Assessing the Suitability of Immunoassay Cut Points from Validation for Use In-Study

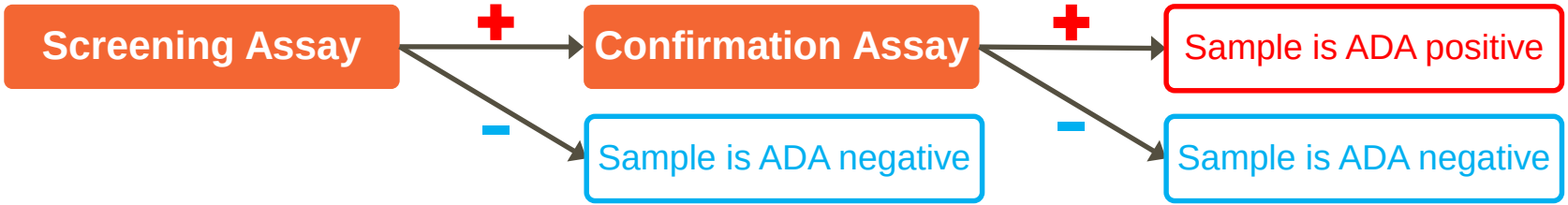
Andrew Gehman
GlaxoSmithKline

Nonclinical Biostatistics Conference
18 June 2019

Anti-Drug Antibody (ADA) Immunoassays

Background

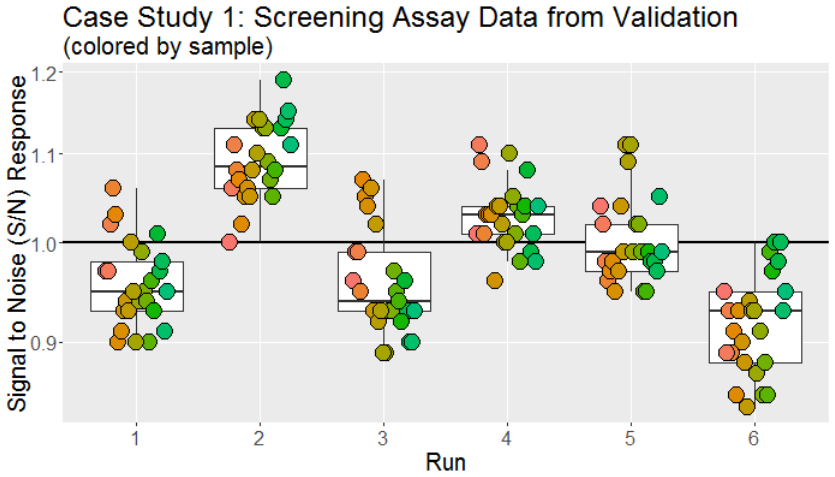
- Plate-based assays
- Assess immune response to biologics
- Tiered approach: screening and confirmation assays use validated cut points



ADA Assay Validation

Experimental Design

- Validation design includes both biological and analytical sources of variability:
  - ≥ 25 drug-naïve pre-clinical samples (≥ 50, if clinical)
  - Each sample tested in ≥ 6 runs
    - By ≥ 2 analysts over ≥ 3 days



Case Study 1: 25 samples in 6 runs			
	Day 1	**Day 2**	**Day 3**
Analyst A	Run 1	Run 3	Run 5
Analyst B	Run 2	Run 4	Run 6

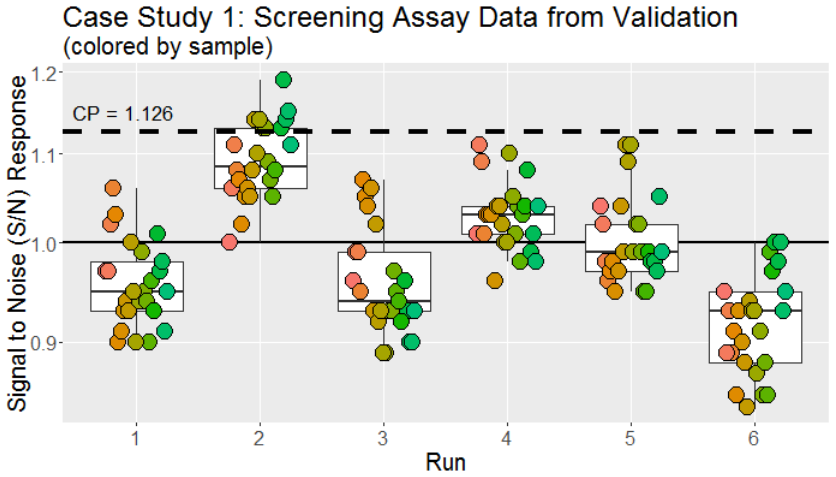
"All studies were conducted in accordance with the GSK Policy on the Care, Welfare and Treatment of Laboratory Animals and were reviewed the Institutional Animal Care and Use Committee either at GSK or by the ethical review process at the institution where the work was performed."

