Biologics Development

The Power of Proper System Suitability Tests
- A Case Study of cGMP Method Improvement

APR2021

Ruojia Li
CASSS Bioassays 2021

Bristol Myers Squibb™

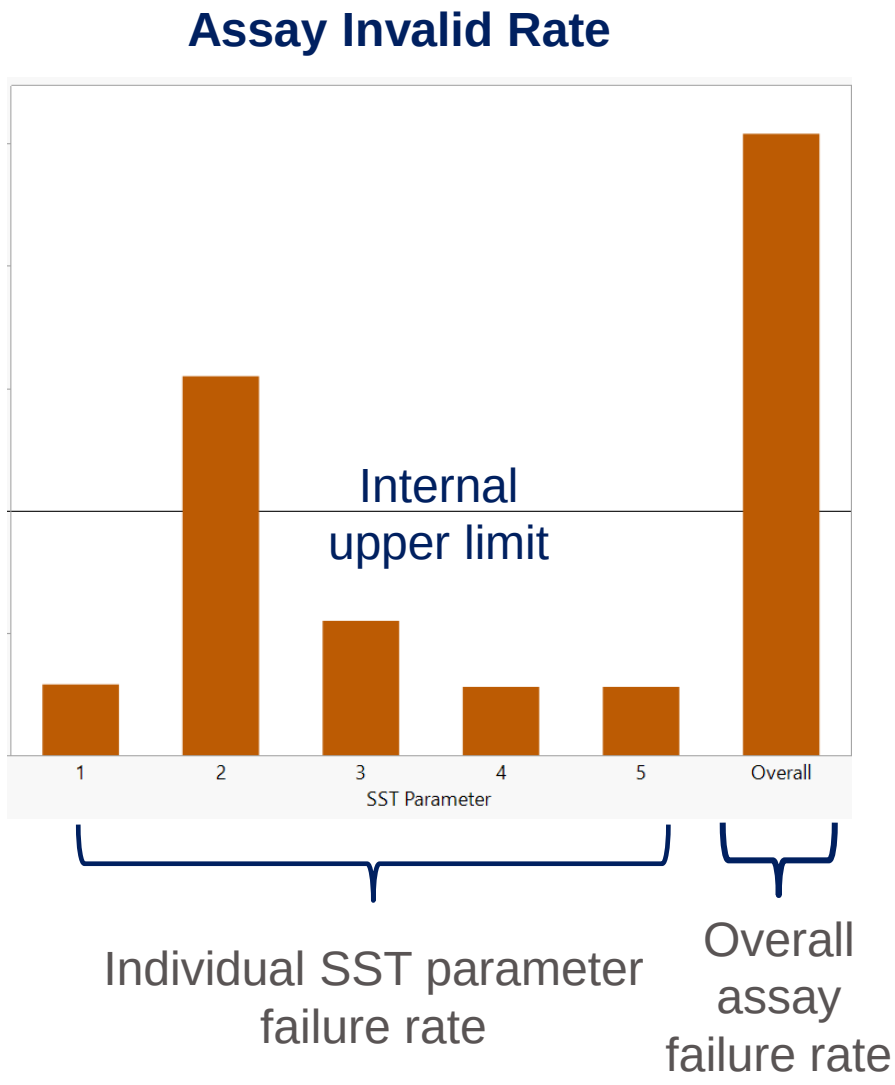
Outline

- Problem statement
- Method assessment
- Method improvement: system suitability updates
- Summary

Problem Statement

- A legacy GMP bioassay method suffers from higher than desirable assay and sample failure rate
  - High assay invalid rate, i.e., failed system suitability tests (SST) that applied to reference and/or control sample
  - Additional sample repeats due to similar sample acceptance failures

The method must be assessed and improved



Potential Root Causes for High Assay Failure Rate

- Undesirable assay data quality
  - E.g., due to non-optimal assay design, assay conditions, etc.
- Inappropriate statistical model and/or data analysis
- Operational errors
  - E.g., due to dilution, instrument, analyst training
- Improper system suitability criteria
  - System suitability parameter
    - Not reflective of assay data quality
    - Can not effectively differentiate good vs. bad assays
    - Not robust. E.g., only applicable to a subset of labs / instruments
  - Acceptance range
    - Not based on representative data set and appropriate evaluation

