BSR/PDA Standard 07-202x, Analytical Procedure Replacement, Transfer, and the Use of Platform Analytical Procedures for Biologics

Draft stage

Revision: V22
Draft Effective Date: 05/01/2024

BSR/PDA Standard 07-202x, Analytical Procedure Replacement, Transfer, and the Use of Platform Analytical Procedures for Biologics

Authors

Stephan Krause (Chair)
Bristol-Myers Squibb

Jayda Siggers (Co Chair)
Health Canada

Maria Susana Albano
Bristol-Myers Squibb

Deborah Aguiar
Sanofi

Cynthia Allen
Health Canada

Rich Anderson
Biogen

Leslie Bumbry
Horizon Therapeutics

Cyrille Chéry
UCB Pharma SA

Kedar Dave
Bristol-Myers Squibb

Jean-François Dierick
GlaxoSmithKline

Cindy Duhigg
Alcon

Christof Finkler
F. Hoffmann-La Roche Ltd

Alexandra Heussner
Vetter Pharma-Fertigung GmbH & Co. KG

Alexey Khrenov
U.S. Food and Drug Administration (FDA)

Claudia Magagnoli
GlaxoSmithKline

Carl Miller
Seagen

77 Most Nahid Parvin
78 CBER, FDA

79 Christopher Shaw
80 AstraZeneca

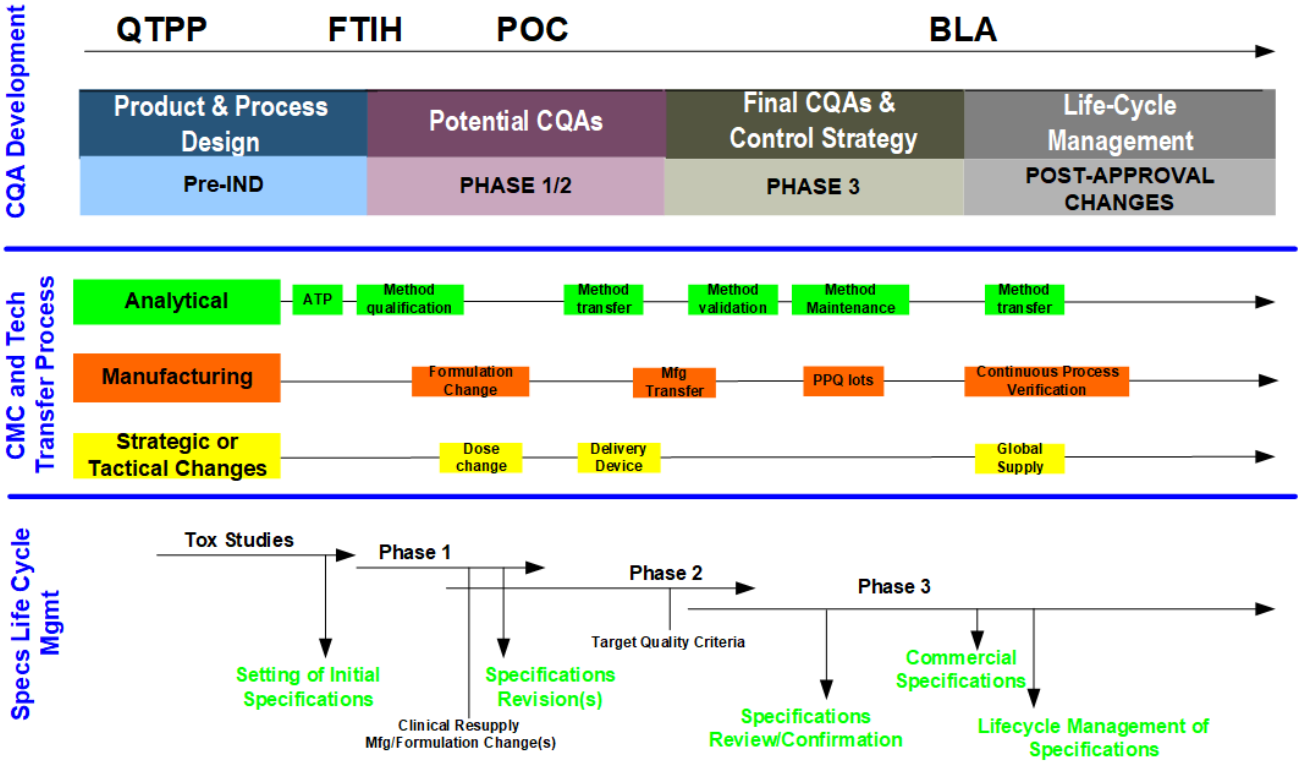
81 Jason Starkey
82 Pfizer
83

Contents

1. Introduction

This standard presents a consensus for current best practices for Analytical Method Transfer (AMT), the use of Platform Analytical Procedure (PAP) methods, and Analytical Method Comparison (AMC). The analytical method lifecycle steps, AMT and AMC, are illustrated in **Figure 1** (below) within the overall CMC development roadmap to and beyond marketing authorization(s). The interdependent development lifecycle steps are aligned with typical clinical study phases and the understanding of critical quality attributes (CQAs). Additional AMT steps and method changes, requiring AMC studies, may occur at various stages before and after the marketing authorization filing. Risk-based and phase-appropriate (where appropriate) AMC, AMT, and/or PAP study considerations are provided in this standard.

Figure 1: **Analytical Method Lifecycle Steps [1]**



In **Figure 1**, a typical CMC development and technology transfer process is shown on a clinical study time scale and separated into analytical procedure ("method" is synonymous with "procedure"), manufacturing, and strategic and/or tactical changes. Manufacturing and method transfers can occur multiple times prior to market authorization. For global drug product (DP) lot release testing, analytical methods may need to be transferred to other testing laboratories and/or into regions requiring import lot release testing.

Manufacturing changes, such as formulation changes to increase product stability and/or product container compatibility, are often implemented following successful product comparability studies. The second process validation (PV) stage (process qualification (PQ) lots) is typically executed with a final control strategy and protocol acceptance criteria based upon commercial-ready specifications. Analytical methods that support release and stability specifications are usually validated prior to the manufacturing of PQ lots. Specifications may need to be revised following the manufacturing of PQ lots, as part of the outcome of the negotiations with regulatory agencies following the market authorization application (e.g., the BLA, NDS, MAA) submission [1].

There are common method performance characteristics required in study protocols for AMT, PAP, and AMC, with some conceptual differences among those studies dependent on the method type [2], in accordance with ICH Q2(R2)[3]. The principles outlined in the standard can be applied to methods developed using a minimal or enhanced approached, as described in ICH Q14 [4]. **Table 1** (below) provides a summary of suggested statistical methodologies and their interpretation to demonstrate successful AMT, AMC, or PAP studies for each performance characteristic. Alternative approaches may be used if justified and explained in detail. More detail is provided in each of the three sections for AMT, PAP, and AMC studies.

Table 1: Performance Characteristics and Suggested Statistical Methodologies for AMT, AMC and PAP Study Protocols for ICH Q2(R2) Method Types

Method Type (ICH Cat. no.)	AMT		AMC		PAP Verification	
	Performance Characteristics	**Statistical Methodology**	**Performance Characteristics**	**Statistical Methodology**	**Performance Characteristics**	**Statistical Methodology**
Identification (cat. I)	**Specificity**	Probability, Chi-Squared (or similar) for Pass/Fail Ratio	**Specificity**	Non-Inferiority or Superiority Pass/Fail Ratio(s) or Probabilities	**Specificity**	Probability, Chi-Squared (or similar) for Pass/Fail Ratio
Impurity - Quantitative (cat. II)	**Accuracy[1]**	**TOST**	**Accuracy[1]**	**TOST**	**Accuracy[1]**	**% Recovery**
	Intermediate Precision[1]	**RSD, other options [2]**	**Intermediate Precision[1]**	**RSD, other options [2]**	**Intermediate Precision[1]**	N/A
	Specificity	N/A	**Specificity**	**% Recovery, other options**	**Specificity**	**% Recovery, other options**
Impurity – Limit (cat. III)	**Specificity[1]**	Probability and/or Chi-Squared for P/F	**Specificity[1]**	Non-Inferiority or Superiority; Pass/Fail Ratio(s) or Probabilities	**Specificity[1]**	Non-Inferiority or Superiority; Pass/Fail Ratio(s) or Probabilities
	Detection Limit	N/A	**Detection Limit**	Non-Inferiority or Superiority	**Detection Limit**	Non-Inferiority or Superiority
Assay (Content, Potency) (cat. IV)	**Accuracy[1]**	**TOST**	**Accuracy[1]**	**TOST**	N/A	N/A
	Intermediate Precision[1]	**RSD, other options [2]**	**Intermediate Precision[1]**	**RSD, other options [2]**	N/A	N/A
	Specificity	N/A	**Specificity**	**% Recovery, other options**	N/A	N/A

[1] At/above QL for Category II and at/above DL for Category III.

2 RSDs (%CV) for Intermediate Precision can be compared from the AMV study results and/or assay control long-term variation. When using a TOST equivalence test, the pooled variation of both methods will impact the confidence interval width for the mean difference. When setting TOST equivalence margins and interpreting TOST outcomes, this should be considered. For example, an equivalence margin overlapping mean difference CI may be observed when the new method variation is significantly higher than the current method.

Controlling the risk(s) associated with AMT, PAP, and AMC studies to product quality continuity and the potential impact to patient and manufacturer follows similarly established concepts [1].The primary goal for risk assessments for analytical lifecycle studies, including AMT, PAP, and AMC study protocols, are to inform:

- The desired level and/or rigor of formal studies to be executed; and,
- The level of method performance needed as manifested in the protocol acceptance criteria.

As variation and uncertainty in test results constitute risk to patient and the manufacturer, they should be evaluated and used to set acceptance criteria to ensure the suitability for use of the analytical method. A relationship of typical variation sources is expressed in **Equation 1** (below). For simplicity, the potential variation sources from the sampling process, transport, and storage, and/or the inconsistency in batch uniformity are part of the manufacturing process variation [1].

Equation 1: Relationship of Common Variation Sources

$$\left[\sigma_{\text{mfg process observed}}\right]^2 = \left[\sigma_{\text{analytical method}}\right]^2 + \left[\sigma_{\text{mfg process actual}}\right]^2$$

The (squared) observed manufacturing process variation is the sum of the (squared) variation sources of analytical method performance and the actual or true manufacturing process variation. As specifications exist for the observed manufacturing process variation, it is critical to understand and control the underlying sources of variation using risk-based acceptance criteria for each of their maximum allowable variation.

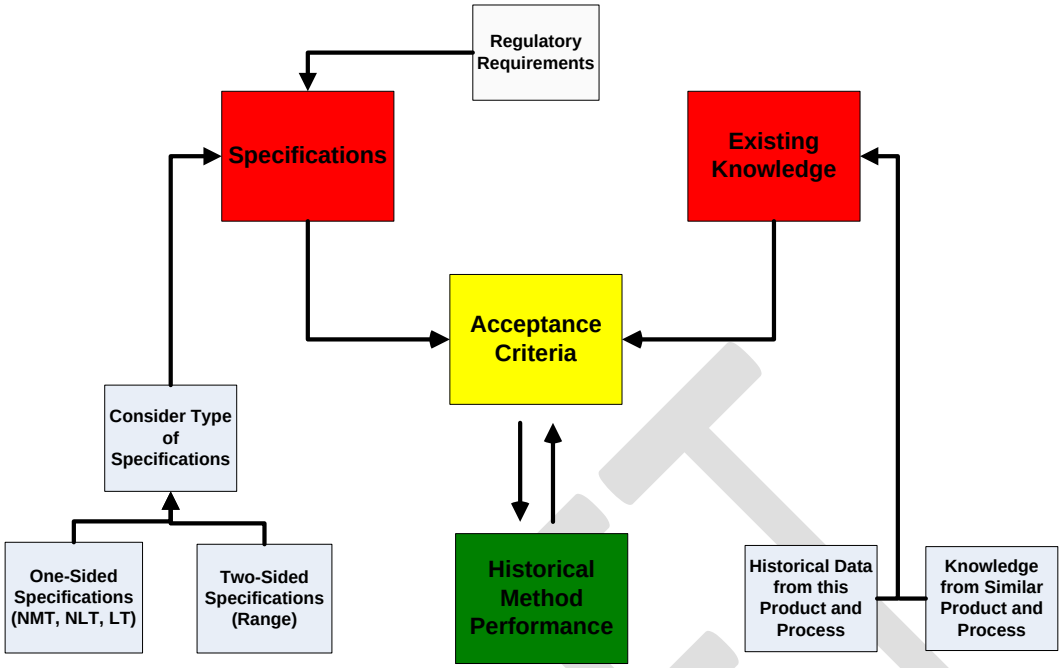
Regardless of the type of study protocol (AMT, PAP, or AMC), the study acceptance criteria are intended to control the maximum allowable risk-based level for bias or variation. Risk-based AMT, PAP, and AMC protocol acceptance criteria should, therefore, be predominately derived from the evaluation of two critical sources - (target) specification and process capability - to ensure an acceptable post-change product quality continuum. Process capability data may include historical data for the product and similar products and process(es), usually expressed in CpK values for commercial manufacturing. Other sources such as regulatory expectations may also impact acceptance criteria and should be considered when applicable. If the consistency of the sampling process, batch uniformity, and product/sample stability is not an integral part of the manufacturing process variation or not known, these variation sources will also need to be considered.

As given in ICH Q14, the product and process understanding leads to the identification of critical quality attributes (CQAs) requiring analytical measurement for control which may be included in the quality target product profile (QTPP). Relevant Analytical test method requirements can be captured in an analytical target profile (ATP) which forms the basis for development of the analytical procedure. ATP is a concept that outlines the minimum acceptable performance characteristics to ensure the method is suitable for the intended purpose. ATP is a development activity and is used to define the analytical procedure attributes and performance criteria for analytical procedure validation (ICH Q2(R2)). The risk-based approach to setting and justifying the study protocol acceptance criteria are based on the current (target) specification and historical method performance, which may include the proven acceptable range(s) (PARs) and the method operable design region (MODR) from a method developed using an enhanced approach. Given that the PARs/MODR for a method developed using an enhanced approach would inform the historic method performance when setting the study protocol acceptance criteria, the principles outlined in Q14 are indirectly captured in the standard.

While formal documentation and submission of an ATP is optional, it can be used to support study protocol acceptance criteria when the ATP lifecycle is aligned with product/process lifecycle change(s). As this standard is intended to be apply to any product development stage, including the commercial stage, and approved products, the risk-based setting and justification for study protocol acceptance criteria are based on current (target) specifications and historical analytical and manufacturing performance as described here.

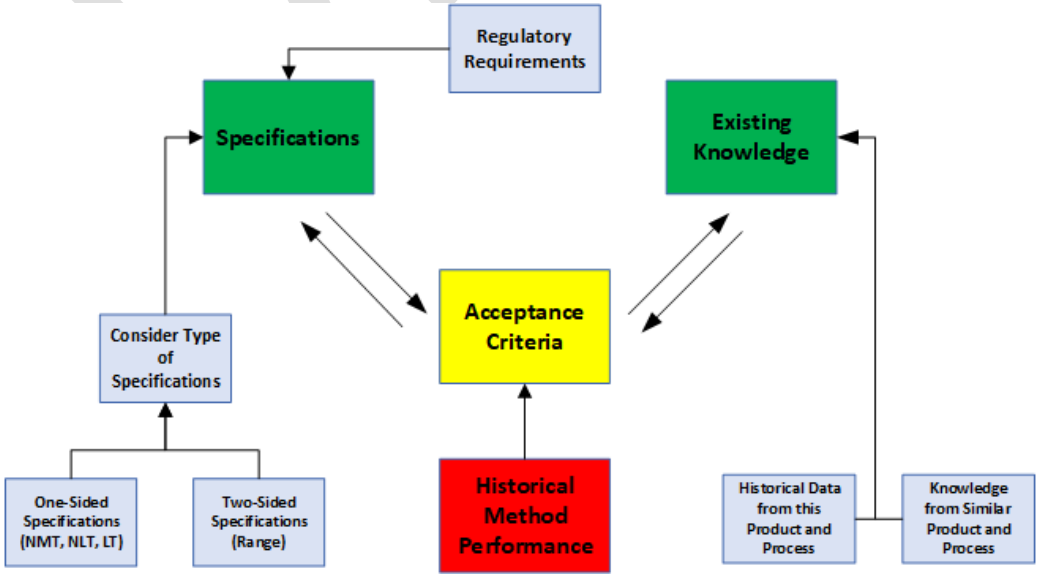
Acceptance criteria for an AMT, PAP, or AMC protocol can be defined using two main concepts – outside-in or inside-out. **Figure 2** below illustrates how protocol acceptance criteria may be conceptually derived [1] for late stage/commercial manufacturing. In the outside-in approach, acceptance criteria are set to assure an acceptable level of maximum (worst-case) performance of individual and total performance attributes. These method performance expectation limits are then compared to the historical method performance capability data. In those cases where limited data points exist to estimate true process and/or method capabilities, other sources such as existing PAP method performance data can be used to assist in setting acceptance criteria. The method performance criteria for AMT, PAP, and AMC should assure that the method will produce sufficiently accurate and reliable results.

Figure 2: Risk-Based "Outside-In" AMT, PAP, and AMC Study Protocol Acceptance Criteria [2]



Whereas **Figure 3** (below) illustrates how protocol acceptance criteria may be conceptually derived for early-stage manufacturing for a new product type (platform manufacturing, and/or platform analytical procedure cannot be used). In the inside-out approach, acceptance criteria for AMT and/or AMC are derived based on limited historical method performance data. For example, the AMT acceptance criteria for the TOST equivalence margins may be +/- 1.0 SDs of the method mean value of the Sending Unit (SU). The maximum allowed bias/difference would then be confirmed versus existing knowledge and specifications.

Figure 3: "Inside-Out" AMT, PAP, and AMC Study Protocol Acceptance Criteria



BSR/PDA Standard 07-202x, Analytical Procedure Replacement, Transfer, and the Use of Platform Analytical Procedures for Biologics

2. Scope

The purpose of this document is to outline standard practices for analytical method lifecycle steps where technical guidance or standardized approaches are currently lacking, including Analytical Method Transfer (AMT) and Analytical Method Comparison (AMC) for replacing methods. This document will also provide a standard practice for the validation, qualification, and implementation of Platform Analytical Procedure (PAP) methods. The proposed standard is intended to support lifecycle management of analytical methods for biologics including manufacturers, testing laboratories, and regulatory authorities.

The information will benefit users by providing the design of consistent and scientifically sound studies, enabling successful AMT, PAP, and AMC implementation, and improving quality of regulatory submissions.

Part 1 (AMT) includes the application of specific AMT models and execution designs. Appropriate statistical methodologies are provided, as well as risk evaluations and the process and rationale for setting risk-based AMT acceptance criteria [2, 5].

Part 2 (PAP) includes a description of the PAP concept and the initial validation study considerations. A standardized risk-based verification concept for follow-on products is described. The PAP validation/verification study design, setting of acceptance criteria, relevant rationale, and documentation practices are provided [6].

Part 3 (**AMC)** includes a description and selection rationale for appropriate comparison models (equivalence, non-inferiority, or superiority). The process and rationale for setting risk-based comparison acceptance criteria is established [2, 5]. Case studies are provided for qualitative and quantitative AMC studies. Additional case studies are provided for replacing existing methods with novel/alternate methods which usually requires corresponding novel/alternative specifications.

This standard is intended to compliment principles described in the ICH Q2(R2), Q12, and/or ICH Q14 [2-4] guidance documents, providing practical technical information.

3. Terms and Definitions

Analytical Method Comparability (AMC)	The documented process that confirms or demonstrates suitable comparability for analytical method replacements. AMC studies are also performed to confirm/demonstrate that a modification to an existing method does not lower the method's performance capabilities to control relevant product safety, efficacy, or quality attributes.
Analytical Method Transfer (AMT)	The documented process that qualifies a laboratory (i.e., receiving unit) to perform an analytical test procedure that originates in another laboratory (i.e., transferring unit or sending unit).
Platform Analytical Procedure (PAP)	A multi-product method suitable to test quality attributes of different products without significant change to its operational conditions, system suitability and reporting structure. This type of method would apply to molecules that are sufficiently alike with respect to the attributes that the platform method is intended to measure [3,4].
Co-validation	Demonstration that the analytical procedure meets its predefined performance criteria when used at different laboratories for the same intended purpose. Co-validation can involve all (full revalidation) or a subset (partial revalidation) of performance characteristics potentially impacted by the change in laboratories [3,4].
Comparative study	When both the sending and receiving laboratories perform a validated analytical procedure on the same manufacturing batch and compare the resultant data between the laboratories.
Performance Verification	Consists of assessing selected performance parameters to provide evidence that the analytical procedure performance complies with the relevant validated procedures (compendia and PAP).