Cut Point from Assay Validation

Screening Assay Cut Point

- Cut point calculated to achieve a 5% false positive rate

- Variety of methods for estimating the 95th percentile:
  - Mean and SD
  - Median and MAD
  - Non-parametric
  - Lower confidence bound



$$y_{ij} = \mu + \delta_i + \tau_j + \varepsilon_{ij}$$

Run i, Sample j

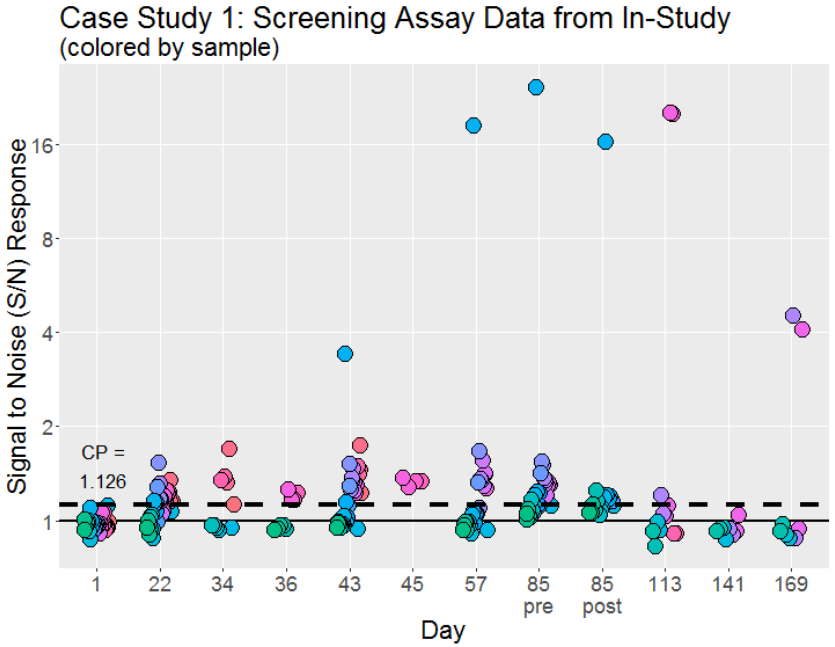
$$\delta_i \sim N(0, \sigma_R^2),\ \ \tau_j \sim N(0, \sigma_S^2),\ \ \varepsilon_{ij} \sim N(0, \sigma_E^2)$$

$$CP = \hat{\mu} + 1.645 * \sqrt{\widehat{\sigma_R^2} + \widehat{\sigma_S^2} + \widehat{\sigma_E^2}}$$