Method Assessment

- Reviewed relevant documents and data to identify potential root cause for high assay failure rate

Data / documents reviewed	Observation	
Method validation report and long-term method performance trending data	Great accuracy and precision	
Method procedure, assay development / optimization DOE data and analysis	No major concern with assay design and conditions	
Large amount of existing assay outputs, including: - Numerical results (curve fit parameters, SST, potency) - Graphs (dose-response data and fitted curves)	- Reasonable statistical model - Acceptable data quality in general (goodness of fit, variability)	
Preliminary review of SST results - Existing system suitability parameters and ranges - Data used to set / justify the SST criteria - Outputs of failed and passed assays	- Some SST parameters do not effectively control assay quality - Some critical SST parameters are missing	

- Potential primary root cause: improper SST criteria

Method improvement plan: thoroughly re-evaluate and re-establish system suitability criteria

Re-Evaluation of System Suitability Criteria

- Review each existing system suitability parameter and the acceptance range
  - Parameter
    - What is the intended purpose of the parameter?
    - Is the intended purpose directly related to the quality of assay results?
    - Does the parameter provide meaningful assessment for its intended purpose, i.e., effectively differentiate desirable vs. unacceptable assay data?
    - Do all the parameters together provide adequate system suitability assessment?
    - Are there any redundant parameters?
  - Acceptance range
    - How was the range determined?
    - What data set and analysis were used to set / justify the range?
    - Was the data set representative? Was the analysis appropriate?

Re-Establishment of System Suitability Criteria

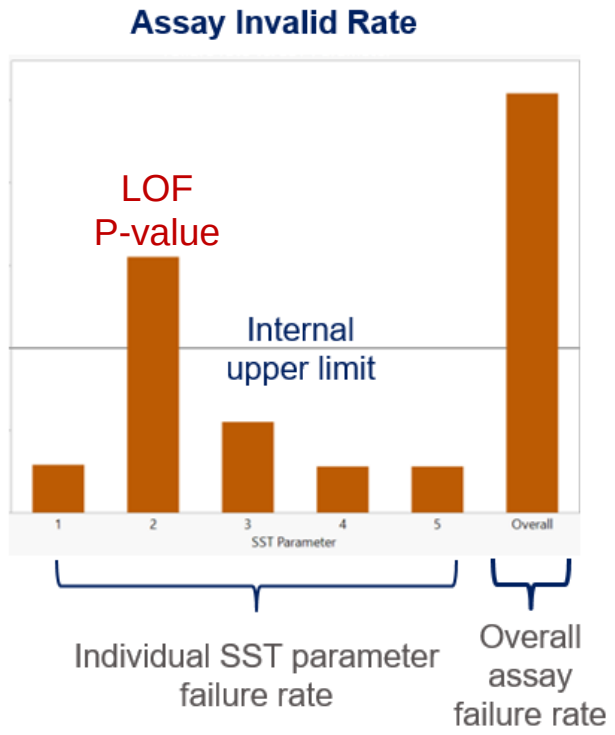
- Based on the re-evaluation, existing system suitability criteria were added, replaced, removed or retained as appropriate

Re-Evaluation Observation	Decision / Action
Critical SST assessments missing	**Add** new criteria to fill the gaps
Ineffective SST criteria	**Replace** existing criteria with properly defined new criteria *
Redundant / non value added criteria	**Remove** with appropriate justification
Properly defined SST parameters and ranges	**Retain** existing criteria

* An example of SST criteria replacement will be presented in the following slides

Example: Replacement of Lack-of-Fit SST Criterion

- Lack-of-fit (LOF) P-value based on ANOVA F test was used in the legacy method to assess goodness of fit
- LOF P-value was the most contributing criterion to assay failures
- The legacy P-value approach has known limitations
  - Tends to over-sensitively reject precise data with adequate fit and retain noisy data with poor fit
- A new LOF criterion (relative LOF error) was established to replace the legacy P-value criterion to provide more meaningful assessment