4. Acronyms and Abbreviations

AM(x)	Analytical method (x) (synonymous with AP(x))
AP(x)	Analytical procedure
AMC	Analytical Method Comparison
AMD	Analytical Method Development
AMM	Analytical Method Maintenance
AMQ	Analytical Method Qualification
AMT	Analytical Method Transfer
AMV	Analytical Method Validation
ATP	Analytical Target Profile
BLA	Biological License Application

CPV	Continued Process Verification
DL	Detection Limit
DOE	Design of Experiments
DP	Drug Product
DS	Drug Substance
FTIH	First Time in Human
HPSEC	High Performance Size Exclusion Chromatography
ICH	International Conference on Harmonization
IMP	Investigational Medicinal Product
LT	Less Than
LOD	Limit of Detection
LOQ	Limit of Quantitation
MOA	Mode of Action
NIST	National Institute of Standards and Technology
NLT	Not Less Than
NMT	Not More Than
OOS	Out of Specification
OOT	Out of Trend
(p)CQA	(potential) Critical Quality Attribute
PB	Plackett-Burman (DOE)
PAP	Platform Analytical Procedure
POC	Proof of Concept
PQ/PPQ	Process (Performance) Qualification
PV	Process Validation
QbD	Quality by Design
QL	Quantitation Limit
QTPP	Quality Target Product Profile
RU	Receiving Unit
SOP	Standard Operating Procedure
SU	Sending Unit

5. Analytical Method Transfer (AMT)

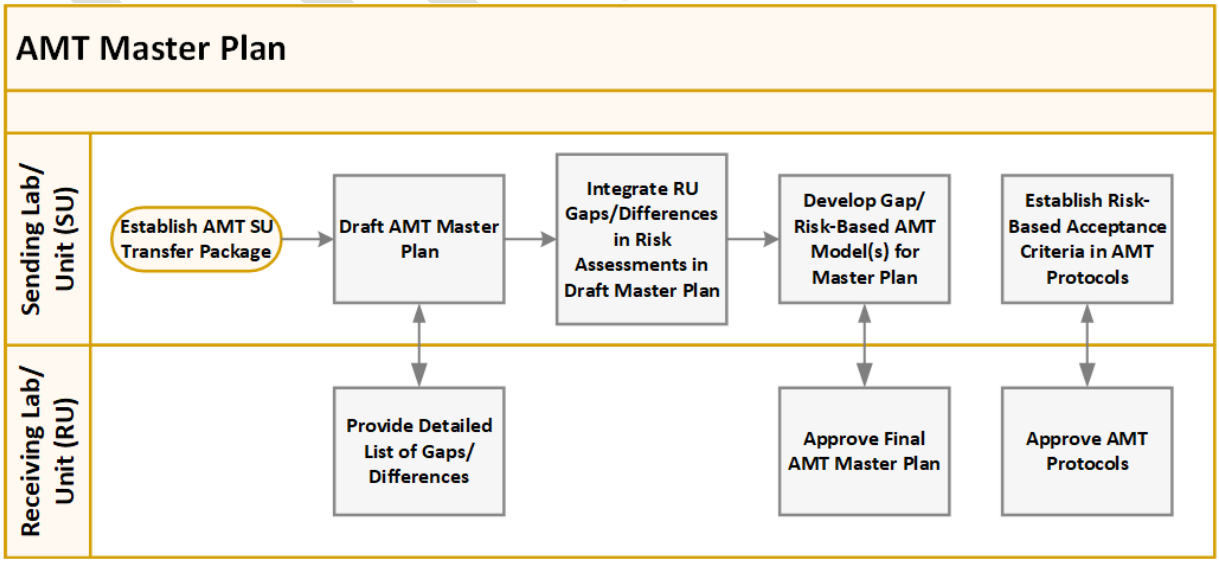
5.1 Description

Analytical Method Transfer (AMT) is a documented process that qualifies a laboratory to use a method that originated in another laboratory. AMT may occur at any point in the method and product lifecycle. The stages of an AMT include an assessment of the new (or receiving) laboratory, development of the analytical method transfer protocol, including justified acceptance criteria, training of the new laboratory (if needed), execution of the transfer data analysis using appropriate statistical tools, and the conclusion(s) documented in an AMT report. Specifically, consideration should be given to the availability of required analytical and supporting equipment, software, critical reagents, standards, controls, and analysts who are skilled in the relevant analytical techniques as well as the qualification status of all materials, equipment and analysts. Pending the development stage, available documentation for the method procedure, method validation report, available historical data, and any prior method transfer reports should be reviewed to assist in the gap assessment of the receiving laboratory.

A master transfer plan is a tool that can be used to document all relevant risks for analytical method transfer; illustrated in **Figure 4**, below. Based upon the assessment of those risks, the transfer strategy for individual methods should be documented in the master transfer plan. The master plan should state the roles and responsibilities of the sending and receiving labs, the gap/difference analysis of the receiving lab, as well as the risk assessment summary and outcome(s), as suggested in **Table 2** below**.** The capabilities of the receiving lab to perform the method should be assessed through a gap analysis exercise and documented in the master transfer plan.

Factors to consider in the master transfer plan for assessing all relevant risks leading to a potentially unsuccessful transfer of each method are outlined in **Table 3 (below)** and illustrated in **Figure 4**. Methods with lower risk, as determined by the method performance characteristics relative to the product risk or the similarity of the new method to a method already established at the receiving lab, may require reduced testing when compared to a higher risk method. Higher risk methods, evaluated as such due to their performance relative to the product risk, complexity of the method relative to the frequency of testing, or lack of familiarity with the methods at the receiving lab, will require a comprehensive AMT study design.

Figure 4: Suggested AMT Master Plan Development



294 **Table 2: Suggested AMT Responsibility Matrix**

Lab	Suggested Responsibilities
Sending lab	- Compile QC/process data. - Provide all relevant documentation related to procedure. - Provide training, as required. - Establish the transfer package. - Write transfer protocol based on requirements of both labs and knowledge of method prior to transfer. - Establish protocol acceptance criteria. - Allocate resources for training and transfer study. - Provide critical reagents and samples, as needed, and all relevant qualification records. - Provide troubleshooting support. - Approve the transfer report.
Receiving lab	- Review the transfer package. - Define the transfer process. - Identify training requirements. - Verify all equipment/systems are available and appropriately qualified. - Inform the sending lab on potential issues identified (such as different suppliers for critical equipment). - Allocate resources for training and transfer study. - Analyze transfer data. - Write the transfer report. - Inform the sending lab of the outcome of the transfer. -Approve the transfer report.

295

296 NOTE: The roles and responsibilities may differ and will depend on the nature of the quality agreement or
297 relationship between the sending and receiving units.

298 **Table 3: Risk Considerations for Analytical Method Transfer**

Risk Evaluation Categories	Risk Variants	Example(s)	Expected Potential Risk/Impact
Attribute(s) Criticality -	CQA, p(CQA), IPC Starting/Raw Material, etc.	DP Purity Release and Stability test: Degradation products during storage.	All DP purity/stability tests transferred are higher risk to patients and/or sponsor. A significant post-transfer bias at RU towards lower values could result in patient safety impact as higher impurity levels may not be adequately controlled (specification(s) unchanged). A significant bias towards higher values may lower the corresponding manufacturability (CpK) and increase OOS likelihood. Representative DP lot samples used in

			a direct SU-RU equivalence study, considering all relevant risk variants, should be tested with risk-based, justified AMT acceptance criteria (see **Figures 2 and 3).**
Intended Purpose for Testing	Routine Release and/or Stability Test, PPQ (only), Process/Product Characterization, etc.	Product/Product characterization test used in product comparability studies.	Risk to potentially observe significant product comparability difference by testing pre-manufacturing sample(s) pre-AMT vs. post-manufacturing post-AMT can be mitigated by testing all product comparability samples post-transfer at the RU. This would lead to lower-risk product AMT testing and less stringent acceptance criteria (if needed).
Product Lifecycle	Commercial, PPQ, Pivotal, Early-Stage Product Development	Potency testing in DP samples for release test at filling site for early-stage (phase 1) development. The method was qualified but has not been validated.	The AMQ extension study is typically executed using significantly fewer comparison testing samples/replicates at this clinical/development stage versus an AMV extension study for late- and/or commercial-stage(s). The equivalence testing acceptance criteria are also typically less stringent as DS/DP specifications are wider than those used for late stage/commercial products. However, the criticality of a potential potency result drift/variation to the clinical trial outcome(s) should be considered. For example, a very narrow therapeutic window may require higher AMT success confidence (more samples/replicates) and tighter equivalence acceptance criteria.
PAP SU/RU experience/ status	PAP already in use in SU and RU, PAP ready at RU but not routinely used (yet), "New" analytical technology/method at RU	PAP Purity test for DS is validated and approved (via market authorization) method for SU and RU	Following a successful previous post-AMV AMT study execution for a highly similar product, the existing PAP experience at the RU with this method performance significantly lowers the risk to patient and/or firm. An AMT waiver could be justified when the relevant PAP verification study execution demonstrated the suitability of the product for use of this PAP method.
Manufacturability **(CpK)**	CpK falls between 1.50 -2.00	DP Purity release test is validated for a commercial product and the relevant manufacturability is relatively high (CpK = 1.80)	A relatively high CpK pre-AMT typically supports the setting and justification of wider equivalence acceptance criteria.
Degree of Automation and/or Risk of Human Error	Manual procedure steps to fully automated Few process steps to many/complex process steps	A manual, complex, multi-day execution potency test is to be used to release test at RU	The RU bias/variation is controlled by analyst-to-analyst process step(s) execution similarity and/or the established criticality and controls for all relevant process steps (ex., pipetting techniques for micro-volumes). Criticality and required procedural control(s) of all variable procedure execution steps may not be established before AMT. See also potency test example under Product Lifecycle above.
Gap or Difference Mitigation	Potential significance of established gaps/differences	Different analytical instrument is used at RU for a DP Purity release test	A significant bias caused by a different instrument for an automated purity test method may lower or raise the corresponding manufacturability (CpK) and OOS likelihood.

5.2 General AMT Strategy

The strategy used for an analytical method to be transferred should be based on the risk factors outlined in **Table 3**. The outcome of the risk assessment will determine the AMT strategy. Common approaches to AMT are described below; other approaches may also be acceptable in certain situations.

a) **Co-validation** – Co-validation is a transfer model that includes method validation data generated at both the sending and receiving laboratories. A potential benefit of co-validation is that it enables method validation and method transfer to be performed at the same time under one protocol, which is advantageous in situations where there is limited time for the qualification of the receiving laboratory. If the RU does not obtain equivalent results when compared to SU, this may impact the overall AMV study success".

b) **Comparative study** – A comparative study is when both the sending and receiving laboratories perform a validated analytical procedure on the same manufacturing batch[es] and compare the resultant data between the laboratories. Acceptance criteria determine the equivalence of the two laboratories. Historical and validation data may be used when appropriate for parts of the method transfer study. The sending laboratory typically has collected a significant amount of historical data for the method performance in addition to test results for the samples to be tested at the receiving laboratory. Acceptance criteria for the AMT should be derived following the process and conditions as illustrated in S**ection 1**, *Introduction* Acceptance criteria can be set based on previous validation/qualification studies and/or recent routine QC testing data with respect to the relevant product or material specifications.

c) **Performance Verification –** Performance verification can be used if the receiving laboratory already performs a PAP method for a similar product or for another type of sample for the same product. In this case, a formal method transfer may not be required. Any reduced or waived prospective study considered should be properly justified**.** This concept is likely to be used for PAPs, which are not product specific. For example, if a PAP has been established in a receiving laboratory for testing MAB-A, and the same PAP is used to test MAB-B, then the PAP suitability can be verified and/or waived as appropriate without a side-by-side comparative test.

d) **Waiver** - A formal AMT study may not always be necessary, and a transfer waiver can be acceptable, provided the scientific justification is valid and documented. Possible scenarios for transfer waivers include, the scenario in which the receiving lab may already perform the same PAP; In this case, a formal AMT study and/or performance verification may not be required.

5.3 Design of AMT Test Studies

Prior to initiating the AMT, a gap assessment should be completed to assess all sources of variation that may contribute to differences in method performance and potential bias in the data. The AMT protocol should outline the study design, specifying method performance characteristics to compare, samples to test, justified acceptance criteria, and the statistical methodology used to evaluate the results.

5.3.1 Selecting AMT Performance Characteristics

The intended purpose of the method should be considered to justify the rationale of the study design and acceptance criteria for each method transfer. **Table 4** outlines the performance characteristics to be compared between laboratories for different types of methods. Other performance characteristics covered during the validation studies may also be considered.

Table 4: Examples of Performance Characteristics to be Assessed during AMT

Type of Methods	AMT Performance Characteristics Examples
Identity tests	System suitability, specificity, qualitative comparison (if applicable)
Impurities (quantitative) – process- and/or product-related	System suitability, precision and accuracy; consider several concentration levels: minimum reportable quantity and / or Quantitation Limit(s) and 120% of the product specification; stability samples may need to be included to assess stability-indicating capabilities, as relevant
Impurities (qualitative, limit)	System suitability, Detection Limit(s)
Assay – content and potency	System suitability, precision and accuracy, range, and stability samples may need to be included to assess stability-indicating capabilities, as relevant

5.4 Sample Selection and AMT Study Design

Representative sample types should be selected, as appropriate, for the application and intended purpose of the method. For example, when comparing stability-indicating methods, degraded samples can be directly compared by both laboratories. Sample preparation, such as different spiking levels, may significantly contribute to variation in method transfer results. Therefore, it is important to prevent inconsistent sample preparation to reduce this potential variation and/or bias during the method transfer process.

It is recommended to use multiple sample concentration levels for products of different strengths and/or matrix variation to ensure that the analytical method performance remains sufficient over these ranges. These extended AMT results may provide additional information as to whether both labs can produce similar accuracy (matching) and precision (reliability) results over the potential range of expected results.

A sufficient number of samples and testing runs should be executed to establish equivalence between the two laboratories. The ability to detect a difference or establish confidence that no difference exists is directly dependent on the number of determinations (number of results from independent runs) for each laboratory. Two common approaches for choosing the sample sizes can be envisaged based on method complexity and its known variability – fixed or variable execution matrix.

A fixed AMT execution matrix does not integrate known method variation and therefore has an identical set of comparative data generated between both laboratories for each method transfer executed. A fixed matrix can be more advantageous when transferring multiple products to/from multiple locations. The fixed number of replicates and acceptance criteria are set for the relative difference between means found at both laboratories or by equivalence testing using two one-sided t-Test (TOST, see **Table 5** below). The study typically addresses at least two independent factors (e.g., analysts and days) known from the AMV studies to (potentially) impact the reliability of routine test results. Intermediate precision at both laboratories can be evaluated from this data set; however, when a more detailed result interpretation is desired at the receiving laboratory, a more extensive set-up may be appropriate. Alternative AMT statistical approaches may also be suitable if appropriately justified. It is recommended that alternative statistical approaches are discussed with the regulatory agency(ies) in advance.

Table 5: Examples of AMT Fixed Execution Matrix

Method Type	AMT Execution Matrix Examples	AMT Protocol Performance Characteristics Examples
Identity	Results for multiple positive and negative samples should be compared when comparing Specificity (differentiation capability). Blind sample testing may be required for non-automated identification systems.	System suitability met, similar or identical differentiation capability should be demonstrated.
Impurities (quantitative) – process- and/or product-related	Two operators and/or instruments on different days; consider spiking at different levels for confirming precision, accuracy, and Quantitation Limit(s).	System suitability met, Quantitation Limit(s) confirmed, TOST (two one sided test at 95% confidence level). Note: Results for different batches may not be pooled unless normalization prior to comparing can be justified.
Impurities (qualitative, limit)	Results for multiple samples below and above the Detection Limit(s) should be compared.	System suitability met; similar Detection Limit's should be demonstrated.
Content uniformity, purity, and/or potency	Two operators and/or instruments on different days. When testing products of multiple nominal strength/concentrations, consider a bracketing approach, using batches of minimum and maximum nominal strength/concentration.	System suitability met, TOST (two one sided test at 95% confidence level). Note: Results for different batches and/or nominal strengths/concentrations may not be pooled unless normalization prior to comparing can be justified.

A variable execution matrix takes into consideration variations in results due to the method and may require a larger comparison data set for highly variable methods (see **Table 6**). The selection of the AMT study design should be considered for a given situation. For example, a variable execution matrix may be advantageous when transferring bioassays with a relatively high degree of test result variation. On the other hand, when transferring multiple assays simultaneously, in support of a product manufacturing transfer, a fixed execution matrix may be the most effective process.

For highly variable methods, such as certain bioassays, the use of fixed matrices may not be adequate; a variable execution matrix should be employed. An appropriate sample size to sufficiently power the study can be determined using the risk-based approach. USP <1033> *Biological Assay Validation* [7] provides guidance on how to determine test sample size based on desired power level (1-Beta) and confidence level (1-Alpha). Typically, Alpha is set at 0.05 and power is desired to be 80%. Sample size is determined based on these values of Alpha and Beta as well as the study design.

Table 6: Examples of AMT Variable Execution Matrix

AMT Design Parameter	Suggested Considerations
How many representative batches – Matrix (number of different sample types and/or batches to be evaluated)	Bracketing of the expected active protein concentration range should be considered. The selected materials should be representative of routine samples. Retain samples, reference standards, samples at the extremes of acceptance limits, stability samples, and/or spiked samples should be used depending on the situation. For impurity tests, samples may be spiked or degraded so that the level of the impurity is below and/or above the QL (and/or specification limit). If samples with a measurable impurity level are not available, then it may be necessary to prepare spiked samples to evaluate the accuracy and precision of measurable amounts of impurity/degradation levels during the AMT studies. If there are differences in the formulation, adequate testing of the range of formulation differences should be included. The rationale for the selection of representative AMT samples should be documented in the AMT protocol.
How many replicates per sample and lab? (Number of independent runs)	The number of replicates depends on the Repeatability and Intermediate Precision performance of the method to be transferred and the desired confidence level(s) for meeting product specifications. The AMV report and other related data sources (for example, routine test results) should be reviewed.
How many Intermediate Precision variability factors?	At least two critical factors should be selected based on prior knowledge of which factor(s) may have the greatest expected impact on test result variation.

5.5 Acceptance Criteria and Statistical Evaluation

Acceptance criteria should be established and justified for the allowed difference(s) between the sending and receiving laboratories prior to the transfer. Risk assessments following similar concepts as those described in **Section 7.5,** *Risk-based approach for acceptance criteria*, should be performed when establishing acceptance criteria. The intended statistical evaluation methodology should be considered. The statistical methodology applied, such as TOST or other approaches, as well as the established acceptance criteria should be justified. Typically, equivalence testing by TOST is applicable for quantitative test methods, while probability testing is applicable for qualitative test methods. Equivalence testing, by TOST is statistically satisfactory if the confidence interval for the difference in means between the two laboratories falls within an acceptable interval $[-\Theta, +\Theta]$. The interval should define the largest difference that can be accepted between the laboratories while not significantly impacting the RU test results. When comparing results from two laboratories, the interval is centered around zero, reflecting the fact that there is no (significant) bias between the laboratories. Based on the two sets of results and the pooled standard deviation from the two laboratories, a confidence interval is calculated for the difference in sample means. The null hypothesis that the means are not equivalent is rejected once the confidence interval is strictly found within the acceptance interval. The two sets of results are therefore considered as equivalent.

For late-stage/commercial AMTs, acceptable differences between laboratories for method performance characteristics of quantitative methods such as Accuracy and Intermediate Precision should be controlled, using an outside-in approach (Section 1, *Introduction*; **Figure 2**), based on historical data and/or previous AMV protocols/reports with respect to the specifications.

For early-stage development AMTs, acceptable differences between laboratories for the method performance characteristics of quantitative methods such as Accuracy and Intermediate Precision should be controlled, using an inside-out approach (Section 1, *Introduction*; **Figure 3**), based on existing method performance knowledge.

5.5.1 AMT Case Study for a Potency Test Method

A validated potency test method used for commercial release of biological drug product is to be transferred from the original QC laboratory to another QC laboratory to release drug product (DP). The analytical method generates potency (dose) results for lyophilized DP. Three nominal doses (500, 1000, and 2000 IU/vial) using an identical formulation are available. Routine release testing is performed using three replicate preparations from each of three vials.

Before analysis the content of a vial is reconstituted with 5.0 mL of water for injection (WFI) and the potency is measured in IU/mL (100 – 400 IU/mL). The analytical method procedure's system suitability criteria include parallelism requirements for the sample, assay control, and reference standard curves. A variable AMT execution matrix is used. The AMT study design, acceptance criteria, and AMT study results are summarized in **Tables 7 and 8,** respectively.

Table 7: Potency Test Case Study - AMT Study Design and Acceptance Criteria

Characteristics Evaluated	**<u>Accuracy:</u>** **The relative difference between lab means (two one-sided 90% CI) should fall between not be less than -Θ = 10% and not more than +/- 10.0%.** The 10.0% difference limit was set and justified with consideration of the product specification(s) (75-125%), manufacturing capability (CpK = 1.20), and the test method's long-term intermediate precision (RSD/CV = 6.0%). **<u>Intermediate Precision:</u>** RSD 6.0 % for all sample types, with appropriate homoscedasticity throughout the potency range (derived from validation results and long-term assay control monitoring). This means that any RSD from a sample of n=8 should not exceed **9.4 %**
Number of Replicates	$N_{replicates}$ = at least **23** independent replicates The confidence interval for the lab-to-lab difference for N determinations to less than the [10.0%, +10.0%].
Samples to test	N_{level} = 3 samples covering the range of potency/dosing results: Lowest dose 500 IU/vial or 100 IU/mL Medium dose 1000 IU/vial or 200 IU/mL Highest dose 2000 IU/vial or 400 IU/mL

Study design for each of n=3 samples	Number of operators, n = 2 Number of days, n = 2 Number of independent replicates per day per operator, n = 2 Number of instruments, n = 1 (only one instrument available at RU) N = 8 in each lab for each of n = 3 potency levels. **Results are converted to "% recoveries vs. expected" to allow pooling for n=3 potency levels.** **N_{Total} = 24 individual observations will be recorded for each laboratory.**

Table 8: Potency Test Case Study - AMT Study Result Summary

Separate and Pooled Potency Levels Evaluated	**Sending lab**		**Receiving lab**	
	Statistical Parameters	**%Recovery vs. Theoretical Potency**	**Statistical Parameters**	**%Recovery vs. Theoretical Potency**
TOST with acceptance criteria [-10.0%, +10.0%]	N1	24	N2	24
	Mean1	101.1	Mean2	99.4
	SD	3.5	SD	4.2
	RSD	3.4	RSD	4.2
	Pooled SD[(2)]	3.9		
	Mean1-Mean2	1.7		
	t-value	1.679		
	Upper 90% CI limit	4 (3.6)		
	Lower 90% CI limit	0 (-0.1)		
Transfer Acceptance Conclusion		Pass		

Figures 5 and 6 illustrate the AMT case study results using a Box Plot format. The boxes represent the 25th - 75th percentile distribution of the results for the two laboratories. Medians (line in the box) and means (cross in the box) are approximately centered while the medians are equidistant from the box hinges, providing a visual indication for a normal data distribution(s) among data points within each laboratory set.