In-Study Experimentation

Experimental Design

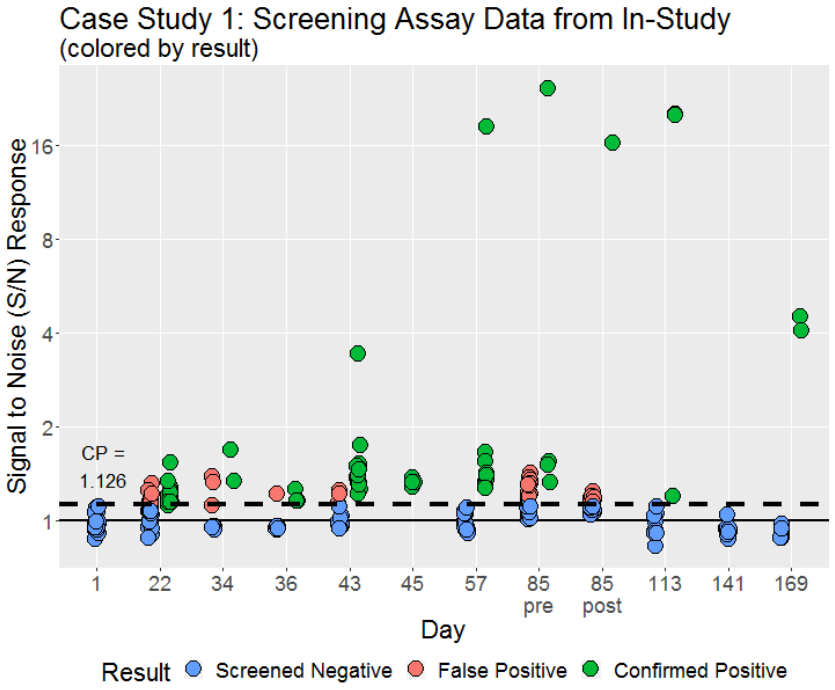
- In-study samples are observed pre-dose and post-dose throughout the study
- In-study design varies
  - e.g., number of samples, doses, and timepoints



In-Study ADA Assessment

Using the Validation Cut Point

- Validation cut point is applied to evaluate in-study samples
- Samples that screen as (“potential”) positives are conclusively assessed in confirmation assay



Suitability of the Validation Cut Point for In-Study Use

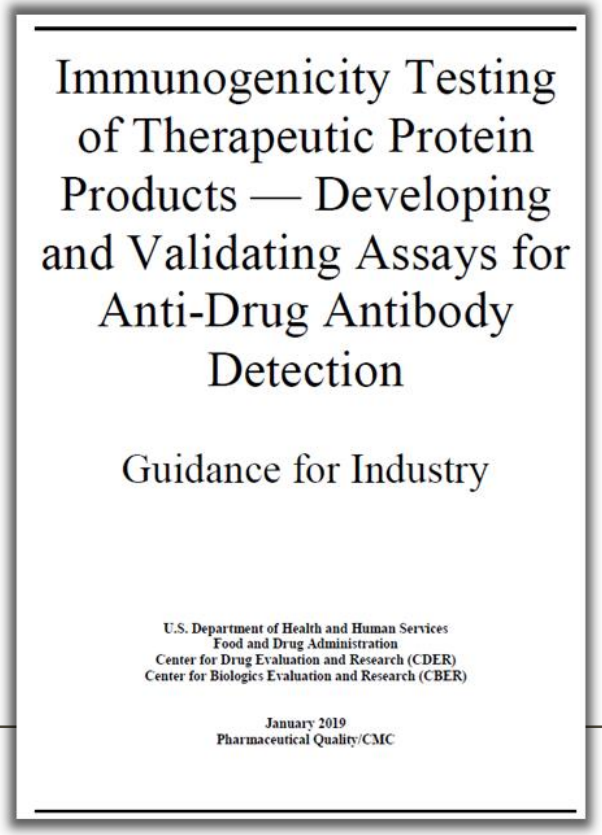
Evaluating the need for a study-specific cut point

- Has the assay changed between validation and in-study testing?
- Are there differences in the animals sampled in each stage?
- Is the validation cut point suitable for in-study use?
- What is the justification for using the validation cut point in-study?

Justification for the Validation Cut Point

– FDA (2019):

"Samples from different populations can have different background activity in ADA assays. Similarly, the background activity can change when samples used to determine the cut-point during assay validation were not obtained and handled in a manner that represents how samples will be obtained and handled in-study. ***Therefore, it is necessary to confirm that the cut-point determined during assay validation is suitable for the population being studied.*** *A sufficient number of samples from the target population should be used, and* ***justification for the number used should be provided****."*


Immunogenicity Testing of Therapeutic Protein Products — Developing and Validating Assays for Anti-Drug Antibody Detection