Original Parameter: LOF P-Value

- LOF P-value based on F test

$$F\ ratio = \frac{SS_{LOF}/DF_{LOF}}{SS_{PE}/DF_{PE}} = \frac{\sum_{i,j}(\bar{y}_i - \hat{y}_i)^2\ /DF_{LOF}}{\sum_{i,j}(y_{i,j} - \bar{y}_i)^2\ /DF_{PE}}$$

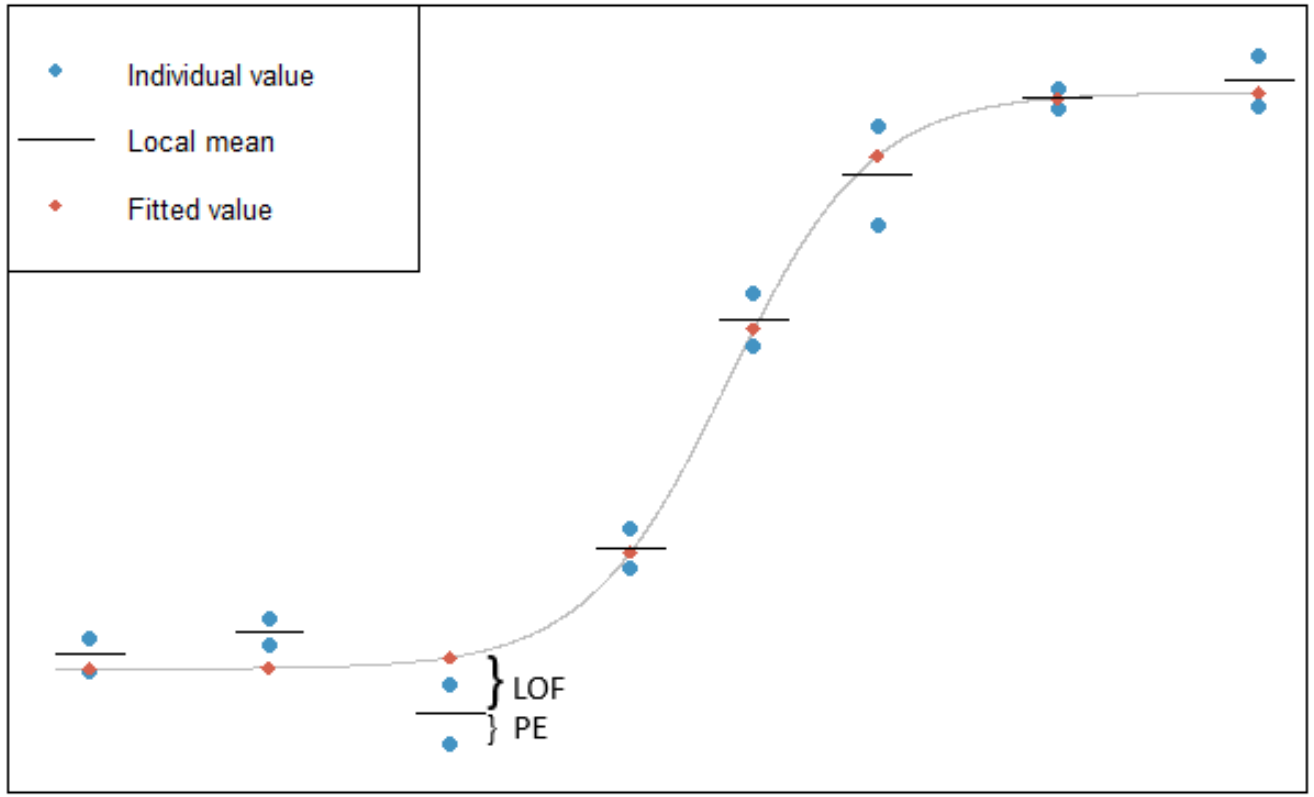
- Assay fails if the LOF term is statistically significant (small P-value)

Notations

- $y_{i,j}$: Individual response value
- $\bar{y}_i$: Local mean of individual response values at given concentration
- $\hat{y}_i$: Fitted value at given concentration
- SS_{LOF}: Sum of squares of LOF error $(\bar{y}_i - \hat{y}_i)$
- SS_{PE}: Sum of squares of pure error $(y_{i,j} - \bar{y}_i)$
- *DF*: Degrees of freedom

Original Parameter: LOF P-Value (cont.)