When interpreting the results in **Figure 5**, one potential outlier (lower open circle outside of the whiskers) is observed in the sending lab; however, this does not change the overall interpretation for the demonstration of lab-to-lab equivalence. Also, the variation in the test results (wider 25th – 75th percentile boxes) appears to be higher in the receiving laboratory for the pooled AMT results (**Figure 5**) as well as for each of the three potency levels (**Figure 6**) which may be attributed to less test method execution experience or could have other reasons. In addition, **Figure 6** shows a relatively large mean SU-RU difference for the lowest potency level. Although this could be a random observation, one of the possible reasons to be considered for prospective monitoring at the RU should be the sample preparation/dilution conditions.

Figure 5: Potency Test Case Study - Box Plot of AMT Study Result Summary

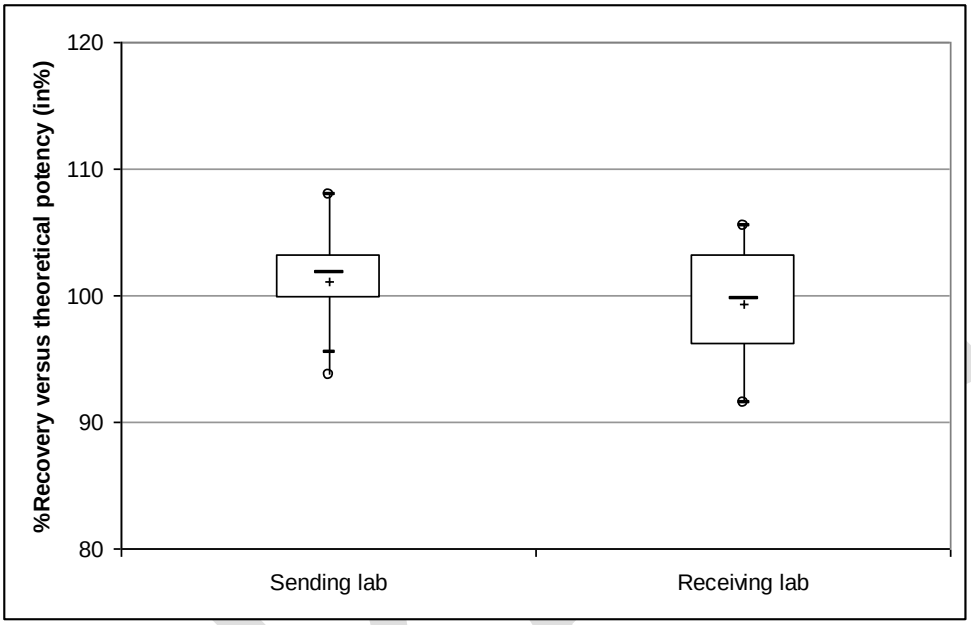
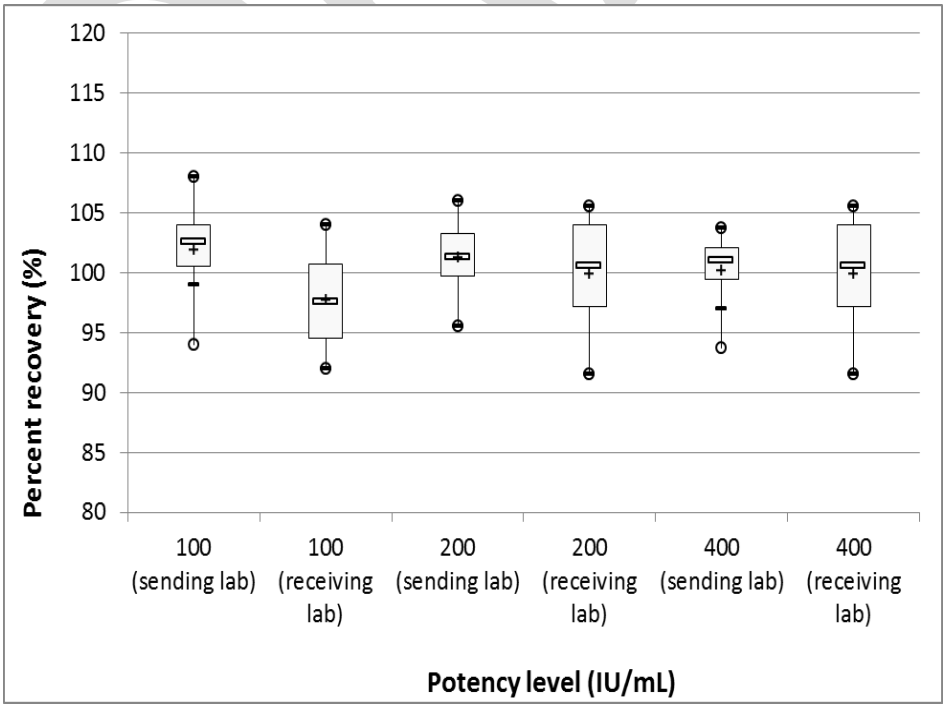


Figure 6: Potency Test Case Study - Box Plot of AMT Study Results for Each Potency Level



5.5.2 Considerations for Sample Preparation

Test samples used in AMT studies should be carefully prepared, shipped, and stored to mitigate any differences in test results due to sample preparation. In addition to sample preparation, shipping, and storage conditions, sample homogeneity and sample stability should also be considered for the AMT studies. Some additional considerations are listed below and in **Section 7.5.3,** *Points to consider when selecting test samples for evaluation of analytical method comparability.*

- The set(s) of sample preparations should be prepared as similarly as possible.
- If applicable, suitable reference and/or control material should be selected and included in each single analytical run.
- Reference material should be sufficiently characterized.
- Sufficient sample and reference material aliquots should be prepared to allow for additional testing in case invalid test results are generated.
- All test samples and reagents should be appropriately documented in accordance with GMP principles. Distribution and storage conditions should be defined based on expected stability of all material to be tested.

5.5.3 Deviations and Failures

Any deviation to the approved AMT protocol should be properly justified and approved. If any of the acceptance criteria stated in the protocol are not met during the execution of the AMT study, an investigation should be performed, and proper corrective and preventative actions implemented. More detail on the failure investigation process is provided in **Section 5.7,** ***AMT Failure Resolution*****.**

5.5.4 Invalid Assays

Assays which do not meet system suitability criteria specified in the method SOP are not included in the analysis of results for comparison to the protocol acceptance criteria. Additional assays should be performed to replace the invalid ones. However, invalidated assays should be described in the AMT report with the rationale for their exclusion.

5.5.5 AMT Study Extension

In case the initial sample size results ($N_{1)}$ have generated unexpected wide confidence interval(s) not allowing for a clear pass/fail conclusion, if it can be justified, the study could be extended whenever no apparent (significant practical bias) root cause exists, with an additional pre-determined set(s) (N_2) of independent runs. All results should then be pooled ($N_1 + N_2$) before final interpretation. A protocol amendment should be approved before execution of additional data sets.

5.6 AMT Failure Resolution

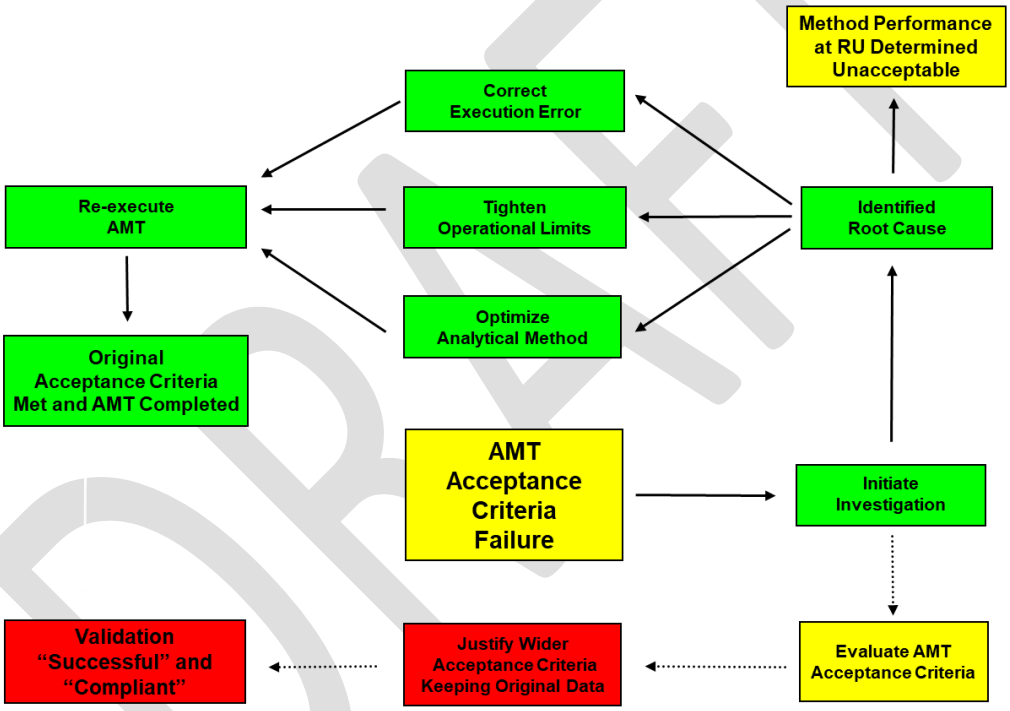
Whenever AMT results fall outside the pre-established protocol acceptance criteria, and thus, an AMT failure occurs, an investigation should be initiated. The investigation process is illustrated in **Figure 7**. Using the upper loop in **Figure 7** leads essentially to three work levels - correct execution error, tighten operational limits, or optimize analytical method. The time and effort required typically increases from the top box ("correct execution error") to the bottom box ("optimize analytical method"). When re-executing after correction of root cause, for example, spiked proteins were partially adsorbed at or before sample preparation to glass containers causing low % recoveries, the actual method performance should not change or improve.

Unexpected high variation may require running the method system under more stringent operational limits. For example, the allowed sample preparation range(s) or overall testing time could be reduced to decrease

variations in degradation or other inconsistencies that impact the test results. An alternative to this is to tighten the qualification requirements for test system components such as the operator proficiency qualification testing requirements. Although these limitations may somewhat indicate that the method performance is not readily reproducible, this should lead to an improvement of the method performance at the receiving unit (RU) as well as the sending unit(s) (SU). Any necessary method performance improvements resulting from tightening relevant operational limits should be implemented at all sites (SU and RU(s)) to prevent procedural and performance variation/drift across testing sites.

When optimizing the analytical procedure, this may have the greatest effect on future method performance across testing sites. It is also the most rigorous recovery process and may require a significant time to complete. Because timely completion of projects is often critical for patients and manufacturers, all aspects should be considered that may impact patient safety, product quality, time required, and chance of success for the recovery process.

Figure 7: AMT Failure Recovery Process



6. Platform Analytical Procedure (PAP)

6.1 Description

A platform analytical procedure (PAP) can be defined as a multi-product method suitable to test quality attributes of different products without significant change(s) to its operational conditions, system suitability, and reporting structure. This type of method would apply to molecules that are sufficiently alike with respect to the attributes that the platform procedure is intended to measure [2].

6.2 Establishment of Platform Analytical Procedures

Different approaches exist to establish a PAP. An established method, developed and validated to control one or more quality attribute(s) in a given product may be applied for a new product. In this case, analytical development for the new product can be abbreviated and focus on the verification that the established analytical method can control the attribute (s) in the new product without significant changes in operating conditions. When products exhibit similar structural properties or are derived from similar manufacturing processes, a PAP may be used after abbreviated development and product-specific verification. In this case an evaluation of existing validation data against a common set of acceptance criteria for respective validation characteristics can be performed to prove that acceptable method performance is achieved independently from the different products and forms the basis of the establishment of the method as a PAP. Typically, when developing and validating an analytical procedure towards suitability for use as a PAP, the procedure is established when it can be used for three different products.

Another approach for the establishment of a PAP is to include multiple products already at the design and execution stage of development and validation. Here the same principles apply as described above as the method should be applicable without significant changes for the different products. Validation data from the different products demonstrate that the validation acceptance criteria can be met for all products used in the study. This approach is often used when an existing PAP is replaced by a more suitable/preferred PAP.

Suitable for use as a PAP are those which by design and intent can generate accurate and reliable results for different products without significant methods changes. The ability of the analytical procedure to be used as a PAP for multiple products is affected by the following factors, which need to be carefully considered:

a) The technology used. Analytical methods that are based on physical principles are usually less sensitive to variation in the matrices and products therefore can be applied to a wide range of conditions, whereas methods targeted to biological functions are often highly specific to the target analyte.
b) The products and matrices subjected to testing. PAPs are more easily implemented in scenarios where tested products are closely related by structure and function and/or matrices are similar.
c) Intended use of the method. The methods for testing of target protein(s) or product-related impurities are generally used on a closely related molecule, whereas methods for process-related impurities or excipients may be used on a wider range of products. Also, qualitative methods (e.g., identity by western blot) are generally easier to use for similar products.
d) The amount of available information about the method. Well-characterized analytical methods for which substantial amount of information is available are more easily converted to PAP. Prior knowledge from literature data describing the performance of the method under different circumstances along with a company's own data may help in preliminary determination of method's range of applicable molecules.
e) The use of reference standards or materials. The quantitative methods which are using product-specific reference standard/material and or controls generally require more effort to establish suitability for use and maintain as PAP.

It is important to understand that the considerations listed above are general and should only be used for assessment of feasibility of using the method in the PAP role. These factors may be helpful in identifying the gaps in existing knowledge and designing appropriate development and verification studies. It should never be assumed, without verification, that a method is fit to test even for closely related products. Examples include methods such as protein content determination by UV spectroscopy, control of size variants by SE-HPLC, excipient quantity in formulated DS and/or DP. Relevant applications include monitoring product-related impurities and process-related impurities. These methods are typically suitable for PAP use if the lack of potential matrix interference can be established. Similar to compendial methods, a previously validated PAP method may not require full validation for each new product or sample type.

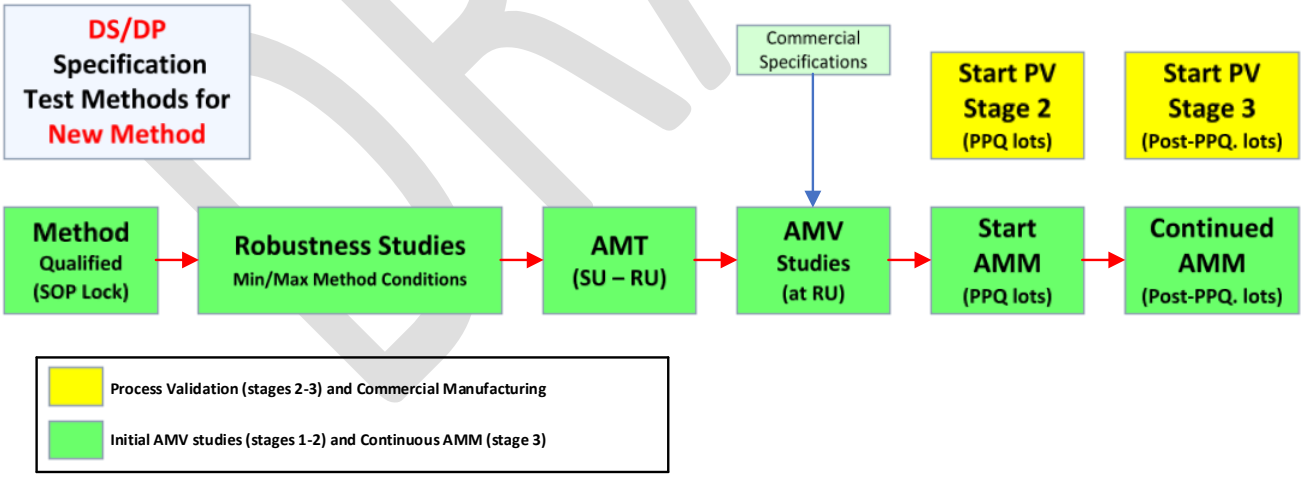
It is important to understand that establishment and use of PAP involves not only analytical development and quality control units, but also other functional units in the integration into the existing company quality systems. Specifically, (1) procedures and criteria allowing assessment of the PAP suitability for multiple products should be developed; (2) strategies for coordinated submission and maintenance of PAP information in multiple regulatory dossiers should be considered; (3) comparative monitoring of PAP performance for all products tested should be integrated into existing and future continued process verification (CPV) programs.

6.3 PAP Concept and PAP Use Benefits

A typical sequence for the lifecycle of a new method is shown in **Figure 8** below. Alternative, risk-based lifecycle sequence steps may also be appropriate for highly accelerated product development programs. The illustrated AMV process can be an ideal step sequence because the method standard operating procedure (SOP) is locked with the AMQ step and all release/stability data for Investigational Medicinal Product (IMP) is generated by a qualified method, confirmed for suitability of its intended use. AMQ can be considered to be equivalent of phase-appropriate validation and is an abbreviated form of AMV studies, using less risk-based acceptance criteria. The AMQ study confirms that the method procedure/conditions is/are suitable for testing of the intended product/sample(s).

The method robustness studies are executed after the completion of any necessary method optimization steps but before the AMT/AMV study execution so that an optimized, robust method is then validated. The AMV studies are then executed prior to PQ studies to ensure PQ studies are performed with validated methods. As a result, this will largely remove the remaining analytical uncertainty when conducting the PQ studies. Like the CPV stage (PV stage 3), the Analytical Method Maintenance (AMM) program can then start with the first PQ lot tested. The illustrations below are used to provide a framework for the process.

Figure 8: Typical Lifecycle Steps for a New Routine Method

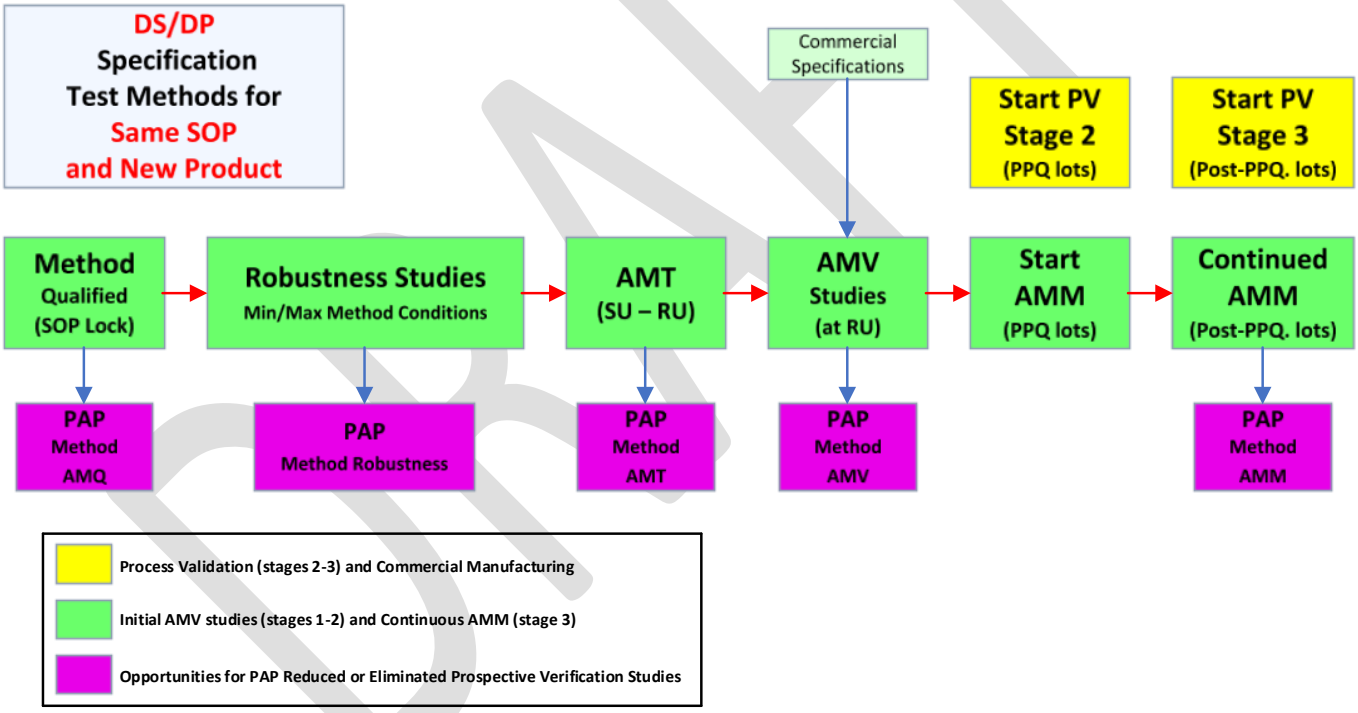


The additional detail provided can then be used to evaluate and illustrate which lifecycle steps can be reduced for PAP methods. The method is qualified prior to release testing and the SOP is then "locked" until further improvements may become necessary. For specific guidance regarding AMQ studies and post-AMQ steps, see PDA TR 57-2 [8] and PDA TR 57 [2] respectively. Robustness studies are recommended to be completed as part of the analytical method development [4] but can be completed any time post AMQ, and pre-AMV. However, completing robustness studies early will result in less overall risk for completing rapid and

successful AMT or AMV studies. The design of robustness studies should be based on a risk assessment, considering available knowledge from development studies and prior knowledge. Parameters assumed to affect the performance of the method should be included in the design of the robustness studies. The outcome of robustness studies informs the definition of the control strategy of the method.

