language, Line 538:
  Accuracy is determined by replicate analysis of samples containing known amounts of the analyte (QCs). …The mean value should be within 20% of the actual value except at LLOQ, where it should not deviate by more than 25%.
- Industry position

➢ The acceptable accuracy range for complex modalities and / or analytical methods requiring complex sample preparation or procedural steps may need to be adjusted from current LBA criteria. Final criteria for assay application to regulated studies should be based on the assay performance data obtained during method validation.

Acknowledgements

- AAPS Bioanalytical Focus Group
- AAPS Ligand Binding Assay Focus Group
- Other AAPS Focus Groups and Section representation
- Global Bioanalysis Consortium
- European Bioanalysis Forum
- Various industrial input