- Intended purpose
  - Assess the adequacy of the dose-response model
- How does LOF P-value work?
  - Compare LOF error to pure error (PE)
  - Assay fails if LOF error is too large compared to pure error
- Limitation of LOF P-value
  - Tends to penalize precise data (with small PE) and propensity to retain undesirable noisy data (with large PE)



LOF error: difference between local mean and fitted value (measures the closeness of the fitted curve to the observed data)

Pure error: difference between individual value and local mean (measured the precision of observed data)

New Parameter: Relative LOF Error

- Relative LOF error

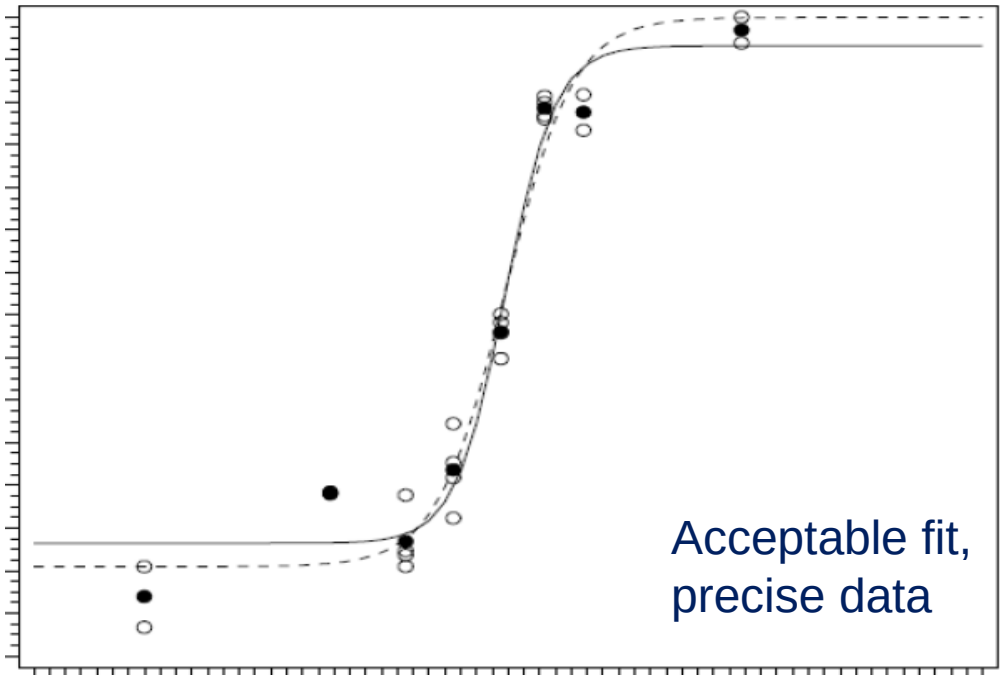
$$\frac{\sqrt{SS_{LOF}/N}}{A_{ref}-D_{ref}} \times 100\%$$

  - LOF error normalized against reference curve window (upper asymptote A – lower asymptote D)
  - Assay fails if relative LOF error is too large