For the initial AMV studies, the method lifecycle steps are typically executed sequentially. As current drug substance/product (DS/DP) specifications frame the intended use of the method, and thus drive the AMV study design and acceptance criteria, it is therefore a critical confirmation step that the suitability of the method's intended use can be extended into any post-AMV DS/DP specification revisions. Following the successful AMV study completion, the analytical method maintenance (AMM, AMV stage 3) typically starts. The method validation stage 3 is practically initiated before process validation stage 3 (CPV) which is typically initiated with the first post-licensure manufactured lot. If the method is used unchanged for a similar new development product, each of the five lifecycle steps can be significantly reduced or eliminated as illustrated in **Figure 9** and described in **Table 9** below.

Figure 9: PAP Opportunities following successful AMV Study Completion



Starting with the reduced PAP AMQ studies of the validated PAP method, all possible reduction opportunities for prospective studies to qualify a new analyte for the PAP for all five method lifecycle steps are summarized below. As some or most of the initially completed prospective study results are used as the foundation for each of the validation stages, additional prospective verification studies are proposed here.

Table 9: (Prospective Study) PAP Reduction Opportunities for Method Lifecycle Steps(s)

AMQ – Similar to AMV studies (see below), AMQ studies can be reduced to confirm that different protein concentration(s) or formulation(s) do not impact significantly the validated characteristics; Accuracy (and inferred Specificity). A spiking study below/over DS/DP specification(s), DL, or QL, as relevant, can be used.
Robustness – previous robustness results can be used without repeating this study since the method has remained unchanged.
AMT - previous AMT results can be used since the method has remained unchanged and has been in routine use at the receiving laboratory since the initial transfer.
AMV – See above under AMQ and also below in the case study.
AMM – Stage 3 of AMV was already initiated at the receiving laboratory for this method. All relevant test system controls have already been established.

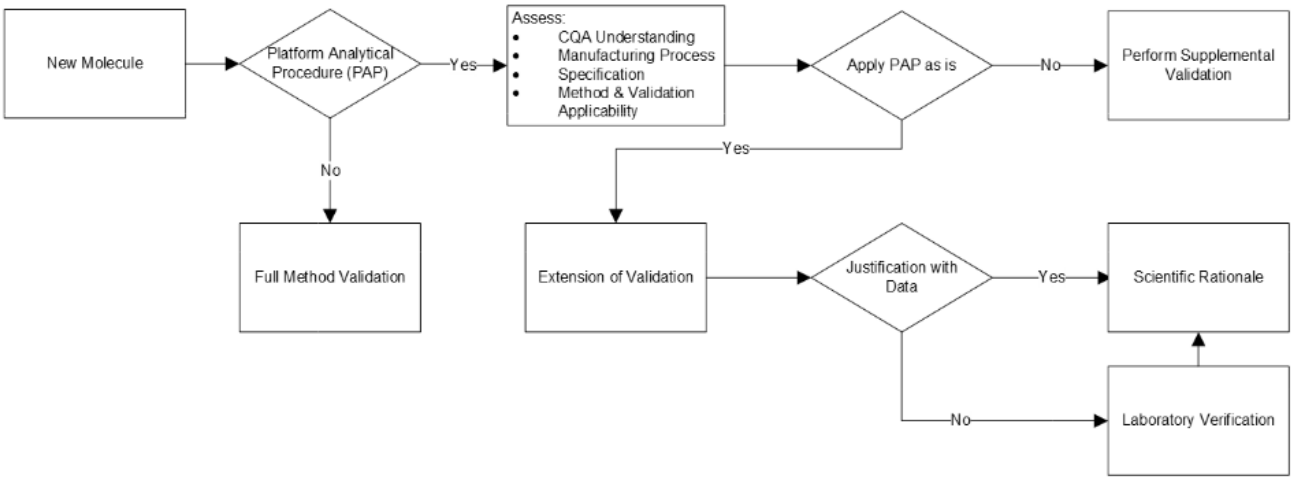
6.4 Decision Tree and Documentation Considerations

Figure 10 below illustrates a decision tree for PAP use for follow-on products and/or product variants. The application of an existing PAP should follow a risk-based assessment (see **Section 6.6**, Risk-based Verification Concept for Follow-on Products) to determine what, if any, supplemental validation activities are needed to establish that a method is suitable for its intended use for the new entity. The first stage in the process is to determine if a suitable PAP exists that can be used that does not require substantial change to the method conditions, reagents or materials used, or the reporting of the result. If there is not a suitable PAP, then method validation according to ICH Q2(R2) should be performed to demonstrate a methods suitability. If a PAP exists, then an assessment of the method should be conducted and documented that evaluates the appropriateness based on prior knowledge of the molecule and performance of the method. For instance, evaluation of the method range and the products acceptance criteria, matrix specificity, and understanding the CQA reportable are critical to form this assessment. Occasionally, minor changes to the PAP may be necessary for a new molecule, and this would require reduced supplemental validation to justify suitability for use. For example, if a new reagent is needed to establish specificity that does not change the overall performance or execution of a PAP, it would be important to demonstrate specificity.

If the PAP does not require any minor changes, then the PAP validation can be extended to include in scope the new molecule. Such extension of validation can be documented in a verification study (see **Section 6.5**, *PAP Verification Studies*) with reduced method performance evaluation to assure precision and accuracy characteristics with the new molecule are consistent with the original validation acceptance criteria. For example, an SE-HPLC method used to report protein high-molecular weight species can be extended to a new molecule through evaluation of method precision of a well-characterized reference material. This can be included in an AMT to establish the PAP in a receiving laboratory.

If the PAP is already established in a laboratory, and does not require changes to the method conditions, operating ranges, or reporting then a scientific justification of establishment of the PAP can be documented based on previous method transfer, and the completion of the assessment that the PAP is suitable for use for the new molecule.

Figure 10: **Decision tree for PAP (follow-on products and/or product variants)**



6.5 PAP Verification Studies

When a follow-on product is suitable for consideration for use with a PAP but cannot be fully justified as being highly similar for testing by this PAP, prospective verification studies should establish suitability (see **Figure 10** above). The method is often internally established as a PAP prior to agency approval using the first product submitted for marketing authorization. When the method is a potential PAP method, supporting data from other relevant products, demonstrating the ability of the method to be a PAP, should be compiled and analyzed as part of a single internal file. When the method is submitted within marketing applications for subsequent product(s), the data generated for prior approved product(s) may be leveraged to support scaled-down validation and verification studies for the current product. This data should demonstrate that the method can be used for multiple products or different sample matrices without modification of the procedure.

Figure 11 below illustrates the different possibilities for the initial AMQ and/or AMV studies and any follow-on post-AMQ/AMV verification studies. Whenever a PAP method is developed with the line-of-sight to use for multiple, similar products/samples, representative product samples can be included in the initial AMQ/AMV study to avoid follow-on verification studies. When conducting AMQ and/or AMV studies for follow-on products, this can occur any time during AMV stage 1, 2, or 3 with respect to the stage of the "lead-product". Any follow-on product, not initially included, introduced post-AMQ (stage 1) or post-AMV (stage 2) should be verified to extend the qualification status or validation status, respectively.

Figure 11: Establishing and Using PAP during Early-/Late-Stage Product Development and Commercial Manufacturing

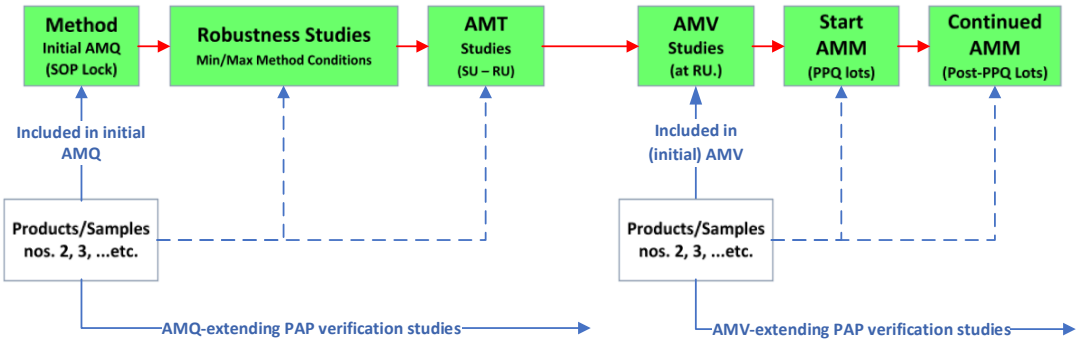


Figure 11 above and **Table 10** below capture the suggested prospective, reduced PAP verification studies to be considered. Some prospective verification studies, covering accuracy and specificity (and QL if needed), are still functionally required for validated methods for this quantitative limit test. The primary intent of the PAP verification study is to confirm a suitable accuracy and specificity method performance for the new product. In addition, a suitable QL is also confirmed by using an appropriate spiking series for the accuracy study. The suggested set of spiked samples provides all relevant method performance characteristics results, is relatively easy to prepare, and can be executed essentially in one day as shown below in **Table 10**. Other method performance characteristics, such as robustness and intermediate precision, which typically take an extensive amount of laboratory work and time, are not required to be repeated.

The prospective study results should be combined in the PAP verification report with relevant retrospective results (initial AMV study and/or AMM results), as illustrated in the example used in **Table 10**. For the marketing authorization submission, the sponsor should submit the initial AMV study report and product-specific verification results and provide a summary table containing all relevant initial AMV and PAP verification results. For the completeness of relevant information contained in the PAP verification report, the following retrospective data should be included:

- All retrospective AMV study results and acceptance criteria are not repeated in the PAP study.
- The long-term AMM assay control performance (% CV) can be used alternatively or in combination with AMV Intermediate Precision.
- Pre-AMV supporting method attributes as relevant.

Table 10: Prospective PAP Verification Study Design Example for ICH Category III Method

ICH Q2(R2) Validation Characteristic	Analyst Number	Day Number	Instrument Number	Validation Design (Spiked Analyte Concentration)
Accuracy	1	1	1	Spike A%/F% (to final %): 0.5, 1.0, 2.0, 4.0% (run each 3x)
Repeatability	N/A	N/A	N/A	From Accuracy
Specificity	1	1	1	Formulation matrix interference (3x) (also inferred from Accuracy)
Linearity	N/A	N/A	N/A	From Accuracy
Assay Range	N/A	N/A	N/A	From Accuracy
QL	N/A	N/A	N/A	From Accuracy

The prospective study results should be combined in the PAP verification report with relevant retrospective results (initial AMV study and/or AMM results). This is illustrated in **Table 11** below. Further eliminating

prospective PAP method validation characteristics should only be considered when the new DS/DP product and the product tested in the initial AMV study have very similar separation profiles and impurity levels. For the typical separation methods, which can be used and verified as PAP methods, the product-specific elution profiles and impurity levels may greatly impact the QL determination or verification.

Table 11: Combined Retrospective and Prospective PAP Validation Results for ICH Category III Method

Q2 and Q14 Validation Characteristic	Retrospective Data/Results	Prospective Data/Results	Option(s) and Consideration(s) (see also Table 10)
Accuracy	No	Yes	Number of spiked levels; Number or replicates
Precision (Repeatability)	No (see option)	Yes	Could consider using initial AMV study results if insufficient replicates (see above)
Precision (Intermediate Precision)	Yes	No	Use long-term AMM assay control data (%CV) instead or in addition to initial AMV study results
Specificity	No	Yes (see option)	Only infer lack of matrix interference from Accuracy study
Range (Reportable range, Response, and Lower range limit)	No	Yes	Number of spiked levels; Number or replicates (see Accuracy)
Robustness	Yes	No	N/A (completed previously)

6.6 Risk-based Verification Concept for Follow-on Products

Quality Risk Management (QRM), according to ICH Q9, should be considered in the design of verification studies for follow-on products. The impact to patients is the most critical factor in such risk assessment. For example, the specifications for the follow-on product may be different than those for the original product and may require improved performance capabilities of the method, so more rigorous verification studies may be required to verify suitability of the method to control the follow-on product.

Identifying any difference(s) between the structure/function and composition of original and follow-on products, is another critical consideration. The relevance of the difference(s) in the analytical method's performance should be assessed. It is further recommended to assess the potential use of the PAP for relevant products in development and to integrate them in validation or verification studies as appropriate.

6.7 Internal and External Submission Documentation Practice Options

The maintenance of appropriate documentation for PAP should allow for controlled additions of new products, verification of PAP method capabilities and managing of regulatory submissions and inspections. It is recommended that a standardized procedure is used for all relevant QC laboratories and test samples using a PAP method. The initial validation report for PAP should remain unchanged as this report is the foundational evidence for acceptability as a PAP method. Any prospective validation/verification studies performed to extend PAP use for new products/samples should be documented per sponsors documentation practice.

6.7.1 For internal documentation

The following may be considered but not limited to support the use of PAPs:

- An overarching strategic document to be developed and made available that includes relevant guidance, terminology, implementation considerations and other use conditions. This may be particularly supportive when operating/testing in multiple and global sites.
- Individual PAP-specific justification/rationale to be focused on test procedure criticality for patient safety and/or manufacturing capability.
- The validation summary contains justification for PAP use, validation acceptance criteria and PAP validation results. Each platform procedure to be maintained by using a continuously updated development and history file.
- Product- and PAP-specific verification study designs and acceptance criteria.
- Standardized PAP verification study protocols and reports.

6.7.2 For regulatory submissions

The following should be considered, but not limited to, to support the use of PAPs:

- The PAP description and the justification for its suitability for the specific PAP method. The extent of the initial/retrospective AMV studies performed to validate the method for similar product/samples relevant for submission.
- Retrospective PAP validation results used and prospective verification results, with sufficient rationale, for the relevant product/sample(s), clearly separating but linking the initial PAP validation study results to the follow-on product verification results (and/or rationale why prospective verification studies are not required).
- Standardized PAP validation summary report content and format.
- In addition, references to other submissions to ensure the assessor recognizes that this PAP has been approved.

7. Analytical Method Comparison (AMC)

7.1 AMC Introduction

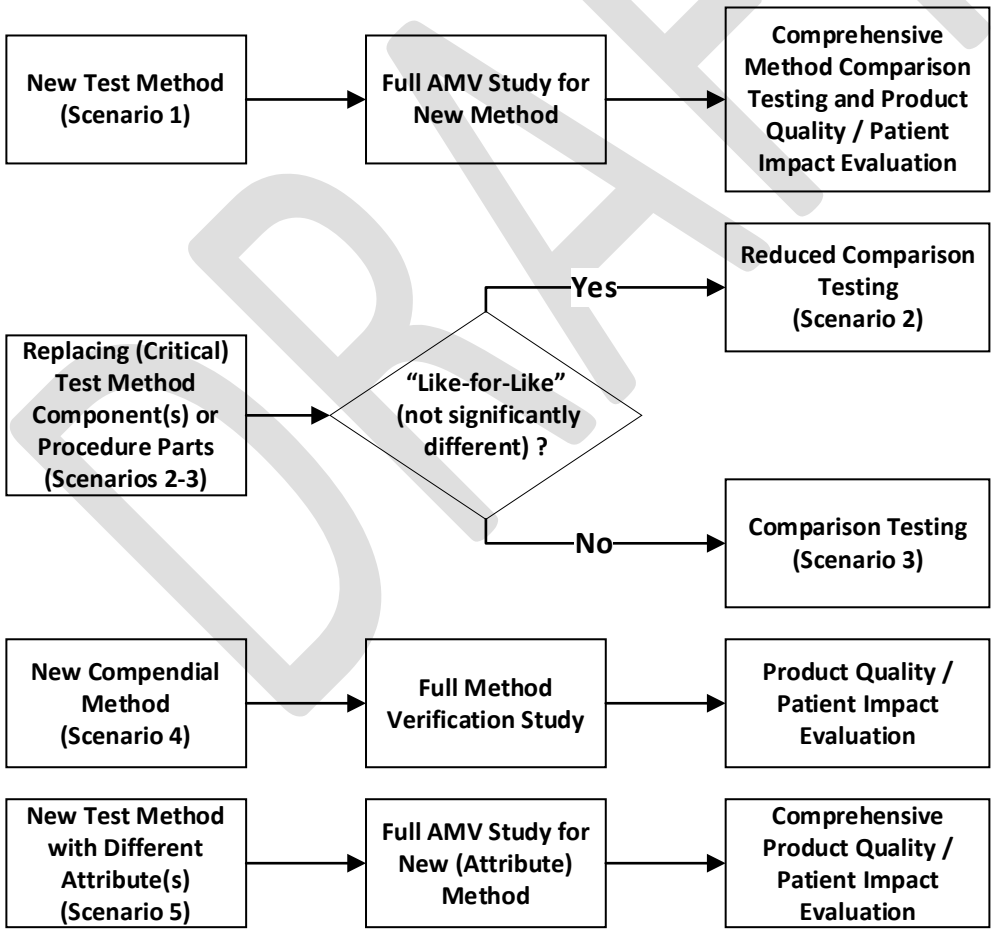
Analytical method comparison (AMC) studies apply when a currently used method is to be changed for a new method. There are multiple reasons to change analytical methods such as the replacement for an improved method or a new technology. The AMC study results are an important part of the entire change information for the new method, as regulatory approval for the use of the new or updated method is contingent upon the submitted AMC studies and relevant AMQ or AMV study results, considering the use of the method and/or stage of development.

A non-inferior, equivalent, or superior analytical method performance with respect to the impact to product safety, efficacy and/or quality should be established for the new or candidate method compared with the current analytical procedure. For late-stage/commercial products, the AMC studies could be included as part of the formal AMV protocol - for the new method - or could simply be executed under a separate protocol after the AMV has been completed. Performing a separate method comparability study after AMV completion may have the advantage that, if the AMV results reveal that a method was not optimized, the method comparison studies can be conducted with the final, representative method procedure.

7.2 Method Change Scenarios and Suggested Qualification/Validation and/or Comparative Testing

Five major change scenarios exist when changing an analytical procedure (refer to **Figure 12** below). For each type of change scenario, the potential risk(s) to the patient and/or sponsor, as well as the comparison concept(s), study design, and acceptance criteria can vary significantly. A comprehensive change control process should be followed; specifically, to perform an assessment for the criticality of change(s) to support a risk-based approach to the comparison study required.

Figure 12: Method Change Scenarios and Risk-Based Comparison Considerations



796 **Figure 12** above illustrates the major change scenarios which require some level of comparison study. Changes
797 to methods can vary from changing non-critical method elements to changing the analytical technology
798 (reporting/analysing different quality attributes). Regardless of whether changes are planned or unplanned the
799 level of risk and assurance of the level of post change control should be similar.

800

801 **Table 12** below lists possible, major change scenario descriptions with suggested comparison study
802 considerations and change examples. Each change scenario is described in more detail below, see also **Table 1**
803 for product specification-based relevant method categories (ex., quantitative limit test) and relevant comparison
804 concepts (ex., non-inferiority).

805

806 **Table 12: Major Change Comparison Scenarios**

Major Change Comparison Scenarios1		**Comparison Study Considerations**	**Case Study Example**
No.	**Description**		
1	Entirely new (replacement) method with similar or different specification limit(s) but same format for reported results.	Full Qualification/Validation plus Comparative Analysis.	Replacing a Quantitative Limit Test Method: Capillary electrophoresis (CE) to replace SDS-PAGE (section 7.6.1).
2	Partially modified method (like-for-like) without specification reassessment.	Full or partial Qualification/Validation and/or relevant Comparative Analysis.	Replacing an instrument from different vendors, different configuration, or upgrading instrument software."
3	Partially modified from Existing Method (not like-for -like) with specification reassessment.	Relevant Comparative Analysis.	Replacement of Host Cell Proteins (HCP)-specific antibody within an ELISA (enzyme-linked immunosorbent assays) method (section 7.6.3).
4	New (Replacement) Compendial method with or without specification reassessment.	Verification of relevant compendia procedure with relevant Comparative Analysis.	Rapid Sterility Test (Bact-T) to Replace USP/EP/JP Compendial Sterility Test (section 7.6.2).
5	Entirely new (Replacement) method with completely different attribute readout and specification limit(s) and format for reported results.	Full Assay Validation/Qualification Verification of method fit for purpose to assess CQA.	Replacing CEX (and IsoQuant Test Methods with (only) CEX (section 7.8.1). Replacing an In vivo Potency Assay with an In vitro G-protein based ELISA for a vaccine (section 7.8.2). Replacing three glycosylation methods with one multi-attribute method, (MAM) (section 7.8.3).