Guidance for Industry

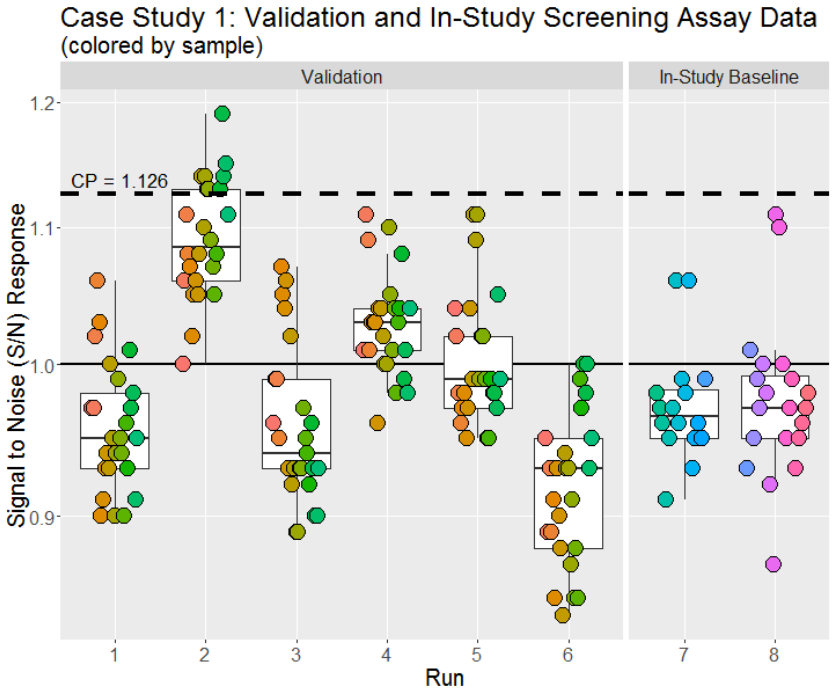
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

January 2019
Pharmaceutical Quality/CMC

Existing Approaches to Evaluate Suitability

Baseline Criteria (using data from in-study baseline)

- Criteria from literature to determine need for a study-specific cut point:
  - If in-study baseline FPR < 2% or > 11% when using validation cut point (*Amaravadi, et. al. 2015 & Devanarayan, et. al. 2017*)
  - If in-study baseline FPR > 15% (*Song, et. al. 2015*)
  - If in-study baseline variances or means differ from those in validation *(Amaravadi, et. al. 2015)*



In-Study Baseline False Positive Rate

Baseline Criteria 1

- **Available in-study baseline data**
  - One observation per sample
  - Samples may be tested in same run or across multiple runs

- **False positive rate evaluation**
  - **Pro**: Allows for an early decision on need for study-specific cut point
  - **Con**: FPR is coarser / more discrete with small sample size

Simple Tests of Baseline Variances and Means

Baseline Criteria 2

- **Method details**
  - Uses all validation data and in-study baseline data (after outlier exclusion)
  - Compares stages by variances (Levene's test) and means (two-sample t-test)

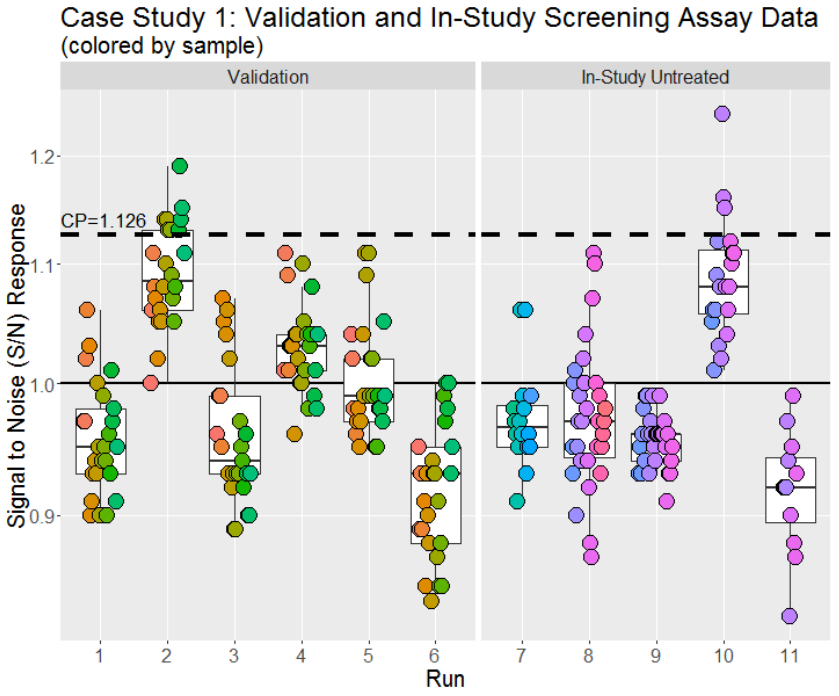
- **Variances and means evaluation**
  - **Pro**: Simple tests that are appropriate with larger experiments
  - **Con**: Does not account for known sources of variability
  - **Con**: Does not allow for an appropriate comparison to validation data
    - e.g., with large run variability and only 1 or 2 in-study baseline runs