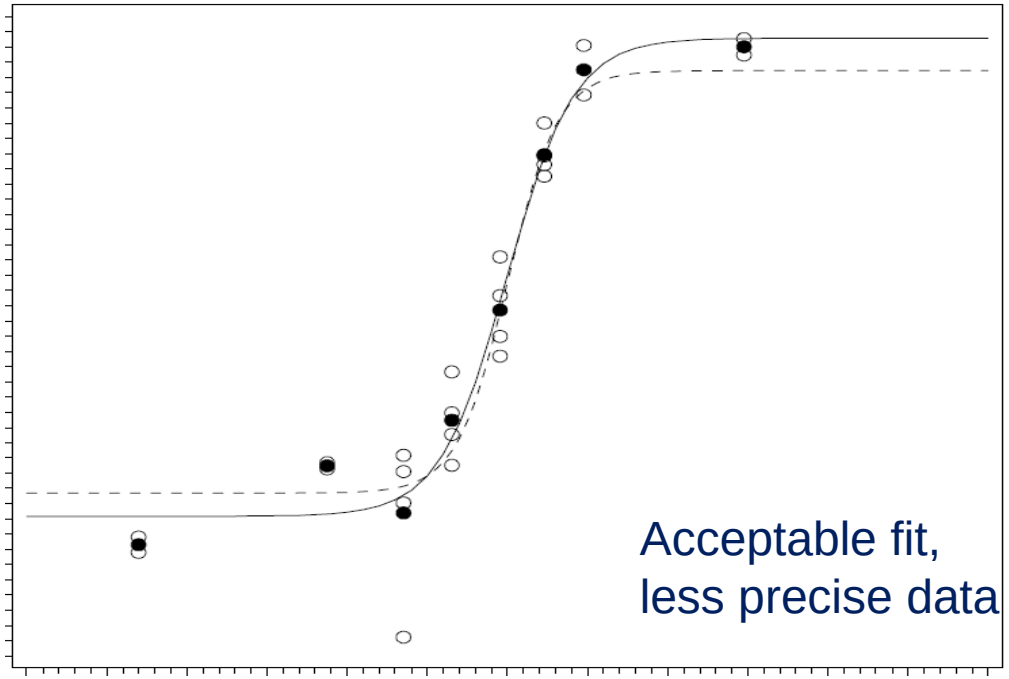
  - A more robust measurement of lack-of-fit
  - Independent of pure error and thus overcomes the shortcomings of LOF P-value
  - Independent of the magnitude of response readings

Li, R., Cai, W. and Zocher, M.S. (2017). A Novel Lack-of-Fit Assessment as a System Suitability Test for Potency Assays, PDA Journal of Pharmaceutical Science and Technology, 71 368-378.

Example: Comparison of Original and New LOF Criteria - Representative Assay Plots



LOF P-value: Failed
Relative LOF error: Passed



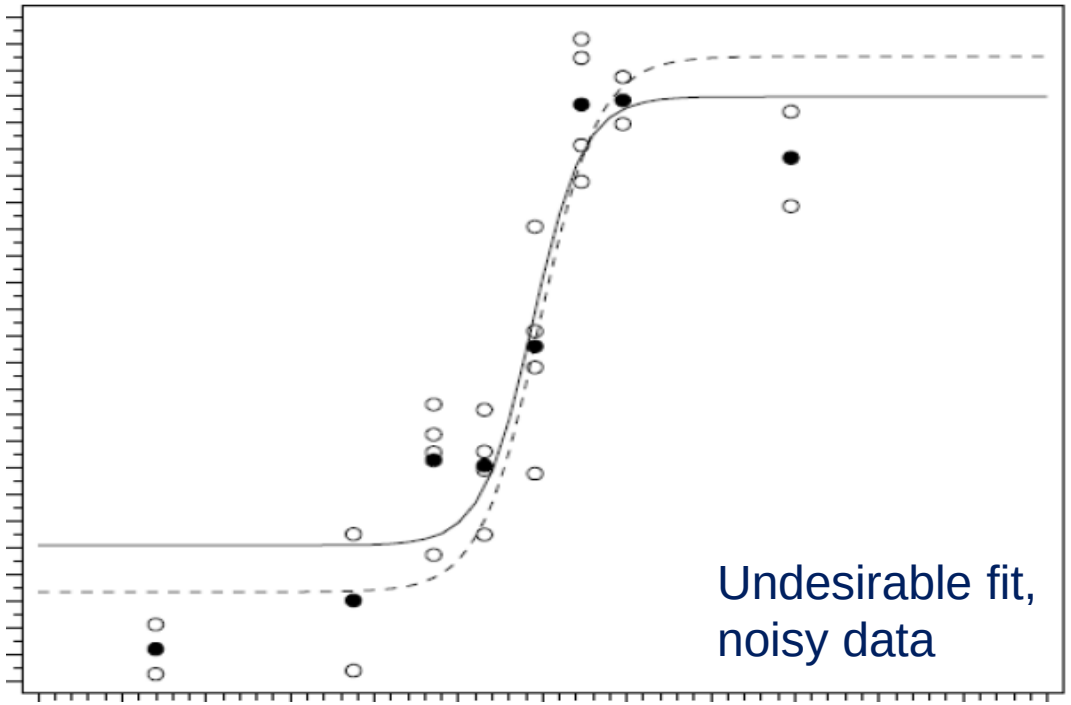
LOF P-value: Passed
Relative LOF error: Passed