Two fundamental analytical control strategy elements should be considered for the qualification/validation studies (or extensions) and/or the comparative testing required to assure a future post-change yields appropriate product safety, efficacy, and/or quality level, described in **section 7.2**. The rigor of the change risk assessments and resulting studies increase with the change impact severity level and the product development stage(s).

Scenario 1 is the replacement of a current procedure by a new procedure. Depending on the product development stage and/or use of current method, the new method should be qualified or validated according to ICH Q2(R2) (see also PDA TR 57 for further practical considerations)[1-3]. Comparative testing should be performed to confirm/validate that relevant method performance characteristics remain either essentially unchanged (vs. risk-based maximum change acceptance criteria), thus resulting in no change in specification limits, or "method performance attributes" have changed significantly, resulting in a reassessment of specifications.

As summarized in **Table 12**, for all quantitative methods and qualitative impurity methods a significant observed mean bias should lead to a reassessment of specifications whenever those were process capability based. A significant reduction in test result variation may also require a reassessment of specifications if those were established based on process capability as actual process/product quality level may have been proportionally widened and may impact patient safety/efficacy. Qualitative method replacements specification reassessments are typically not required as specifications are not based on process capabilities. ***Scenario 1*** also applies when a compendial procedure is replaced with an in-house procedure.

The comparative study used for ***Scenario 1*** should include representative test samples representing routine test samples and potential attribute levels close to (and ideally beyond) relevant release and stability limits.

Scenarios 2 and 3 apply when the procedure is modified from the current procedure. The two scenarios are distinguished by the extent of the modification(s) and the resulting qualification/validation study requirements for the modified procedure.

In ***Scenario 2***, significant procedural change(s) require a full or partial re-qualification/validation study. Conceptually similar to scenario 1, and with consideration of method category- and/or change-depending risks, a comparative study is also required to confirm/validate that relevant method performance characteristics remain either equivalent and/or non-inferior (vs. risk-based maximum change acceptance criteria).

A comprehensive evaluation of the change(s) should include the potential impact on the accuracy and/or specificity of test results and the change in intermediate precision. The comparative study design should cover all relevant method performance characteristics and the qualification/validation study results. Whenever a significant test result shift is observed, a reassessment of specifications may be required to adjust the desired control level as needed.

In ***Scenario 3***, critical test system elements (ex., critical reagents, signal output detection device) are changed without significant procedural changes, thus not requiring additional qualification/validation studies. A comparative study may be required to confirm/validate that relevant method performance characteristics remain equivalent (vs. risk-based maximum change acceptance criteria). For those significant change-impact outcomes resulting in a test result shift, a reassessment of specifications may be required to adjust the desired control level as needed.

Scenario 4 describes the process and considerations when changing from a currently used pharmacopeial procedure to a different pharmacopeial procedure. The sponsor should consider the implications, if relevant, changing within a local pharmacopoeia versus between pharmacopoeia (ex., USP => EP). Comparative testing should be considered if the change could potentially impact any established specifications and/or product quality levels. Those cases can be covered by the strategies proposed for categories 1, 2 and 3 as illustrated above. The same risk assessment principles will drive the strategy for validation/verification of the candidate analytical procedure and comparison of the current and candidate procedures.

As pharmacopeial procedures are typically only verified with limited studies and method history/knowledge may be limited, the potential lack of detailed knowledge about the method performance characteristics of pharmacopeial procedures may be extensive compared with other validated methods. This potential gap should be considered for comparison studies.

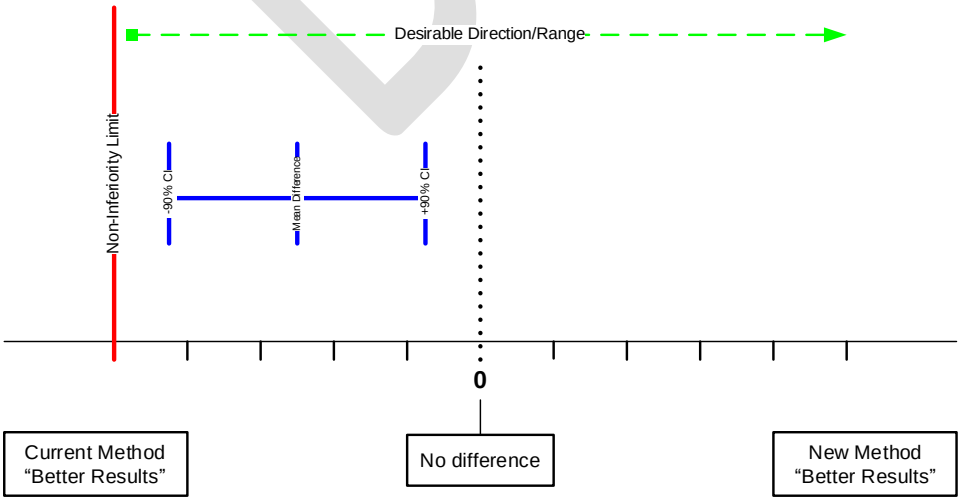
Scenario 5 considers the pertinent factors to be established when replacing a validated method with a novel/alternate method which may result in a change in specification limits or establishment of novel specifications. More than one alternate method may be required to assess the CQA (ex. identity testing via enzymatic activity combined with molecular STR genotyping). The alternate method must be fully validated in accordance with ICH Q2(R2) and demonstrated to be fit-for-purpose.

Comparative testing may not be appropriate; however, the current and proposed analytical procedures should be performed concurrently to acquire detailed knowledge to appropriately establish acceptance criteria so that the assessment of product quality is not negatively impacted. The intended purpose of the analytical procedure (characterisation, in-process control, release and/or stability) will govern the level of control required; therefore, samples must be representative of routine test samples (in-process, release, stability samples). Alternative approaches and risk assessment principles are discussed in greater detail in Section 7.7, *Comparison of New Technologies to Existing Technologies*.

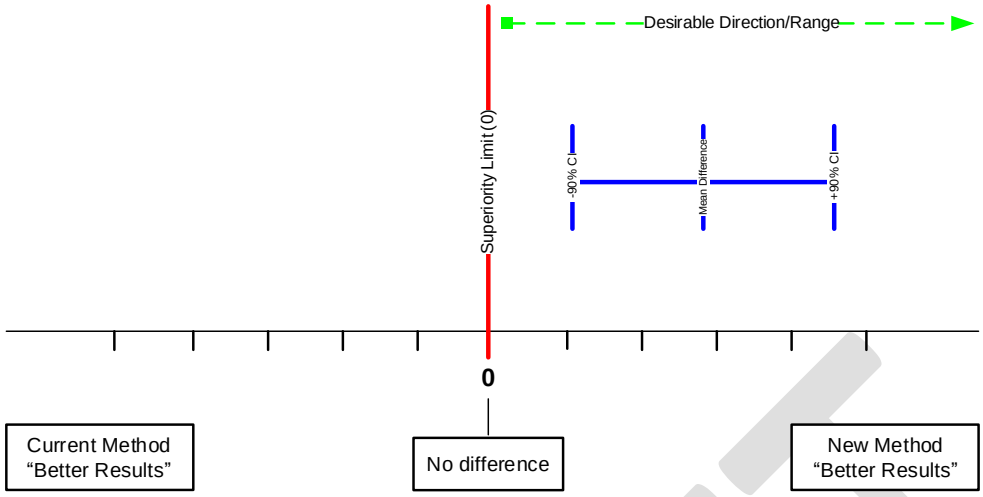
7.3 AMC Models and Rationale

AMC is established to be required when the change impact assessment and the resulting conclusion that the identified change(s) and change outcome(s) is/are not like-for-like changes. A like-for-like change would typically be sufficiently similar with a high degree of confidence and not anticipated to significantly impact method performance attributes (ex, test result accuracy/bias, reliability). The appropriate AMC model is dictated by the method category (ICH Q2(R2) I-IV) and method performance characteristics compared. The three AMC models are illustrated in **Figure 13, Figure 14,** and **Figure 15** below. Conceptually, qualitative results/data comparison typically requires the use of non-inferiority and/or superiority models which typically compare pass/fail probabilities (by using a 90% CI TOST for the mean difference in probability [in %, ratios, or counts]). The use of a non-inferiority model versus superiority should be justified. In both models, the 90% CI for the mean difference should fall to the right of the pre-set (desirable direction) of the non-inferiority limit or superiority limit. Although more commonly used as two-sided CI, a one-sided CI could also be used whenever a one-sided specification exists. The equivalence model is applied for quantitative direct comparison testing where the 90% CI for mean difference by a TOST test is typically used.

Figure 13: Non-Inferiority AMC Model



884 **Figure 14: Superiority AMC Model**

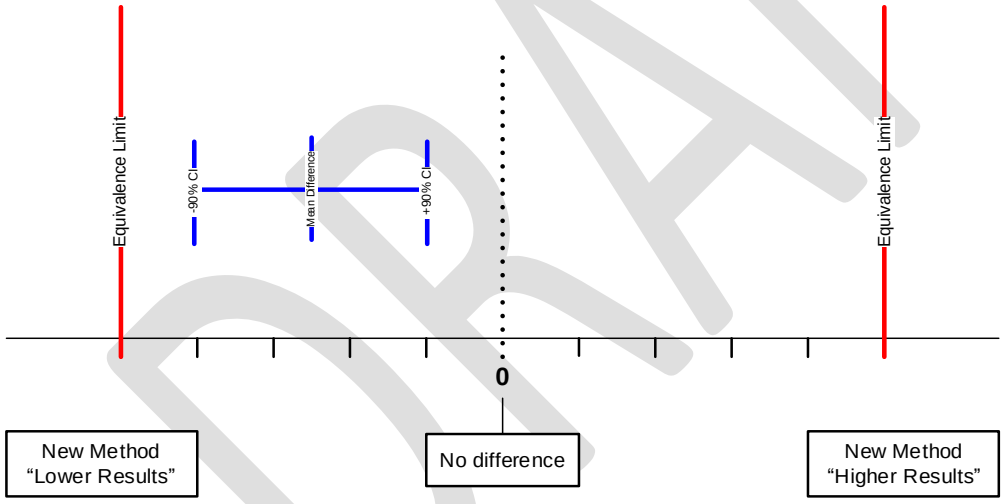


885

886

887

888 **Figure 15: Equivalence AMC Model**



889

890

891 ## 7.4 Risk Based Study Design and Acceptance Criteria

892 For most AMC studies, a direct comparison testing will be the best option to demonstrate that the new method
893 will be equivalent, non-inferior, or superior for each method category. The direct comparison approaches are
894 first conceptually described in sections 7.4.1-3. Risk-based study design options (sections 7.4.4-5) and
895 acceptance criteria (section 7.5) for the comparison of required method performance characteristics are given
896 below. For those method comparisons when correlations between measurement values are not possible because
897 of the change in the analytical technology and reported results (ex., different units), alternative approaches for
898 comparative analysis are necessary to establish continuity, as relevant, between the methods.

7.4.1 Direct Comparison Testing

The four general ICH Q2(R2) method categories can be grouped into two greater categories, qualitative and quantitative methods. A qualitative method provides qualitative results (pass/fail, yes/no, or results reported simply as "less than" the action or specification level), whereas a quantitative method is expected to provide results accurately and reliably, reported in the same units as the specifications.

Qualitative methods are expected to provide a current representative result (ex., "pass") at a high probability. For qualitative limit tests, a low Detection Limit (DL) is desirable as it increases the likelihood for observing positive results at low analyte concentrations – refer to **Table 1** in section 1, *Introduction,* that summarizes the method performance characteristics to be compared.

7.4.2 Qualitative Tests

All qualitative tests should contain, at minimum, a comparison of positive-to-fail ratios of spiked (low) analyte concentrations. This will ensure a comparable level of specificity of both methods. For the DL, both hit-to-miss ratios can be compared at very low analyte concentrations using probability statistics. When comparing qualitative data, noninferiority or superiority models should be used, and three possible outcomes are illustrated below [1].

- **Inferiority**. A particular performance characteristic compared provides significantly inferior results for the new method, therefore failing to demonstrate AMC.
- **Noninferiority**. The new method performs at a comparable level. The new method could be superior, equivalent, or insignificantly inferior. All three outcomes are acceptable outcomes to demonstrate noninferiority.
- **Superiority**. The new method is superior. When testing for superiority, only this outcome is acceptable.

7.4.3 Quantitative Tests

For all quantitative methods, the method performance characteristics accuracy and precision (intermediate precision) should be compared. Similar to the comparison of qualitative data, the comparison of intermediate precision for quantitative tests could have three acceptable outcomes (noninferiority, equivalence, or superiority). Depending on the prespecified allowable difference set and justified, a significant shift in results may require a change in the release specifications or other possible adjustments before the new method can be used for release testing. The demonstration of comparable accuracy (or result matching) of a method will therefore require the use of an equivalence model [1].

When comparing quantitative data for accuracy, two possible outcomes are illustrated below:

- No equivalence. The observed statistical difference (e.g., 90% two-sided confidence intervals) is not within the predefined acceptance criteria. The new method may be acceptable if specifications changes are justifiable or other adjustments can be made.
- Equivalence. The statistical difference between both methods is completely enclosed in the acceptance criteria (i.e., the new method performs at a comparable level).

7.4.4 Study Designs

The success of the AMC study relies on the design and acceptance criteria of the study. The following general considerations for AMC studies should be based on change and impact using a risk-based approach.

- Inclusion of critical method variation factors typically seen in intermediate precision studies to simulate and represent long-term variation.

- The number of representative products or relevant test material lots/samples, and any replicates of those used given the desired statistical confidence in concluding passing or failing AMC studies.

- The setting and justification acceptance criteria for relevant method performance characteristics compared.

7.4.5 Considerations for Phase / Stage Appropriate Comparative Studies- Post Approval Change Demonstrating AMC in Post-Validation Studies

Pre- and post-validation AMC studies are conceptually similar. The comparison model(s) and method performance characteristics evaluated are essentially the same for any product development stage and a commercial product. However, the comparative testing sample size, study design, statistical significance (if relevant), and acceptance criteria used should be commensurate with the risk(s) to patients and sponsor.

Comparative testing acceptance criteria and/or protocol may not be required for early-stage method changes but any necessary test data selection as study outcome should be justified and included in the change control documentation. Any comparative testing conducted simultaneously with routine material and/or product testing should be captured with appropriate documentation as required in the relevant GMP quality system/procedures.

7.5 Risk-based approach for acceptance criteria in AMC Studies

The success of a comparison study is pre-established with suitable acceptance criteria as justified in the comparative testing protocol. As described in Section 1 (*Introduction*), the acceptance criteria should be risk-based and set to ensure a continuous product safety, efficacy, and/or quality.

7.5.1 Approaches to consider for AMC Studies

Multiple approaches for establishing acceptance criteria can potentially be justified and applied to confirm/demonstrate a continuous level(s) of required method performance characteristics (see also Section 1, *Introduction*, **Table 1**). For example, using an "inside-out" approach in early-stage product development, where acceptance criteria are derived from current method performance, the established acceptance criteria should still ensure risk-based continuous control following this method change, if comparative study results fall close to failing the comparative testing acceptance criteria. For late-stage/commercial product development, an "outside-in" approach, where acceptance criteria are derived from the specification and relevant, current process capability, could be used where test results of the post-change method are not expected to be significantly different from the current method, thus not requiring a potential specification reassessment.

7.5.2 Risk factors to consider for AMC Studies

Several risk factors are described in **Table 13** below. The criticality of the current and future method is usually captured in the product control strategy and is reflected by the (potential) critical quality attribute level/score, which in turn is based upon the potential severity to harm patients [9 - 12]. Similar to the established product attribute severity assessment level, the associated risk factor(s) for patient safety/efficacy will be a fixed contributor as manifested in relevant specification(s). As specifications are typically risk- and process-capability-based in late-stage development and often tighter going into commercial product manufacturing, acceptance criteria should be justified in reference to relevant specifications.

Aligned with the manufacturer's and/or sponsor's risk evaluation practice, the test method change evaluation should ideally follow this practice and could be performed qualitatively or quantitatively.

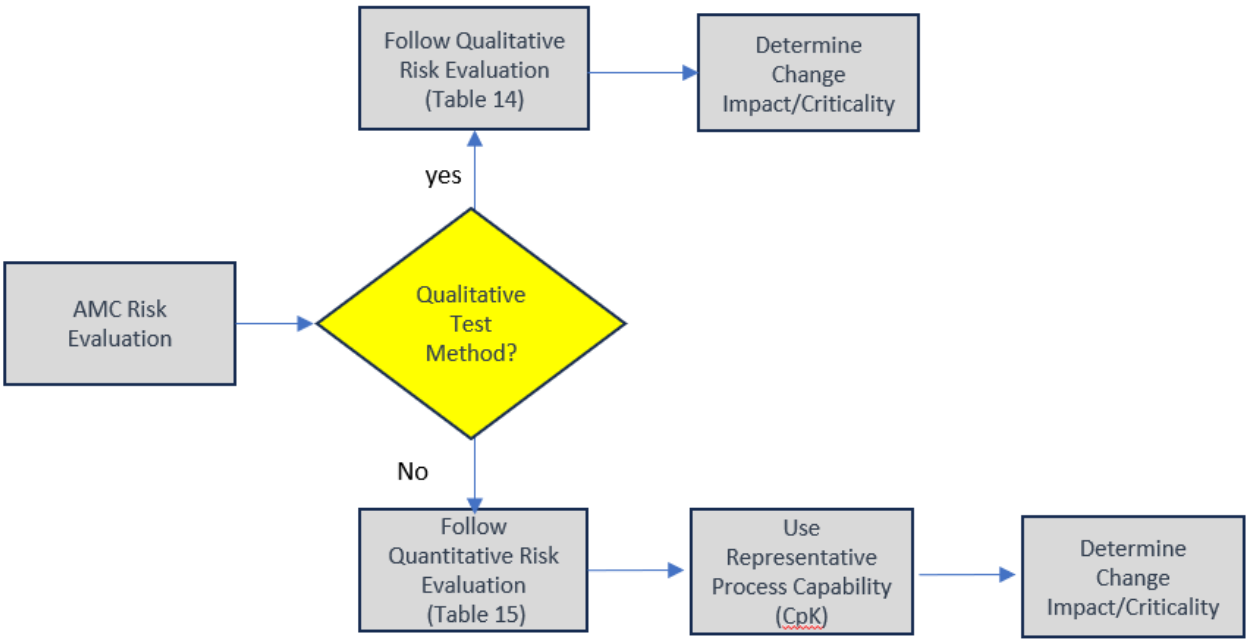
980 **Table 13: Risk Evaluation Factors and Examples for AMC Studies**

Risk Evaluation Categories	Risk Variants	Examples	Expected Potential Risk/Impact
Attribute(s) Type - Intended use	Identity, Safety, Purity, Quality, Potency, and Stability	a. Safety test: Sterility test using new rapid microbial method. b. Quality test: Excipient concentration at final production stage. c. Purity/Stability test: Degradation products during storage.	a. Potential risk to patients and firm is very high if sterility test provides false negative results. b. Potential risk to patients is relatively low if the quality test provides inaccurate results as excipient is quantitatively added during production. c. Potential risk to patients is high if stability test is incapable to measure all degradation products
Attribute(s) Criticality - Intended use	CQA p(CQA) (CPP, KPP, NKPP)	a. Safety test: Sterility test using new rapid microbial method. b. Quality test: Excipient concentration at final production stage. c. Purity/Stability test: Degradation products during storage.	a. Potential risk to patients and firm is very high if sterility test provides false negative results. b. Potential risk to patients is relatively low if the quality test provides inaccurate results as excipient is quantitatively added during production. c. Potential risk to patients is high if stability test is incapable to measure all degradation products
Closeness to Finished Product	DP, DS, API, Intermediates, Starting/Raw Material	Purity/Safety test: An HPSEC method is used for quantitation of protein aggregate levels. A second electrophoresis method provides similar results for aggregate levels.	If second method routinely supports the results of the primary method, the risk to patients may be lower if the primary method provides inaccurate results.
Intended Use	Routine Test, PPQ, Process/Product Characterization	Purity Test: Fermentation impurities are measured before purification and after purification.	Early-stage inaccurate impurity results from less reliable method are lower risk to patients if late-stage testing provides more accurate results.

Product Lifecycle Commercial, PPQ, Pivotal, Early-Stage	Commercial, PPQ, Pivotal, Early-Stage	Potency Test: Potency testing in drug substance samples.	The potency results of in-process samples collected may be affected by the actual sampling process and/or hold times before testing. This risk may therefore be higher to the firm as test results may not be representative of drug substance batch prior to filling.
Replacement Method "Status"	PAP (MAA)- "approved" "New" analytical technology/method	Purity test: APT HPSEC method is used to test in-process samples.	Current QC experience with this method performance should lower the risk to patient and/or firm if the effect of different sample types is insignificant.

Figure 16 below describes several risk evaluation factors, and the risk assessment process is presented as a qualitative assessment to avoid over- or under-representation of patient and/or sponsor risks for outcomes of the comparative testing. For example, replacing a critical safety test such as the USP/EP compendial sterility test with a new-generation rapid sterility test is an extremely high patient risk and clearly outweighs all other risk evaluation factors. The resulting study design, sample size, and statistical confidence level required should therefore be commensurate with this high level of patient risk, regardless of the product development stage. The study acceptance criteria should ideally be developed with respect to the specification(s) associated with the current method and the risk evaluation categories in **Table 13** (and **Figure 16**) below should remain consistent. Each risk component must be evaluated, and the process documented. Similarly, replacing an identity method can also be a high-risk change if the production facility is used for multiple similar products or products with a high degree of cross-reactivity. As an example, for a lower risk to patients and/or sponsor, the replacement of an early-stage product characterization test which does not support a (p)CQA, less rigorous comparative testing may be justified to support this change. Refer to **Table 14** and **Table 15** (below), for additional details on AMC studies for qualitative and quantitative test methods.