Proposed Approaches to Evaluate Suitability

Full-Study Criteria (using data from in-study baseline and post-dose)

- In-study untreated false positive rate
- Model-based tests of untreated variances and means
- Full in-study positive predictive value

Focus on small in-study experiments (# samples < 50)



In-Study Untreated False Positive Rate

Full-Study Criteria 1

- **Available in-study untreated data**
  - One observation per dosed samples (at baseline);
    all observations over time per placebo samples (at baseline and post-dose)
  - Samples likely tested across multiple runs
- **False positive rate evaluation**
  - **Pro**: Increases total number of observations for FPR calculation
  - **Pro**: Contains more representative variability (esp. run-to-run)
  - **Con**: Can only be evaluated at end-of-study
  - **Con**: Forces unblinding of dosed groups for statistical analysis

Model-based Tests of Untreated Variances and Means

Full-Study Criteria 2

- **Method details**
  - Uses all validation data and untreated in-study data (after outlier exclusion)
  - Fit a single mixed-effects model to validation and in-study untreated data
    - Allow mean and variance parameters to vary by stage
    - Compare stages by total variances (F-test) and means (mixed-model contrast)

- **Variances and means evaluation**
  - **Pro**: Accounts for known sources of variability
  - **Pro**: Contains more representative variability (esp. run-to-run)
  - **Con**: Can only be evaluated at end-of-study

Full In-Study Positive Predictive Value

Full-Study Criteria 3

- **Available in-study data**
  - All observations over time for all samples (at baseline and post-dose)
  - Samples tested across multiple runs
- **Positive predictive value (PPV) evaluation**
  - PPV = # of Confirmed Positives / # of Screened Positives
  - **Pro**: Leverages all available in-study data
  - **Con**: Can only be evaluated at end-of-study
  - Note: Best used with balanced in-study design
  - Note: Unable to assess negative predictive value

Case Study 1

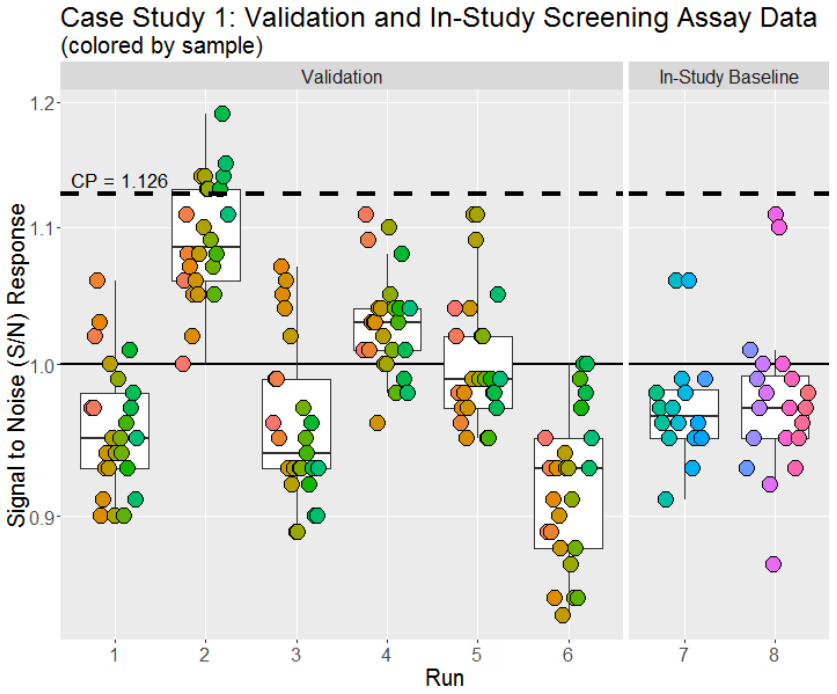
Validation and In-Study Baseline False Positive Rate