Individual response
Mean response
Fitted curve
Reference curve

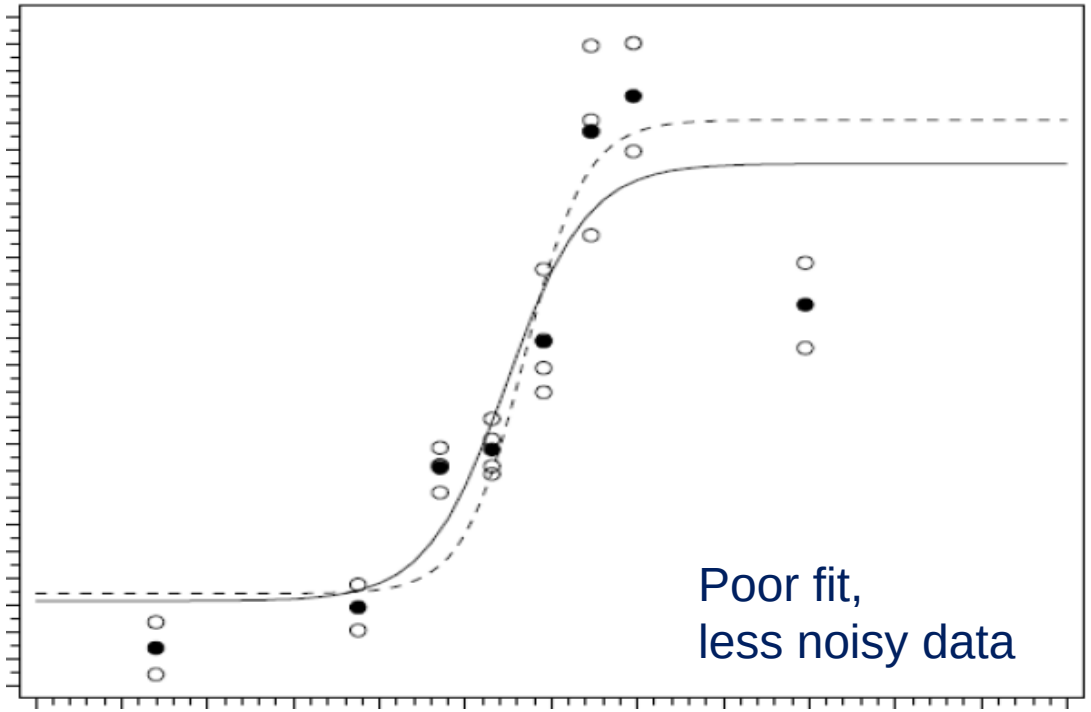
LOF P-value tends to over-sensitively reject precise assays with acceptable fit

Relative LOF error retains assays with acceptable fit regardless of noise level

Example: Comparison of Original and New LOF Criteria - Representative Assay Plots



LOF P-value: Passed
Relative LOF error: Failed



LOF P-value: Failed
Relative LOF error: Failed

LOF P-value could retain noisy data with undesirable fit

Relative LOF error rejects assays with unacceptable fit regardless of noise level