1006 **Figure 16: High-Level Decision Process for AMC Studies**



1007

Table 14: Examples for Qualitative Method Comparison Studies

Method Type (ICH Cat. no.)	Examples	AMC					Rationale/ Comments
		Risk/ Impact Level (using Table 11)	Performance Characteristics	Method Comparison Study Design	Statistical Methodology	Acceptance Criteria	
Identification - Present/Absent (cat. I)	Safety Test, replace existing 5-day Sterility Test with faster 2-day Sterility Test	(Very) High	Specificity	Compare detection probabilities for selected spiked microbial organisms at low levels (above/below 5-day Sterility Test DL)	Non-Inferiority of 2-day Test for Pass/Fail Ratio(s) or Probabilities	Justified (-) margin for individual or pooled one-sided 90% CI mean difference between 5-day and 2-day test results.	Non-inferiority comparison model justified as more frequent testing will be scheduled, and test results obtained faster. Pooling of test results groups for selected microorganisms justified if pass/fail probabilities are similar among spiked organisms.
Impurity – Limit (cat. III)	Replace existing Host Cell DNA (q)PCR test used to test and release DS with an advanced automated (q)PCR and more sensitive test method. Product to be administrated to patients during pivotal clinical phase.	High	Specificity	Compare detection probabilities for selected low spiked microbial spike concentrations and/or sample dilution levels at low levels (above/below 5-day Sterility Test DL)	Non-Inferiority of automated test for pass/fail low spike concentration levels bracketing existing DL.	Justified (-) margin for individual or pooled one-sided 90% CI mean difference between pass/fail probabilities between both methods.	Non-inferiority comparison model justified as new test method is automated, more reliable, and DL expected to be lower (superior).
			Detection Limit	Compare established DL of automated test to existing test method using same methodology for establishing DL.	Superiority of automated test (lower DL) established as lower-than current method's spike levels can be consistently detected using a one-sided 90% CI t-test to assure at least 95% confidence.	Established DL for automated method should be significantly lower statistically (at 95% confidence level) than existing method's DL.	Automated method is to be implemented as an improved and more sensitive test method. Superiority should be established given the expected lower DL for Host Cell DNA.

Table 15: Examples for Quantitative Method Comparison Studies

Method Type (ICH Cat. no.)	Example (using CpK-values) Consider Release and Stability (if relevant)	AMC					
		Risk/Impact Level (from above)	Performance Characteristics	Method Comparison Study Design	Statistical Methodology	Acceptance Criteria	Rationale/Comments
Impurity - Quantitative (cat. II) Replacing a Quantitative Limit Test Method: Capillary electrophoresis (CE) to replace SDS-PAGE (see also section 7.6.1)	CpK = 1.7(+) for Release and/or Stability	High R/S Test; Commercial. CQA test Etc.	Accuracy	Test n=30 representative DP lots (pooled or alternated, if needed) to compare mean release test results.	TOST for matched-paired t-test	TOST Equivalence Margin to +/- 0.50% (reported % results)	DP release CpK of 1.7 was determined for the one-sided, lower specification limit (LSL) since the product specification(s) is/are one-sided: The current CpK of 1.7 was used to establish and justify the DP release equivalence acceptance criteria (0.50%).
			Intermediate Precision	Same data as above, plus AMV data	TOST (see Accuracy) RSD	TOST Equivalence Margin to +/- 0.50% (reported % results) From AMV study: RSD (%CV) of CE method to be NMT + 1.0% of SDS Page Method	The intermediate precision of both methods impacts the width of the 90% two-sided CI for the mean difference between the two method results. This equivalence test model therefore integrates the total analytical error (TAE) of each method into the outcome of the analytical bias (mean difference) study results. RSDs (%CV) for Intermediate Precision of the from the AMV study results were compared from the AMV study results and/or assay control long-term variation. Intermediate precision (% CV) was improved with the CE method (lower % CV).
			Specificity	AMV data (retrospectively)	% Recovery, other options	From AMV study: CE test method specificity acceptance criteria	No significant matrix interference to be observed.

			Stability-Indicating	Compare n=x degradation rates or degradation level(s) at t=X	TOST for pooled degradation rates (if pooling can be justified). If data cannot be pooled, compare mean and worse-case degradation rates.	Maximum acceptable difference marginfor the pooled (n=3) DP stability regression slopes one-sided 90% prediction interval for CE method to be NMT (-)0.40%.	The stability (end of shelf-life) equivalence margin of 0.40% was established and justified similar to Accuracy. The degradation regression line slopes from n=6 representative DP stability lots were pooled as none of the regression line slopes were statistically different p (>/=0.05). The lower 90% one-sided prediction interval of the average (n=6) slope for the EOSL (4 years) of 0.60% was subsequently used to establish the maximum acceptable limit (0.40%) for the allowable difference in the pooled (n=3) DP stability regression slope one-sided 90% prediction interval for the new method.
Assay (Content, Potency) (cat. IV) Relative Potency Binding-Assay to be replaced with Cell-Based Relative Potency Assay	CpK = 1.4(+) for Release and 1.2(+) for Stability Current specification (60-140%) is same for Release and Stability	(Very) High R/S Relative Potency Test. Pre-Pivotal CTM (change implemented via pivotal IND/IMPD amendment) CQA test Etc.	Accuracy	Side-by-side test (concurrently if possible) all a representative (total n=26) pre-pivotal CTM release/stability samples to establish MoA-based new specification(s) for Release and Stability.	TOST for matched-paired t-test	Equivalence margins for the two one-sided 90% CI for mean difference to be 79-121%.	Established to account for new relative potency assay non-inferiority limit (13% RSD) for Intermediate Precision. Non-inferiority (+3% RSD vs current assay) is justified as new assay is MoA-but based on relative potency (vs MoA-based, qualified reference standard). Limit(s) are set to not exceed current specification(s) as product has relatively narrow therapeutic window. Failing to pass equivalence test may require a re-evaluation of potential wider specifications for pivotal CTM given the desired therapeutic dose window.

			Intermediate Precision	Same data as above, plus AMV data (retrospectively)	TOST for Equivalence and RSD based on Non-Inferiority	Equivalence margins for two one-sided 90% CI to fall inside 81-119%. % RSD ≤ 13% (AMV)	See above under Accuracy Non-inferiority (+3% RSD vs current assay) is justified as new assay is MoA-but based on relative potency (vs MoA-based, qualified reference standard). Limit(s) are set to not exceed current specification(s) as product has relatively narrow therapeutic window.
			Specificity	AMV data (retrospectively)	% Recovery, other options	No significant matrix and/or process impurity interference for validated assay range (50 – 150% relative potency).	The specificity or lack of matrix and/or potential process impurity interference was determined to be insignificant and considered acceptable for both methods
			Stability-Indicating	Compare n=x degradation rates or degradation level(s) at t=X	TOST for pooled degradation rates (if pooling can be justified and sufficient stability lots/data is available). If data cannot be pooled, compare mean and worse-case degradation rates.	See Accuracy Degradation rate(s) or levels for desired shelf-life not to exceed Non-Inferiority limit of 5% (vs current method)	See Accuracy Failing to pass equivalence test and /or Non-Inferiority limit may require a re-evaluation of potential wider specifications for pivotal CTM given the desired therapeutic dose window. Release and Shelf-life specifications may need to be re-evaluated.

In addition to the management of the risk/impact severity represented by the change, the control of the consistency of quantitative results in relation to the product specifications must also be considered. This can be done by measuring the process capability index or process capability ratio, the ability of a process to produce output within specification limits [2]. The concept of process capability only holds meaning for processes that are in a state of statistical control. Process capability indices measure how much variation a process experiences relative to its specification limits and allows different processes to be compared with respect to how well an organization controls them.

The acceptance criteria set to control a change might also consider the maintenance or improvement of the Cpk value and the control of the consistency of results in relation to the product specifications. A low Cpk before the change may result in more restrictive criteria for the comparability study or will require a different study outcome interpretation. Refer to **Table 16**, for suggested acceptance criteria when informed by process capability.

Table 16: Examples of AMC Acceptance Criteria for Quantitative Methods Using an Inside-Out Approach

CpK Value	Suggested Acceptance Criteria based on Current Method
Above 2.00	TBD (widest possible ACs); Example: TOST equivalence margin(s) to be an equivalent CpK reduction of 0.67 (CpK post-AMC is still above 1.33)
1.50 – 2.00	TBD (relatively wide ACs)
1.00 – 1.50	TBD (typical late-stage/commercial ACs)
Below 1.00	TBD (tightest possible ACs) Example: TOST equivalence margin(s) to be an equivalent CpK reduction of < 0.33 (CpK post-AMC is only < 0.67)

7.5.3 Points to consider when selecting test samples for evaluation of analytical method comparability.

Samples used in the AMC studies should represent the product that is routinely tested and consider the criticality of the attribute and intended purpose of the method. General considerations for test sample selection and AMC study execution are given below.

- Samples from different batches and relevant production steps should be selected to represent manufacturing variability.
- Samples of different age, and/or forced degraded samples should be selected to demonstrate the new method is stability indicating and support the comparison of degradation rates.
- Samples representing relevant product variants and/or historical ranges of critical product-related impurities should be included.
- The samples should cover the specification range if impact to product specifications is anticipated or if the bias between methods is variable across the range.
- Test sample stability over the duration of the study should be considered.
- The level of effort for the study design and formality of documentation should be commensurate with the risk to the product specification, hence the criticality of the analytical method (see **Table 13**).

Whenever possible, side-by-side testing should be performed for AMC studies. If routine testing is relatively frequent and the test method change(s) and/or replacement can be implemented with a prospective change plan/protocol, routine samples could be tested concurrently with the new method. This would avoid generating any potential re-testing results with the current test method. Whenever concurrent testing cannot be done or may not be sufficient, the following AMC testing options should be considered.

- **Re-testing of DS/DP** (released GMP lots). This should be avoided for AMC (and AMT) studies, except for, characterization testing, and DS/DP stressed stability condition(s). Test sample considerations, using alternative (representative) material is given below. Additional acceptable options could also be considered.

- **Expired DS/DP:** Available samples can be conveniently used for side-by-side testing without further preparation. No DS/DP lots in distribution (or filings) will be implicated if OOT/OOS results are obtained. When OOS/OOT results are observed in AMC testing, this could be challenging to evaluate since a reference value is less reliable (only estimated/extrapolated).

- **Reference standard and/or assay control(s):** Available samples can be conveniently used for side-by-side testing without further preparation. Reliable reference values exist so that comparison acceptance criteria can be set with relatively high degree of confidence.

- **Stressing/Spiking of DS/DP:** OOS results do not really apply and may be intentional for some samples to simulate testing and samples close to and/or above the OOS level(s). These manipulated samples are recommended to compare test methods at the point of failure and are typically used for AMQ/AMV studies to bracket the product specifications. However, these AMC samples are by themselves not sufficient to evaluate test method performance and/or bias for routine DS/DP results.

- **Blending of DS/DP lot samples:** Available samples can be relatively easily prepared and used for side-by-side testing. No DS/DP lots in distribution (or filings) will be implicated if OOT/OOS results are obtained. When OOS/OOT results are observed in AMC testing, this could be challenging to evaluate since no direct reference values (expected results) are readily available due to blending of multiple lots.

An appropriate sample size should be determined using the risk-based approaches, as outlined in TR57 and USP's proposed general chapter PF 35(2) <1033> *Biological Assay Validation* [2,7]. Sample size should consider method complexity, variability, and type (quantitative versus qualitative), attribute criticality, prior knowledge, and practicality. The sample size must support the power comparison/confidence needed to demonstrate equivalence or other AMC study conditions.

If new process-related impurities or product-related substances are detected with the replacement or modified method, samples from historical batches should be analyzed to demonstrate that the newly detected impurities or substances are a result of an increase in the sensitivity or selectivity of the replacement or modified method, and not a result of a change to process-related impurities or product-related substances.

7.6 AMC Case Studies

7.6.1 Case study 1 - Replacing a Quantitative Limit Test Method: Capillary electrophoresis (CE) to replace SDS-PAGE

Case study 1 represents scenario 1 described in **Table 12**. A capillary electrophoresis (CE) method was validated and compared to an approved SDS-PAGE electrophoretic method used to control product-related impurities in a commercial monoclonal antibody (mAb) drug product, with the intent to replace it. Using **Tables 15 and 16** for the risk evaluation and manufacturability, respectively, the outcome of the AMC is briefly summarized in **Table 17**, below. The risk evaluation requires a maximum level of confidence that the new method performs equal or better, thus using a high number of comparison test samples.

Table 17: **Risk Categories and Evaluation for Case Study 1 – Replace SDS-PAGE with CE**

Risk Evaluation Categories	Risk Variants	Case Study Example/Outcome
Attribute(s) Type	Purity and Stability	Controlling protein impurities and degradation products during manufacturing and storage.
Attribute(s) Criticality	CQA	Controlling a CQA with potentially severe patient safety.
Closeness to Finished Product	DP	Final control point for product release and shelf-life
Intended Use	Routine Test	Primary purity release and stability test to assure a safe product
Product Lifecycle	Commercial Production	Large patient population and highest possible patient impact. Highest regulatory expectations for approval.
Replacement Method "Status"	Current PAP replaced with new method (potential new PAP)	Extensive method performance knowledge with current APP method. Far less method performance knowledge/certainty with new method.

The method performance characteristics (quantitative limit test) are directly compared in AMC studies or indirectly from AMV study results following the suggestions in Section 1, *Introduction*, **Table 1**.

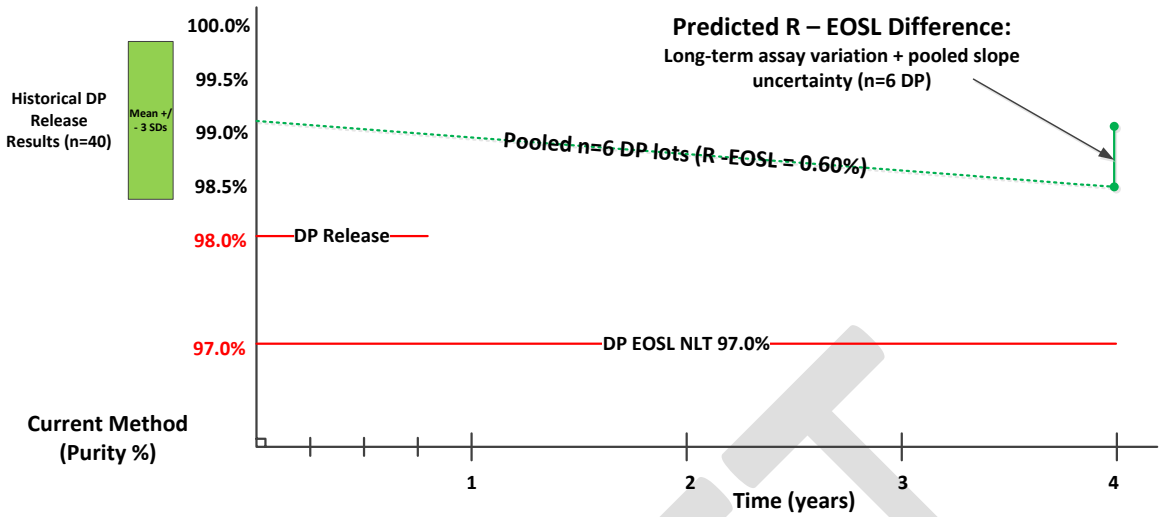
- **Accuracy** ("matching") is directly compared through product sample testing (release and stability).

- **(Intermediate) precision** is compared historically and concurrently by comparing the long-term assay control variation of a drug product-representative material. As expected, the intermediate precision (% CV) was improved with the CE method (lower % CV). The intermediate precision of both methods impacts the width of the 90% two-sided CI for the mean difference between the two method results. This equivalence test model therefore integrates the total analytical error (TAE) of each method into the outcome of the analytical bias (mean difference) study results.

- **Quantitation limit(s), Assay Range(s), Specificity,** and **Stability-Indicating** were compared indirectly by a retrospective comparison of the method performance characteristics from the AMV studies. As expected, the QL of the CE method was slightly lower while the upper QL was similar between methods, and thus, the Assay Range was suitable for the intended use. The specificity or lack of matrix and/or potential process impurity interference was determined to be insignificant and considered acceptable for both methods.

Figure 17 below graphically summarizes representative historical DP lot release results (n=30) and DP stability results (n=6 lots). In this figure, the current manufacturing capability (mean +/- 3 SDs) has been significantly better than the desirable minimum capability as typically expressed by a process capability (CpK) value of 1.00. Similarly, the predicted stability changes/degradation measured by the current method are significantly smaller than the worst-case maximum DP release-to-stability (EOSL) differential ($DP_R - DP_S = 1.0\%$).

Figure 17: Evaluating Process Capability (CpK) for DP Release and Stability – Case Study 1 [13]



A suitable statistical methodology (TOST) was selected given the risk evaluation outcome and the quantitative limit method category considerations. Using the commercial process capability of the current method, CpK = min x [(USL – mean/3σ), (mean – LSL/3σ) , where: σ is the standard deviation; upper specification limit = USL; lower specification limit = LSL, a DP release CpK was determined for the one-sided, lower specification limit (LSL) since the product specification(s) is/are one-sided: The current CpK of 1.7 was used to establish and justify the DP release equivalence acceptance criteria (0.50%). The stability (end of shelf-life) equivalence margin of 0.40% was established and similarly justified.

$CpK_{(LSL)}$ = (99.2% - 98.0%) / (3 x 0.23%) = 1.7

The stability (end of shelf-life) equivalence margin of 0.40% was established and similarly justified. The degradation regression line slopes from n=6 representative DP stability lots were pooled as none of the regression line slopes were statistically different p (>/=0.05). The lower 90% one-sided prediction interval of the average (n=6) slope for the EOSL (4 years) of 0.60% was subsequently used to establish the maximum acceptable limit (0.40%) for the allowable difference in the pooled (n=3) DP stability regression slope one-sided 90% prediction interval for the new method.

The long-term assay variation has been included practically into the stability data of the current method over extended time period (> 4 years) and was simulated by testing the new method with varying critical method variation factors (instruments, days, etc.). In addition, accelerated and/or stressed DP stability samples exceeding the EOSL specification limit of 97.0% were compared by both methods to ensure that the unchanged DP specification(s) will be suitable to maintain a similar product purity level at and beyond the specification level(s).

Any prospective retesting of DP lots used in commercial product distribution for the purpose of method comparison studies should be carefully considered. Other options may need to be considered, such as concurrent testing when the new method is also used in testing DP lots routinely (in parallel), and/or, using representative historical release/stability data for the current method while prospectively testing only with the new method. Deliberate pooling of DP lots can be used to avoid different duplicate results, see also section **7.5.1, *Approaches to consider for AMC Studies*.**

Both methods are to be run simultaneously (side-by-side) for each of a total of n=30 reported results. Results are then compared by two-sided matched-paired t-test statistics with pre-specified equivalence limits of plus/minus 0.50% (% = reported percent and not relative percent).

Figure 18: Establishing and Justifying Release and Stability Equivalence Acceptance Criteria – Case Study 1

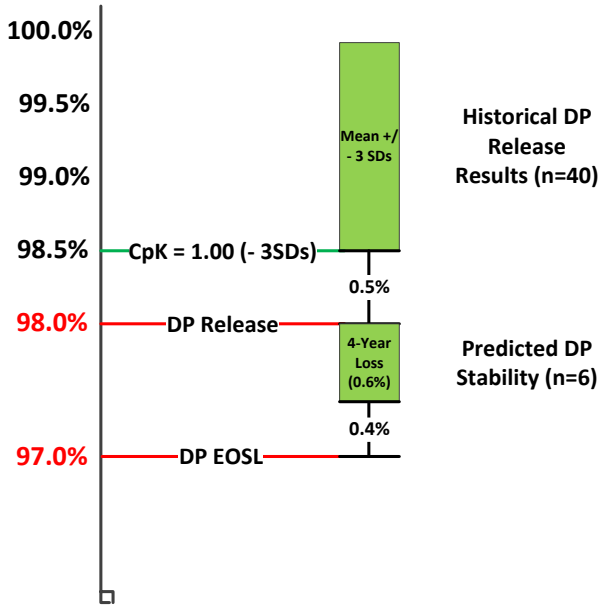


Figure 18 above illustrates how release and stability equivalence acceptance criteria were set and justified. A limit of plus/minus 0.50% was chosen for the equivalence category between both impurity levels from the analysis of historical data with respect to the current specifications (for SDS-PAGE), as illustrated in **Figure 19** below. The historical DP release data (last n=30) shows a current manufacturing capability significantly greater than a CpK of 1.0 (> 3 SDs). Using the DP release and stability specifications with the intent to leave those unchanged, this maximum difference of 0.50% allowed will ensure that a desired minimum DP manufacturing capability remains acceptable (CpK ≥ 1.0), as illustrated in **Figure 19**.

Although, practically, a one-sided equivalence test could have been applied due to the anticipated one-sided bias towards lower purity results by the new method, this atypical option was not selected as the actual lower acceptance criteria and outcome(s) would have been identical.