Validation:

- 25 pre-clinical samples tested in each of 6 runs
- Screening cut point = 1.126
- Validation FPR = 5.4% (8/149)

In-study baseline:

- 36 pre-clinical samples each tested once over two runs
- Baseline FPR = 0.0% (0/36)
- Baseline variance test: $p < 0.001$
- Baseline mean test: $p = 0.034$

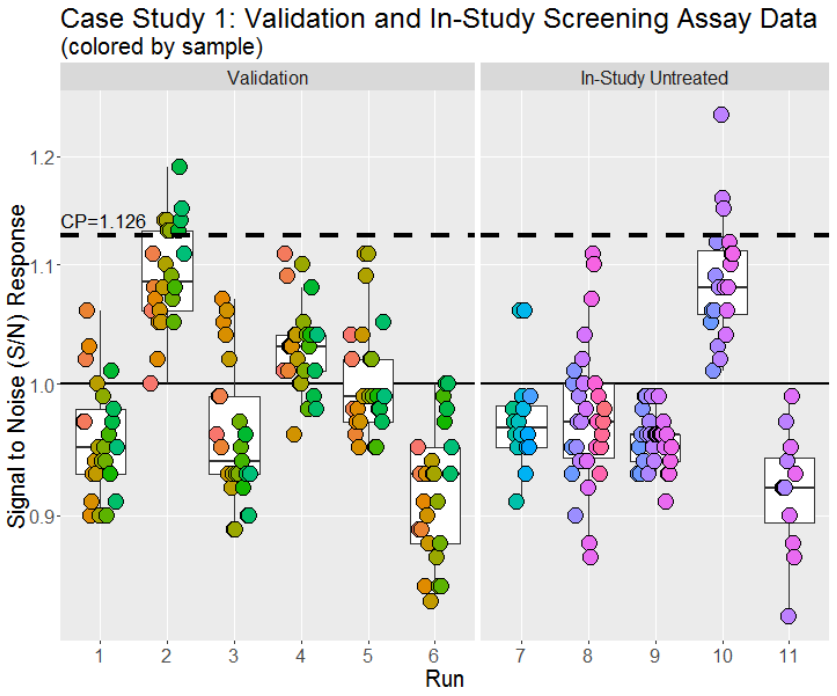


Case Study 1

In-Study Untreated False Positive Rate

In-study untreated:

- 36 pre-clinical samples
  - 10 in placebo group observed at 7-10 different time points
  - 26 in dosed groups observed once at baseline
- Samples tested over 5 runs total
- Untreated FPR = 2.8% (3/108)
  - All 3 confirmed negative
- Untreated variance test: p = 0.356
- Untreated mean test: p = 0.723

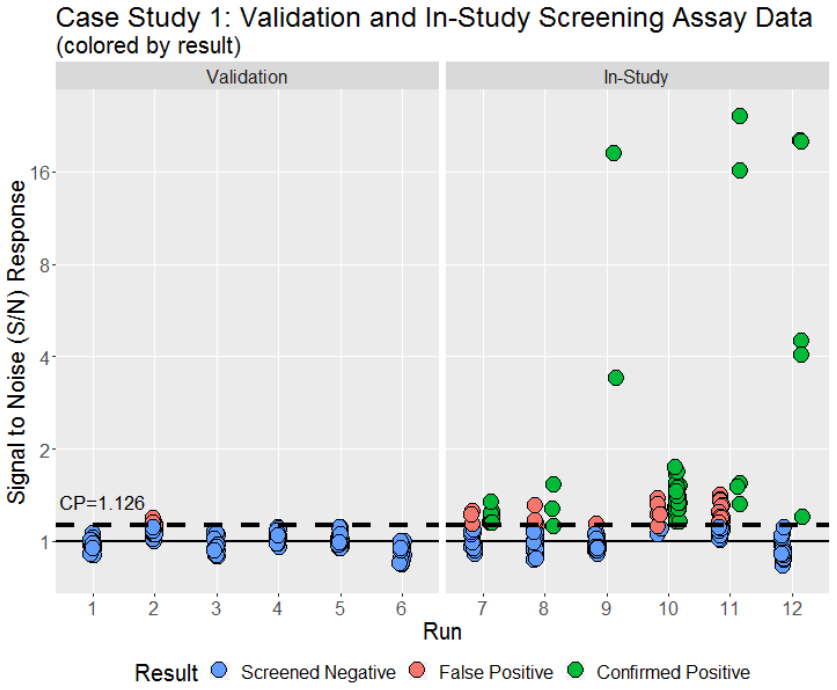


Case Study 1

Full In-Study Positive Predictive Value

In-study (full study):