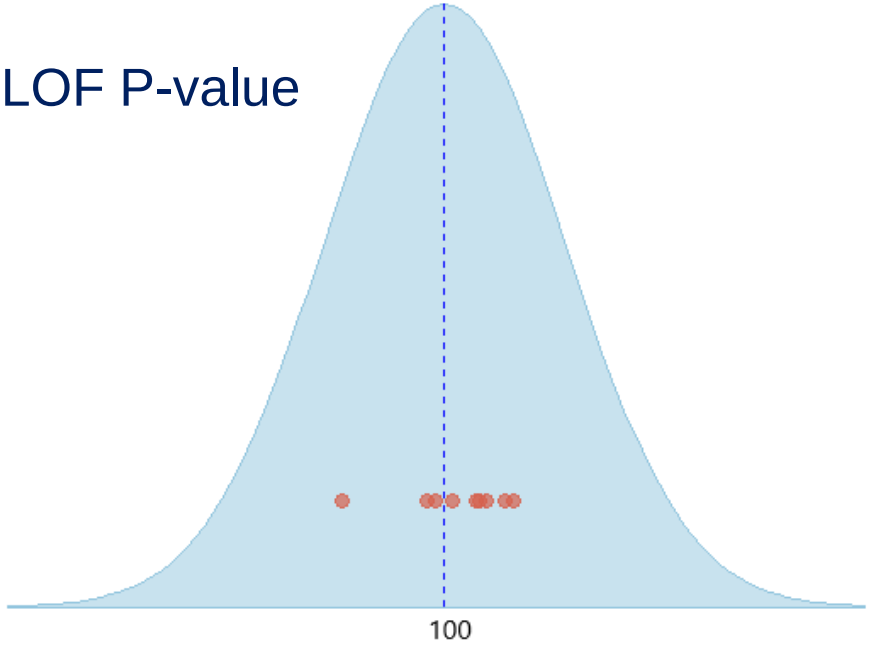
Example: Comparison of Original and New LOF Criteria - Passed vs. Failed Results

Blue: Distribution of QC potency recovery results that passed LOF test

Red: Individual QC potency recovery results that failed LOF test

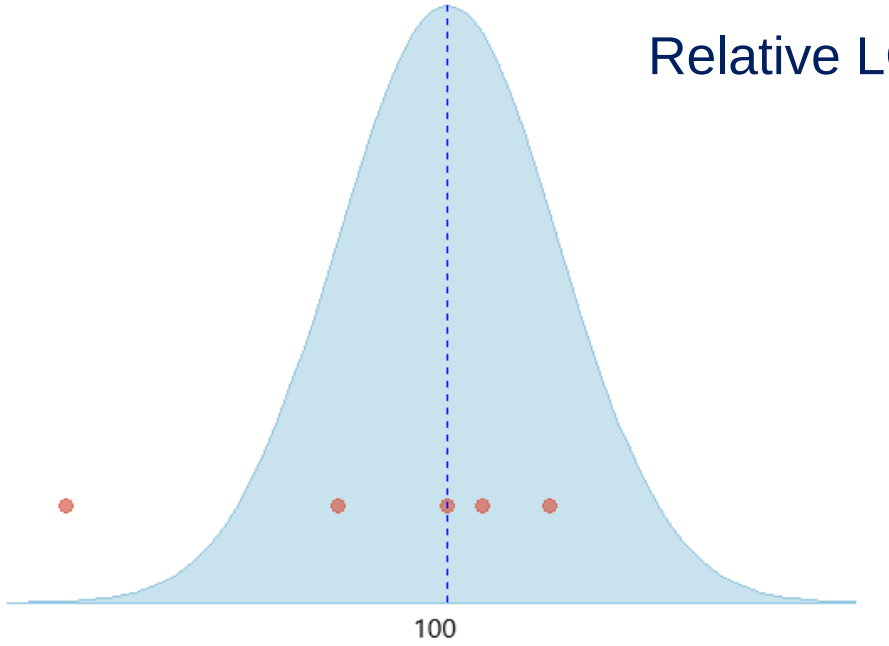
(Data source: Method validation)