The comparison of stability test results was done separately. Similar to the release test result comparison, a delta of +/- 0.40% was set for the equivalence margins between both impurity levels for stability testing from the analysis of historical stability data (n=6 DP lots), see also **Figure 20** below.

The option to pool release and stability data when comparing both methods was not selected as a time/temperature-dependent diverging of relevant degradation/stability data was expected based on increased sensitivity (lower QL) of the new method. DP release and stability data were therefore compared separately with separate acceptance criteria, established from the corresponding specification limits and based on manufacturing capability and predicted stability changes over the approved DP shelf-life. The average DP stability changes was 0.6% when using the current method with a specification limit differential of 1.0%. As a diverging, proportionally higher degradation/change rate was anticipated for the new method at the end of shelf-life, the maximum difference for the new method (vs. old method) was therefore limited to 0.4% at 4 years shelf-life.

Figure 19: AMC Equivalence Testing Results for DP Release

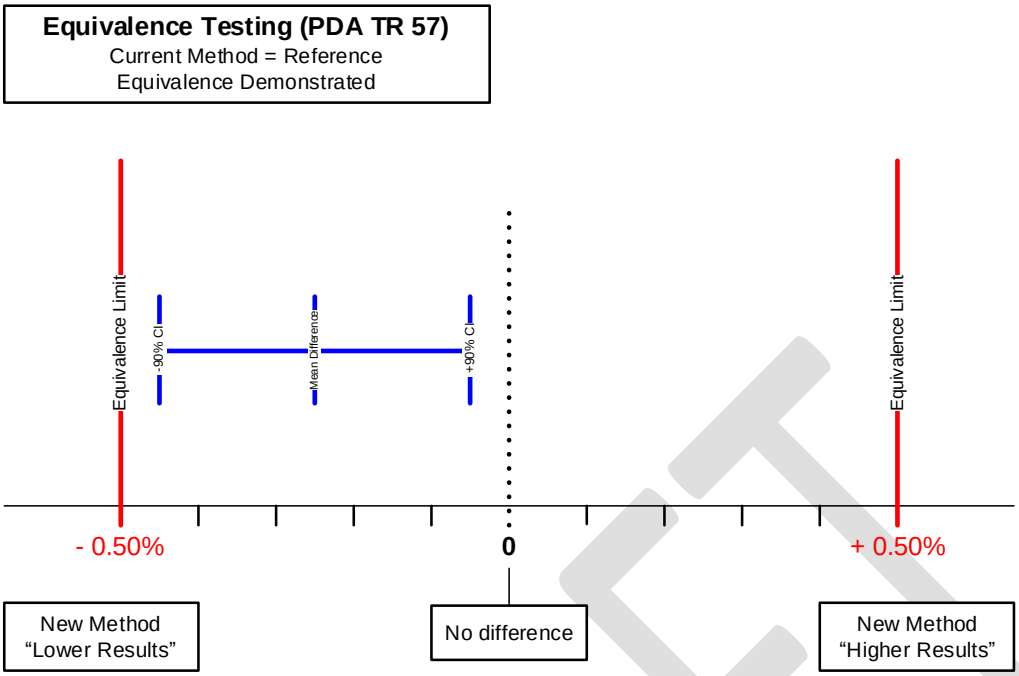
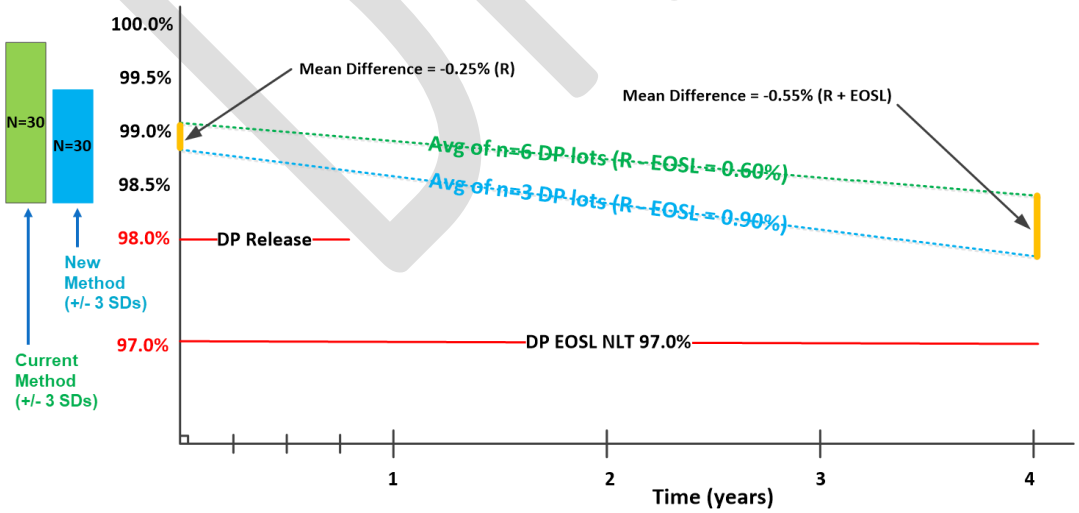


Figure 19 above, illustrates the DP release equivalence testing results. While statistically significantly different (90% confidence; two-sided), the 90% confidence interval falls entirely into the (+/-) 0.50% equivalence margins and, therefore, is considered a practically insignificant difference and passes the acceptance criteria of the equivalence testing. The mean difference for the new method was determined to be (-) 0.25% and future DP results are expected to have lower purity results by this bias when using the new method.

Figure 20: AMC Results for DP Release and Stability Testing – Case Study 1



AMC study results, illustrated in **Figure 20** above, show again the equivalence testing results for DP release with a mean difference of 0.25% (yellow bar on x-axis). Also shown are the distributions (mean +/- 3 SDs) of release testing results (n=30) for each method. Although the new method provided a bias of test result means of (-)0.25%, the tighter release result distribution resulting from the desirable decreased method variation, as indicated with the blue column above, does provide a manufacturability improvement ($CpK_{(new\ method)} > CpK_{(current\ method)}$).

The combined difference for the expected future total difference at the DP end of shelf-life (EOSL of 4 years) was estimated to be (-)0.55%. The net difference for DP stability (4 years) was (-)0.30% and within the acceptance criteria. When testing with a more sensitive and reliable method, the time- and condition-dependent degradation estimates for both methods were shown to diverge. The new method may detect a degradation of a product and/or change at an earlier timepoint. The overall manufacturability has been at least maintained when considering both DP release and stability specification limits ($CpK_{(new\ method)}$ no less than $CpK_{(current\ method)}$). As release and stability comparison acceptance criteria were set based on the desired outcome to leave the drug product/release specifications unchanged, the AMC results demonstrated the desired outcome.

Figure 21: Potential Equivalence Testing Outcomes and Interpretation

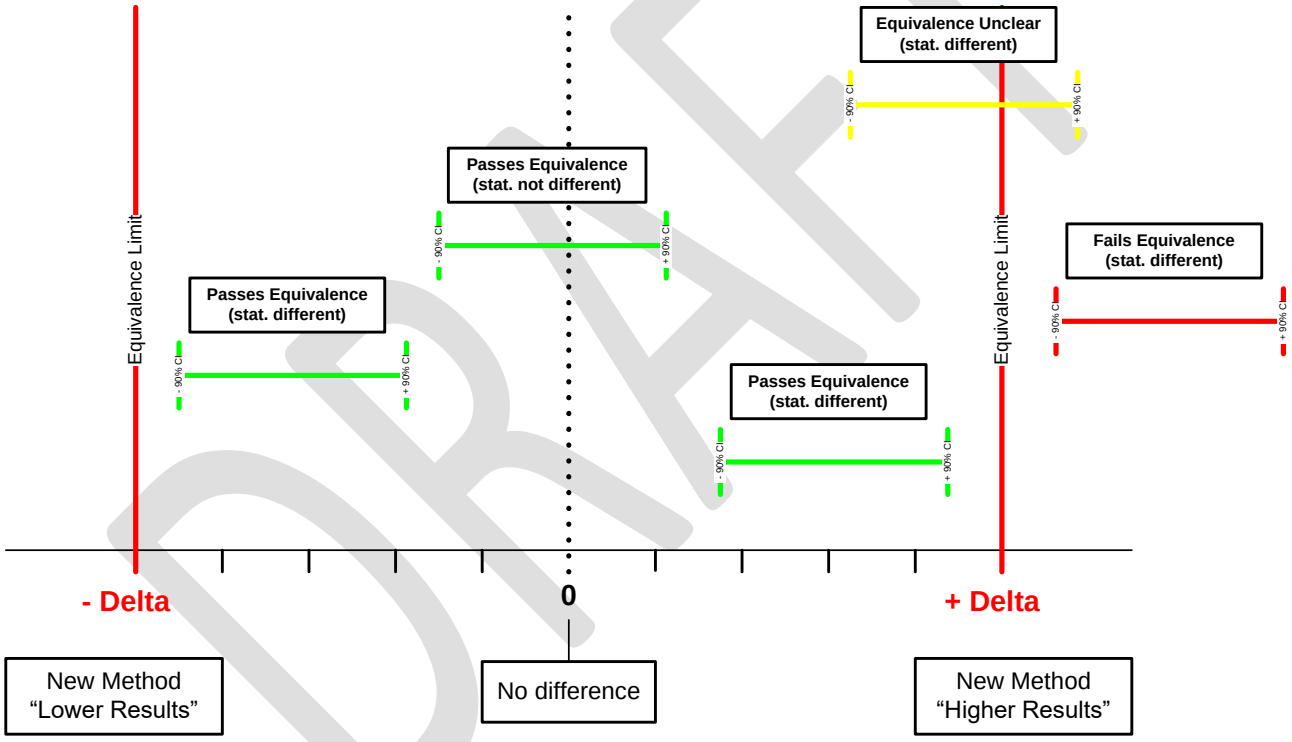


Figure 21 illustrates hypothetical AMC equivalence test study outcomes. Moving from left to right in the above figure, the AMC study results for the 90% CI for the matched-paired mean difference (left CI in green), using a two one-sided t-test (TOST), showed that we passed the equivalence margins (delta), although a statistically significant difference has been observed (at 90% confidence level). Based on the set-up of the AMC study, this method change can therefore be submitted or agency-approval without the request for a specification revision.

The next possible outcome (moving to the right) also passed the equivalence margins, but in addition, also failed to show a statistical difference. Like the most-left CI in **Figure 21**, the third (green) CI also passed equivalence margins while observing a statistically significant difference. Thus, all three (green) outcome scenarios will allow us to change methods without having to change specifications.

The 90% TOST CI overlapping the (+) equivalence delta is, in this specific acceptance condition set-up, the most undesirable outcome as we cannot demonstrate a risk-based equivalence, post-change impact for patient safety without changing specifications. On the other hand, we also do not have convincing justification given our comparison results and acceptance condition-based justification. Recovery options to consider for the least straightforward outcome could be, for example, to consider protocol amendment-driven additional comparison testing if the lot/material source sample size was relatively small.

The failing equivalence AMC study (red CI) will require a specification revision to compensate for the statistically significant, as well as the CpK-based practically significant, bias observed. Although the acceptance criteria were failed, convincing evidence exists that a specification revision is required when using the new method. The pre-approval regulatory submission for the new method should therefore include a request for a specification revision based on the AMC study outcome.

7.6.2 Case Study 2: Replacing a Compendial test with an Automated test: Rapid Sterility Test (Bact-T) to Replace USP/EP/JP Compendial Sterility Test

Case study 2 represents scenario 4 described in **Table 12**. A faster and technologically advanced Bact-T method for (upstream in-process) sterility testing is to be implemented and intended to replace the currently used compendial USP/EP/JP method for an autologous CGT product during late-stage product development. The rapid sterility test was validated and the detection capability for all potential, relevant ATCC microorganisms and manufacturing plant isolates is compared to the compendial EP/USP Sterility Test.

Using **Table 18** for the risk evaluation, the outcome is briefly summarized in **Table 19**. The risk evaluation requires a very high level of confidence that the new method performance is not inferior to that of the current compendial method. Non-inferiority model use can be justified based on the highly desirable faster testing and product release for autologous CGT product receiving patients. A highly significant reduction of from the current 14/28 days to ≤ 5 days for the routine testing duration is the primary justification for using a non-inferiority model. A relatively high number of comparison testing samples is required to demonstrate non-inferiority.

Table 18: Risk Evaluation for Compendial Sterility Test Replacement with Bact-T Test Method

Risk Evaluation Categories	Risk Variants	Case Study Example/Outcome
Attribute(s) Type - Intended use	Safety	Assuring sterile in-process intermediates and final product.
Attribute(s) Criticality (Intended use)	CQA	Controlling a significant CQA with potentially extremely severe patient safety impact.
Closeness to Finished Product	Intermediates and DP	Final control point for product release.
Intended Use	Routine Test	Primary sterility release and stability test to assure a safe product
Product Lifecycle	Late-stage Product Development (and intended for Commercial Manufacturing)	Large patient population(s) and highest possible patient impact. Highest regulatory expectations for approval.
Replacement Method "Status"	Current compendial method replaced with already agency-approved method which could potentially be a new PAP)	Extensive method performance knowledge with current compendial method. Significant industry experience exists using this newer technology for biologics and/or CGTs. Far less method performance knowledge/certainty with new method.

The method performance characteristics (qualitative test) are directly compared in AMC studies or indirectly from AMV study results following the suggestions from **Table 1.**

- **Detection Probabilities** for selected spiked microbial organisms at low levels (above/below compendial Sterility Test DL(s)).

- **Specificity** was compared simultaneously, directly by a retrospective and prospective comparison data from the AMV studies and prospective comparison testing. As expected, the detection capabilities and inferred specificity of the Bact-T method was not inferior for all relevant microorganisms and test samples (intermediates and DP). The specificity or lack of matrix and/or potential process impurity interference was determined to be acceptable and non-inferior with respect to the compendial sterility test.

7.6.2.1 Comparison testing model, study details, acceptance criteria, and result interpretation.

- The current proportions (presence or absence of microbial growth) of the USP/EP method is approx. 77% (23% false negative results with SD of approx. 5.0%) probabilities for selected spiked microbial organisms at low levels (above/below 5-day Sterility Test DL), see **Table 19**.

Non-inferiority comparison model was used and justified as more frequent testing will be scheduled, and test results obtained faster.

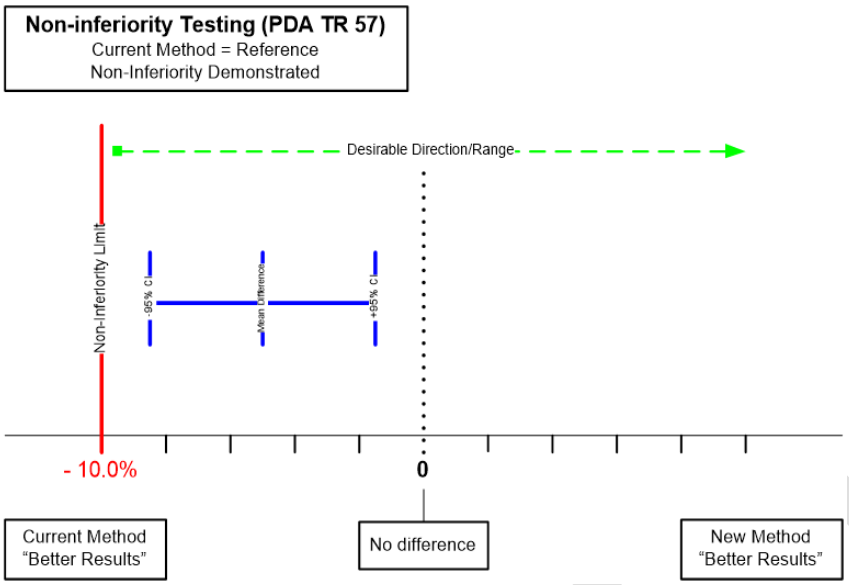
Justified non-inferiority (-) margin for individual or pooled one-sided 95% CI mean difference between compendial.

- Using a Proportions Test, the non-inferiority comparison acceptance criterion for a one-sided (lower) margin at the 95% confidence level (p=0.05) was established to be (–) 10.0%. The -10.0% limit versus the compendial (current) method was set and justified based on the compendial method performance (2 x EP/USP method standard deviations: 2x 5.0% = 10.0%) for the detection of (pooled) reference microbial organisms and/or plant isolates. Pooling of test results groups for selected microorganisms can be justified if growth/no-growth probabilities are similar among spiked organisms as confirmed in our case study. Comparison results are listed in **Table 19** and illustrated in **Figure 22** below.

Table 19: Pooled Growth/No-growth Percentages (Proportions) for Compendial and Bact-T Test Results

Method	Present	Total Samples (n)	Present/Absent Ratio
Candidate	**225**	**300**	**0.75 (75%)**
EP/USP	**232**	**300**	**0.77 (77%)**
Statistical Results			
Difference = p (new method) - p (EP/USP) Estimate for difference: -0.023 (-2.3%) **95% lower confidence interval limit for difference: -0.080 (-8.0%) (Limit = -10.0%)**			

Figure 22. Comparison of Compendial vs. Bact-T Detection Capabilities (Pooled Proportions) [13]



As demonstrated with a Proportions Test, illustrated in **Figure 22,** the 95% lower confidence limit for the justified pooled mean difference was above the non-inferiority limit of –10.0%. All three possible outcomes below are acceptable for the justified AMC comparison model used.

Non-Inferiority: The 95% lower CI falls entirely above the non-inferiority limit but does not overlap with the 0 difference (as illustrated in **Figure 22**).

Equivalence: The 95% CI overlaps with the 0 difference and this failure to show a statistically significant difference can be interpreted as equivalent.

Superiority: The 95% CI falls entirely above the non-inferiority limit but does not overlap with the 0 difference.

7.6.3 Case Study 3: Replacement of Host Cell Proteins (HCP)-specific antibody within an ELISA (enzyme-linked immunosorbent assays) method

Case study 2 represents scenario 3 described in **Table 12**; modifying a methos in a way that may affect the specificity or selectivity of the method. The removal of residual Host Cell Proteins (HCP) from the drug substance to an acceptable level is one of the goals of the biotherapeutics manufacturing process. HCP are the proteome of the production cell, and they are co-produced with the active molecule during cultivation of the cells in the bioreactor. These impurities might cause immunogenicity in certain patients, some may as well as act as enzymes and thus may negatively influence the stability of the drug product. HCP therefore need to be quantified. Host cell proteins are commonly quantified by ELISA using polyclonal antibodies.

The replacement of HCP antibody requires special consideration. Because the population of the anti-HCP antibodies might be drastically different between polyclonal antibodies (even generated using the same antigen material), the analytes detected (the independent proteins) are different and quantitative output of the assays are different as well. Additionally, to ensure detection of maximum number of HCPs potentially present in the drug substance, the polyclonal antibodies are generated using upstream process material which is also used as an ELISA standard. As the number of HCPs present in drug substance is much more limited, the accurate absolute

quantification of HCPs by ELISA is not feasible. Accordingly, the comparison activity is focused on the performance characteristics of the method rather than on direct comparison of quantitative outputs, establishing correlation between the 2 methods, and demonstrating the ability of the new method to differentiate between high and low levels of HCP in the sample. Replacement of HCP ELISA commonly requires reassessment of statistically established acceptance criteria.

The most common scenarios of HCP antibody replacement in the product lifecycle are shown below.

A -- Change Between Clinical Phases.

It is common to change antibodies used for HCP ELISA during product development This situation is summarized in Table "Case Study 3".

In the first clinical phases of a project, a commercially available cell line-specific HCP kits are generally used, while in a later phase, when control strategy for commercial manufacturing is developed, a process-specific kit, tailored for the specific process, is generated. In some instances, comparability between commercial and process-specific ELISA can't be established due to significant differences in antibodies used. In this case new acceptance criteria developed for process-specific assay using clinical trial data, to confirm safety on the new acceptance criteria. Of note, if capability of commercial assay is superior to process-specific one, commercial antibody may be used through the product lifecycle.

Table 20: Case Study 3– Risk Evaluation for the replacement of a generic HCP ELISA to a process specific ELISA

Risk Evaluation Categories	Risk Variants	Case Study Example/Outcome
Attribute(s) Type - Intended use	Safety (immunogenicity) Stability (HCP can have enzymatic activity)	Assuring HCP quantity is systematically below a given threshold
Attribute(s) Criticality (Intended use)	CQA	Controlling a significant CQA with potentially severe patient safety impact.
Closeness to Finished Product	DS	Final control point for DS release
Intended Use	Routine Test	Final control point for DS release
Product Lifecycle	Early-stage Product Development (and intended for Commercial Manufacturing)	Generic anti-HCP antibodies and HCP standard used in Phase I-II changed into process specific anti-HCP and HCP standard
Replacement Method "Status"	Generic method to commercial process specific method	Knowledge of the individual HCPs present. Coverage of the process specific ELISA (ie, number and proportion of specific HCP that are recognized) is very well understood

B-- Change during commercial manufacturing.

The risk increases if the change occurs during the commercial phase. The change can be triggered by the exhaustion of the stock of anti-HCP antibody. It may necessitate the production of new antibody batch and replacement of the method. This specific case requires a more stringent approach.