- 36 pre-clinical samples observed at 2-10 different time points
- Samples tested over 6 runs
  - Each run with at least 28 samples
- Full In-Study PPV = 60.9% (56/92)



Case Study 1: Support for Validation Cut Point?

Baseline Criteria

- Baseline FPR = 0.0% (0/36)
- Baseline variance test: $p < 0.001$
- Baseline mean test: $p = 0.034$

Conclusion → **Need in-study CP**

Full-Study Criteria

- Untreated FPR = 2.8% (3/108)
- Untreated variance test: $p = 0.356$
- Untreated mean test: $p = 0.723$
  - 95% CI on mean ratio (0.90,1.07)
- Full In-Study PPV = 60.9% (56/92)

Conclusion → **Suitable validation CP**

Case Study 2

In-Study Baseline and Untreated False Positive Rate

Validation:

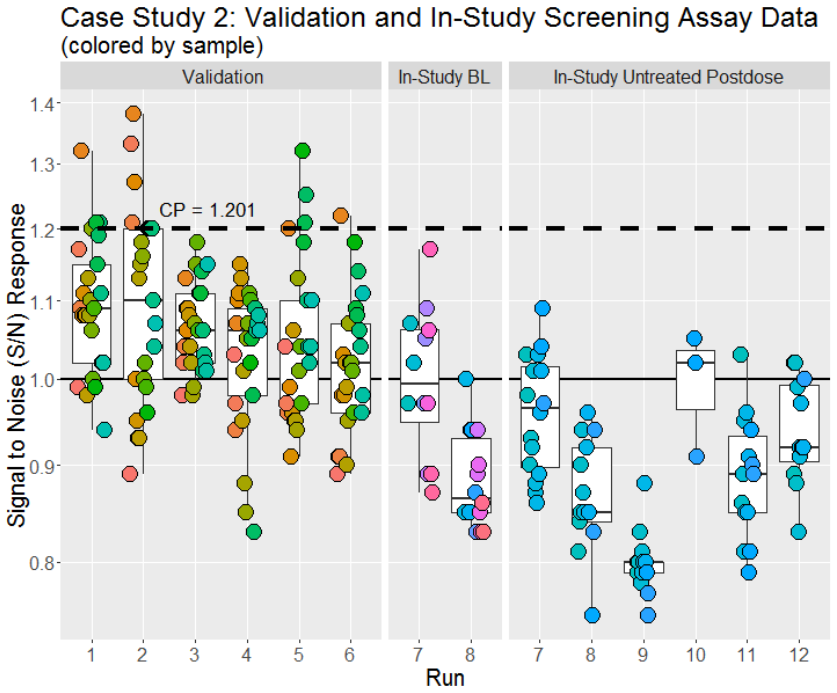
- 25 pre-clinical samples tested in each of 6 runs
- Screening cut point = 1.201

In-study baseline:

- 26 pre-clinical samples each tested once
- Samples tested over 2 runs
- Baseline FPR = 0.0% (0/26)

In-study untreated:

- 8-9 more reps for 8 placebo samples
- Samples tested over 6 runs total
- Untreated FPR = 0.0% (0/97)



Case Study 2: Support for Validation Cut Point?