LOF P-value



Overall great potency recovery results from assays failed LOF P-value

Relative LOF error

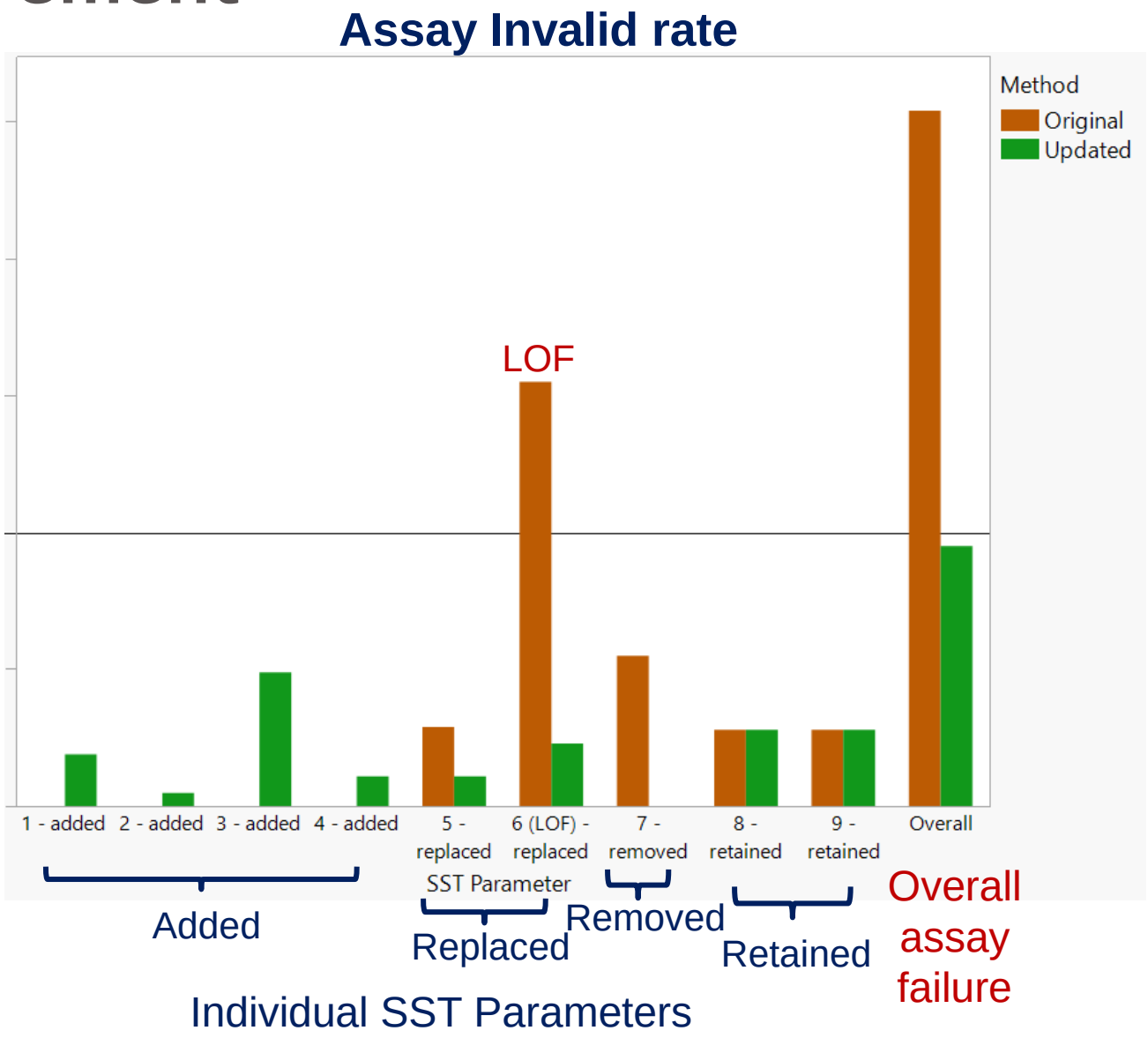


Some inaccurate potency recovery results from assays failed relative LOF error

Outcomes of Method Improvement

The method was significantly improved with updated SST criteria (added, replaced, removed or retained)

- Adequate and more meaningful SST assessment
- Overall assay invalid rate reduced by more than 60%
- Same great accuracy and precision
  - Based on retrospective analysis of historical data



Implementation of the New SST Criteria

- Documentation of SST updates and justification
- Data analysis software updates and re-validation
- Validation amendment
  — Re-assess existing validation data (with updated SST applied) against validation criteria
- Method change control and filing

Summary

- A legacy cGMP bioassay suffered from high assay failure rate
- Improper system suitability tests was identified as primary root cause
- Without any wet lab work, the quality and success rate of the legacy method were significantly improved by implementing state of the art updated system suitability criteria
- The case study clearly illustrated the power of proper system suitability tests

Acknowledgement

Bioassay Center of Excellence

Weiguo Cai

Jeff Glenn

Isam Qahwash

Lihong Mcaleer

Shelley Fang

Hai Zhang

LeeAnn Benson-Ingraham

Marcel Zocher

Thank you

Bristol Myers Squibb™