Baseline Criteria

- Baseline FPR = 0.0% (0/26)
- Baseline variance test: p = 0.221
- Baseline mean test: p < 0.001

Conclusion → **Need in-study CP**

Full-Study Criteria

- Untreated FPR = 0.0% (0/97)
- Untreated variance test: p = 0.897
- Untreated mean test: p = 0.005
  - 95% CI on mean ratio (0.79, 0.94)
- Full In-Study PPV = 75.8% (25/33)

Conclusion → **Need in-study CP**

Case Study 3

In-Study Baseline and Untreated False Positive Rate

Validation:

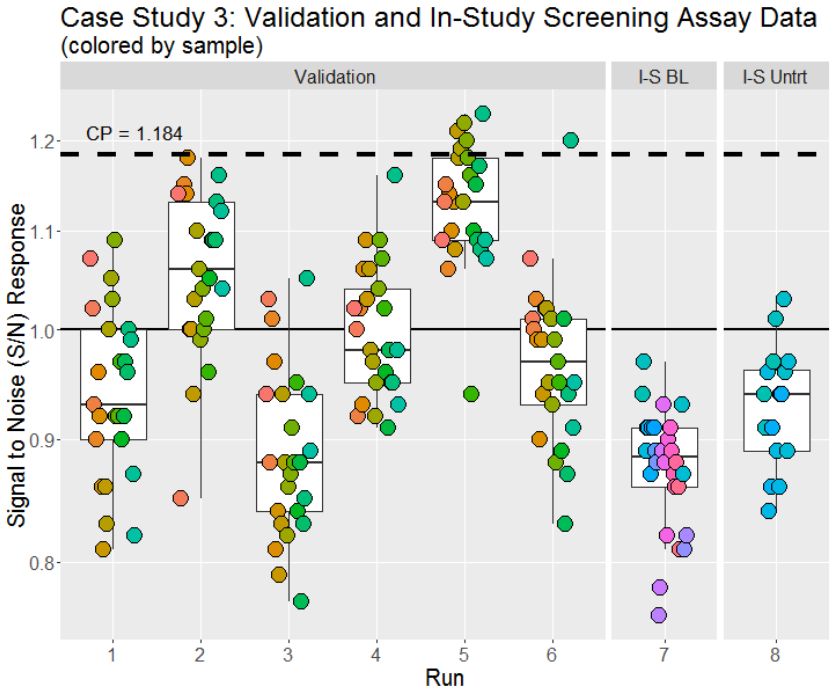
- 25 pre-clinical samples tested in each of 6 runs
- Screening cut point = 1.184

In-study baseline:

- 28 pre-clinical samples each tested once
- Samples tested in 1 run
- Baseline FPR = 0.0% (0/28)

In-study untreated:

- 2 more reps for 8 placebo samples
- Samples tested over 2 runs total
- Untreated FPR = 0.0% (0/44)



Case Study 3: Support for Validation Cut Point?

Baseline Criteria

- Baseline FPR = 0.0% (0/28)
- Baseline variance test: $p < 0.001$
- Baseline mean test: $p < 0.001$

Conclusion → **Need in-study CP**

Full-Study Criteria

- Untreated FPR = 0.0% (0/44)
- Untreated variance test: $p = 0.025$
- Untreated mean test: $p = 0.061$
  - 95% CI on mean ratio (0.81, 1.01)
- Full In-Study PPV = *NA* (0/0)

Conclusion → **Want more data**

Takeaways

When evaluating the suitability of a validation cut point for in-study use…

- Consider multiple bases of evidence
- Use all available in-study data
- Account for the study design / sources of variability (especially Run)
- Increase in-study sample size when possible (# of samples or # of runs)

Acknowledgements

Thank you!

GSK Research Statistics

- Mark Burgert
- Mandy Bergquist
- Katja Remlinger

GSK Immunogenicity

- Scott Hottenstein
- Kimberly Florence

References

- Amaravadi, Lakshmi, 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 (2015): 3107-3124.
- Devanarayan, Viswanath, et al. "Recommendations for systematic statistical computation of immunogenicity cut points." The AAPS journal 19.5 (2017): 1487-1498.
- FDA. "Guidance for Industry: Immunogenicity Testing of Therapeutic Protein Products —Developing and Validating Assays for Anti-Drug Antibody Detection." (2019)
- Jani, Darshana, et al. "Recommendations for the development and validation of confirmatory anti-drug antibody assays." Bioanalysis 7.13 (2015): 1619-1631.
- Song, Sam, et al. "Understanding the supersensitive anti-drug antibody assay: unexpected high anti-drug antibody incidence and its clinical relevance." Journal of immunology research 2016 (2016).
- Zhang, Jianchun, and Robert J. Kubiak. "Clinical Assays for Biological Macromolecules." Nonclinical Statistics for Pharmaceutical and Biotechnology Industries. Springer, Cham, 2016. 263-294.