The extent of comparability studies depends on comparative performance of the polyclonal antibodies between new and old assays. The coverage of antibody against upstream HCP proteome and DS is commonly established using 2D-PAAGE, and bridging study is performed, quantifying HCPs in the set of samples by both methods. Of note, the samples should be representative of all materials tested, as the scenarios below may be different for materials at different process stages. The following scenarios are possible:

a) Comparable coverage, comparable results: assay is replaced without the change of specification.
b) Comparable or better coverage, different results, good correlation: assay is replaced, acceptance criteria are adjusted based on bias observed. Additional study on extended set of samples recommended to confirm acceptance criteria.
c) Significantly different coverage, different results, poor correlation (lack of observable systemic bias): suitability of method may still be established, and new acceptance criteria is established based on testing of extensive set of samples covering maximum range of HCP variability.
d) Poor coverage, including lack of binding to proteins present in the DS, different results, lack of correlation (no sensitivity to previously detectable differences) – assay is not suitable.

7.7 Comparison of New Technologies to Existing Technologies

Existing standards and guidance for analytical method comparison have traditionally focused on like-for-like direct comparative study designs, with the goal of demonstrating that the reportable parameter is unaffected. Statistical methods for assessment of method comparability in this framework have been established and are comprehensively described in USP <1010>, TR57 [2, 14] and the AMT chapter of this standard. However, there exists alternative approaches in which fundamental changes to the analytical methodology, or in which the demonstration of non-inferiority/superiority cannot be made with traditionally applied comparative analyses. One example is the implementation of in vitro assays to replace in vivo assays as outlined in the *Ph. Eur. 5.2.14* [15]. There is a concerted effort to develop in vitro methods to replace animal assays yielding a better comparison of product quality attributes, substantially reducing assay variability as well as time and resources required to execute the method. Quality attributes of the product will be assessed differently using an in vitro method therefore a 1-to-1 comparison is not appropriate. Key quality attributes necessary to ensure product safety and effectiveness must be sufficiently assessed using the in vitro method in place of the in vivo assay. The new method must ensure comparability of quality attributes between commercial lots and those found to be safe and efficacious in clinical studies. Alternative approaches to comparative analysis are necessary when there is a fundamental change in analytical principle with the replacement method, where the correlation between the methods is unlikely and the conventional approach to bridge the method(s) is not viable. The comparative analysis in the alternative approach typically focuses on the method performance, e.g., the new method is superior because it is more sensitive or because it offers a better control with less variability than the original method. This standard provides guidance and strategies with specific applied examples for such alternative approaches.

The replacement study may lead to a specification setting exercise as impact on the specification limit is likely, and data from this activity could be used for new specification setting. The study design and sample selection, therefore, should contain an adequate number and appropriate selection of samples to fully understand the difference in results between the methods for the intended use. For example, an in vivo potency assay has the potential to measure complex functional responses whereas one or more in vitro assays may be required to replace an in vivo assay resulting in the need to establish novel specifications with respective lower and upper limits.

7.8 Scenario 5 Case Studies

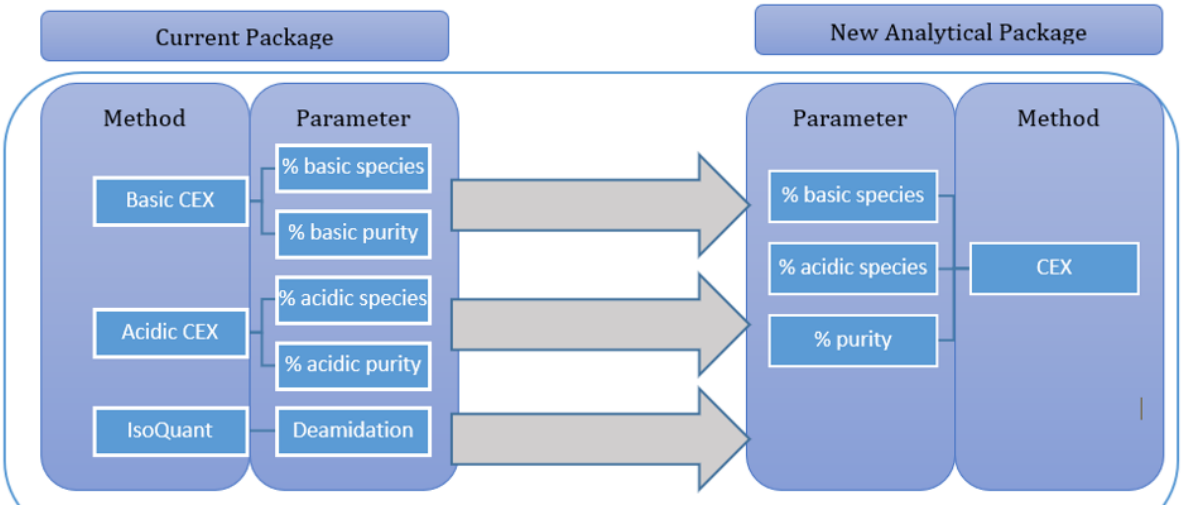
All case studies presented in **Section 7.8** represent **scenario 5** described in **Table 12**; new replacement method that monitors or controls for a different attribute readout, specification limit(s), and format for reported results.

7.8.1 Case Study 4: Replacing CEX and IsoQuant Test Methods with (only) CEX.

Charge variants are sometimes deemed as CQA for therapeutic antibodies. In the following case, two cation exchange (CEX) methods were used to control three CQA, namely acidic, basic and main charge variants; an additional method was used to quantify deamidation, which also appears in the charge variants. The complexity of this package led to logistics difficulties in the QC lab and was the trigger to rationalize and simplify the testing package, without of course losing any control on the quality of the product. This is an example of a complex change in the analytical package where the attribute readout (i.e., percentage basic, acid or purity) remains unchanged and therefore falls into scenario 1 (see **Table 12**).

The two methods, acidic and basic CEX, dedicated to the control of acidic species and basic species respectively, were merged into one. A third method, IsoQuant, used to control the level of deamidation, was also merged into the CEX method, by proving that one of the CEX peaks was correlated to the level of deamidation (see **Figure 23** below). The complexity of the change between methods required, among other things, a deep understanding of the identity of the sub-peaks, which was supported by characterisation studies and also cross injections of the subpeaks, from a variety of samples (stressed and non-stressed, at release or shelf life) between the current and new package to demonstrate that the control with the new analytical package was maintained (see **Table 21** for a summary).

Figure 23: Illustration of the current and new analytical package and quality attribute that are controlled for Case Study 4.



The outcome of the risk evaluation given in **Table 21** below is that a high level of confidence is required to demonstrate that the new method is not inferior to the current ones. It is noteworthy that the criticality of the current and future method remains unchanged because the CQA remain unchanged, even if the number of readings changes.

Table 21: Risk Categories and Evaluation for Case Study 4

Risk Evaluation Categories	Risk Variants	Outcome
Attribute(s) Type - Intended use	Purity and Stability	Controlling protein impurities and degradation products during manufacturing and storage.
Attribute(s) Criticality (Intended use)	CQA	Controlling 3 CQA that indicate process consistency: acidic, basic and main charge variants
Closeness to Finished Product	DP	Final control point for product release and shelf-life
Intended Use	Routine Test	Primary purity release and stability test to assure a safe product
Product Lifecycle	Commercial Production	Large patient population and highest possible patient impact. Highest regulatory expectations for approval.
Replacement Method "Status"	Current 3 methods replaced with 1 new method	Extensive method performance knowledge with current and new methods. Extensive characterization data required to demonstrate that the same species (in lieu of the "vague" CQA acidic and basic species) are covered by the new method. Extensive knowledge of the bias between the different readings.

Because of the complex merging of the three methods, a one-to-one replacement strategy was deemed inadequate. This example of scenario 1 required a maximum method comparison testing, but the comparison could not be done according to the concept of equivalence, especially because 5 readings yielded by the current package could not be correlated to 3 using the new method.

The advantages of the new method (increased accuracy and robustness, testing simplification) largely outweighed the risks (see **Table 21**). A specification revision was the next step in this example (see **Chapter 7.9** for more details). The revision was alleviated by the determination of the bias between the current and new methods at the specification limits, factoring the bias in the new specification limit and verifying the soundness of the new limit by using samples from pivotal clinical studies and by proving that the new package came to the same conclusion in terms of within / out of specification.

7.8.2 Case Study 5: Replacing an *In vivo* Potency Assay with an *In vitro* G-protein based ELISA for a vaccine.

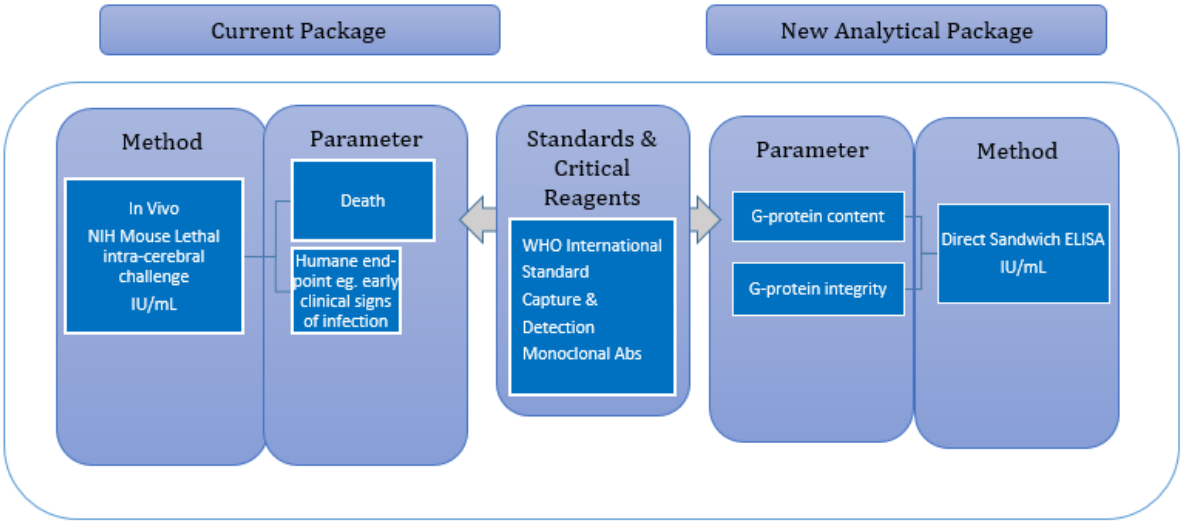
There are challenges when replacing an *in vivo* assay with an *in vitro* method to assess product quality attributes. A one-to-one replacement is only possible when both assays measure the same CQA therefore multiple *in vitro* assays may be required. Regardless, a change in method begins with in-depth product knowledge including the manufacturing process (e.g., CPPs, in-process controls) and the functional response of the drug product. A suitable comparability study design/protocol is required to determine that the alternate *in vitro* method will not negatively impact the assessment of product quality and provide the same level of post change control. Risk factors (**Table 13**) and approaches to consider (section 7.5) should be determined to establish appropriate study acceptance criteria demonstrating the new (*in vitro*) method is not inferior to the current (*in vivo*) method (**Table 22**).

The *in vivo* NIH mouse lethal intra-cerebral challenge test for the standardisation of rabies virus vaccine is used to assess potency as per EP 0216 [16] and WHO TRS 941Annex 2 [17]. This mouse protection assay was developed in the 1950's. There are several limitations including high variability with defined confidence limits of 25-400% making the test inappropriate to assess batch-to-batch consistency. Additionally, the assay has lethal

endpoints, requires the handling of live virus, thus presenting safety issues for workers. The mouse assay takes several weeks to complete, overall creating a very cumbersome method.

Various ELISA methods offering a rapid alternative, are in development to quantify rabies virus glycoprotein (G-protein) content to be used in place of the *in vivo* potency assay (see **Figure 24** below). Prior to the implementation of an *in vitro* ELISA, product characterization is imperative *i.e.*, the antigen conformation and content required to induce an immune response in animals must be determined. It is accepted that protection afforded by rabies vaccines is due to the production of virus-neutralising antibodies against the trimeric transmembrane glycoprotein following immunization. The ELISA is based on monoclonal antibodies capable of specifically recognizing the native form of the rabies viral G-protein which is responsible for the induction of neutralizing antibodies. Two sites, site III and site II of the G-protein are important to elicit a protective immune response therefore the ELISA employs two distinct antibodies, a coating (capture) antibody and a detection antibody directed towards these two sites. It is important to assess the specificity, binding affinity, and avidity of the antibodies toward the critical conformational epitopes of the antigen. The new assay should be validated in accordance with ICH Q2(R2) (Table 1, cat. IV). Demonstration that the alternative method meets validation performance criteria is not sufficient to imply comparability with a compendial method.

Figure 24: Illustration of the current and new analytical package and quality attributes that are controlled for Case Study 5.



Use of an *in vitro* method to replace a compendial *in vivo* method is conceptually similar at both pre- and post-market stages of product development. Both involve parallel *in vitro* and *in vivo* testing to obtain a comprehensive data set demonstrating reliable method performance. This example to assess glycoprotein content for a marketed rabies vaccine having a consistent manufacturing process and a known safety and efficacy profile employs a control strategy according to consistency testing. A head-to-head comparison of clinical lots might not be possible as the batches used in clinical trials to demonstrate efficacy and safety may no longer be available or are expired, therefore a consistency approach allows for the continuity of the clinical experience [16]. Parallel testing of a maximum number of batches (*i.e.*>200) covering a significant manufacturing period must be included in the comparability study. Furthermore, to adequately assess assay limits, sample selection (section 7.5.3) is of the utmost importance. The study design must include samples to demonstrate the assay's ability to the discriminate between potent and subpotent vaccine batches such as using

vaccine samples that have undergone forced degradation treatments. An assay's ability to discriminate compliant versus non-compliant lots suggests the method's stability indicating potential which is required for the assessment of vaccine stability through end of product shelf-life.

While there is no way to demonstrate a direct correlation between the NIH *in vivo* assay and the ELISA, due to high assay variability of the *in vivo* method and the fact that the assays measure different biological parameters, there needs to be an agreement between the two sets of data, *i.e*. a similar trend observed between the ELISA and the NIH *in vivo* test results. In the current example, both assays make use of an international standard that has been established for both the *in vivo* NIH mouse potency test and *in vitro* glycoprotein ELISA, facilitating parallel testing of vaccine lots representative of manufacturing capabilities. In addition, novel product specifications can be established for the new *in vitro* method (*i.e*. mean ± 3SD).

Table 22: Risk Evaluation for Compendial *in vivo* Potency Test Replacement with *in vitro* ELISA

Risk Evaluation Categories	Risk Variants	Case Study Example/Outcome
Attribute(s) Type - Intended use	G protein content (Potency)/ Stability	Assuring potency of final drug product
Attribute(s) Criticality (Intended use)	CQA	Controlling a significant CQA with potentially extreme severe patient and product impact.
Closeness to Finished Product	DP	Final control point for product release
Intended Use	Lot Release/Stability Test	Primary potency release and stability test to assure product efficacy
Product Lifecycle	Late-stage Product Development intended for Commercial lot release	Large patient population(s) and highest possible patient impact. Highest regulatory expectations for approval.
Replacement Method "Status"	Current *in vivo* compendial method replaced with novel *in vitro* method	Consistency approach ensuring product safety and efficacy requiring accurate product characterization and product specific criteria set using lots representative of manufacturing variability. Extensive historical data using highly variable *in vivo* compendial method, less industry experience using this newer *in vitro* method. International collaborative study is ongoing to further validate the transferability and robustness of the *in vitro* ELISA. Novel product specifications required.

7.8.3 Case Study 6: Replacing three glycosylation methods with one multi-attribute method (MAM)

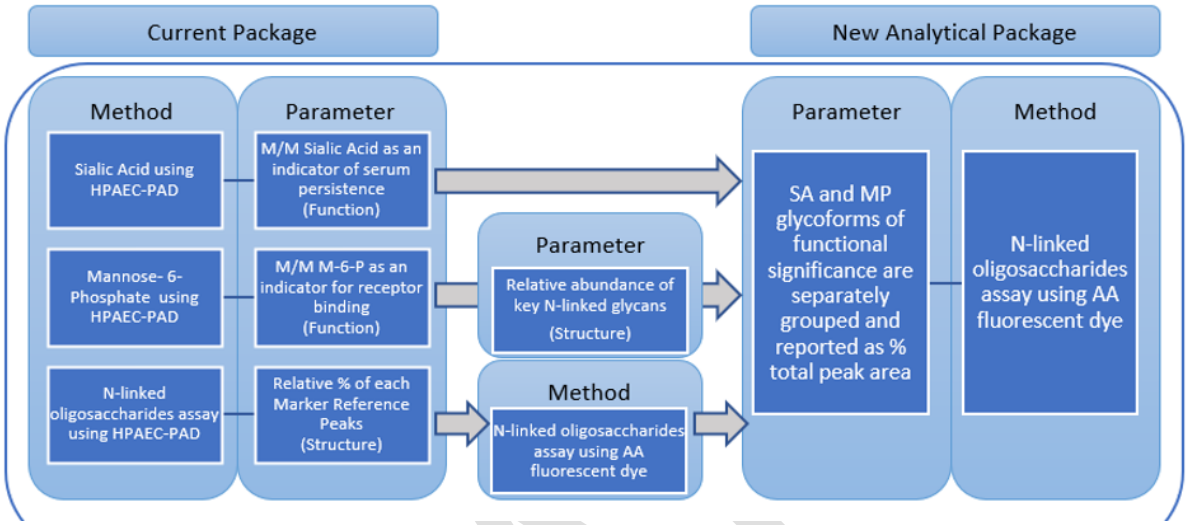
More than two thirds of therapeutic proteins are glycoproteins—proteins that contain glycans or chains of carbohydrates. These chains become attached to the protein via a post-translational modification process called glycosylation. Glycosylation plays an important role in structure, function, absorption, half-life, clearance, and safety of therapeutic proteins.

In this example, we illustrate how three assays used for glycosylation composition and profiling using HPAEC-PAD technology were replaced with one, more robust method using AA labelling to monitor glycosylation (see **Figure 25** and **Table 23** below).

The process for replacing the methods was achieved using a phased approach. In Phase I, (intermediate package), the high-pH Anion Exchange Chromatography (HPAEC) method with online pulsed amperometric detection for unlabelled glycans was replaced with N-linked oligosaccharides profiling assay employing fluorescent dye labelling. As part of the continuous improvement effort on control strategy, two additional

assays that provide overlapping information for monitoring glycosylation, Mannose-6-Phosphate and Sialic Acid, were re-evaluated using data compiled from all commercial lots to propose a simplified grouping approach for reporting N-linked glycans distribution using the AA-labeled oligosaccharide assay for Phase II, (new analytical package).

Figure 25: Illustration of the current and new analytical package and quality attributes that are controlled for Case Study 6.



Phase I

The AA labelling method was intended to replace the non-labelling oligosaccharide profiling method to monitor product lot to lot structural consistency. The AA labelling method represents a more updated analytical approach commensurate with industry current standard and is the most informative in terms of discerning glycosylation species heterogeneity as well as providing robust relative quantitation.

Analytical method comparability included a side-by-side comparison of new to registered method using the traditional approach for statistical analysis of significance, however, the eight marker reference peaks identified in the original method were selected based mainly on the capability of the method to discern peak entities and did not directly correspond to the two functional glycosylation critical quality attributes of the molecule. The AA labeling method was able to discern 11 major AA-labeled peaks, (8 of which with functional relevance), structural identities of these key N-linked oligosaccharide species were determined by HPLC coupled with mass spectrometry. As a result, extensive scientific justification was provided for the intermediate package that included rationale for superiority of the new method based on its capabilities of improved analytical discriminate properties for relevant functional peaks. New specifications for the AA-labelling method were proposed, based on a somewhat limited data set (n=24), and submitted and approved in one jurisdiction.

Phase II

Even though the new AA-labelling method was approved in one jurisdiction, there were still multiple markets that required release testing using the original HPAEC acid release oligo method. Which meant, concurrent testing of both methods was performed for a period. In that time, a much broader data set was developed and utilized to provide the rationale for the next phase of process control continuous improvement, a proposal to consolidate the glycosylation specifications into two key groups corresponding to two key functions of the product, serum persistence and target cell uptake. The simplified grouping approach allows establishment of release criteria for meaningful lot disposition decisions based on current understanding of functional significance of the product glycoforms and could replace the current release paradigm routinely used to monitor product glycosylation: (1) acid-released sialic acids (N-acetyl neuraminic acid and N-glycolyl

neuraminic acid); (2) acid-released mannose- 6-phosphate,and (3) in the remaining markets, enzyme-released N-linked oligosaccharides assay using HPLC with pulsed-amperometric detection.

Key glycan peaks as monitored by the AA-labelling method were identified based on biochemical identity and predicted glycan structure as determined by LC/MS.

Product experience, as well as published literature on the native enzyme indicated that glycoforms with terminal mannose-6-phosphate (M6P) and sialic acids can impact efficacy. In the new grouping proposal, glycoforms with terminal phosphorylated mannose and glycoforms with terminal sialic acid are combined into two respective groups based on functional significance.

The peak group totals are converted to unit-less values and termed “mannose phosphorylation index” (MP index) and “sialylation index” (SA index). As such, the MP index is calculated by combining peak area percentage values of the MP peaks, and SA index is calculated by combining peak area percentage values of the SA peaks.

In each of the two cases, the index gives a snapshot of the distribution of glycoforms in a single parameter relevant to functional significance. This simplified approach improves robustness of the specification reporting, as minor variation in biologically similar glycosylation species (ex. branched positional isomers) will not materially alter the reportable results. Overall, this allows for robust and straight-forward tracking and trending of the product and process for the commercial operation.

Orthogonal data, including process change characterization assay glycopeptide analysis by LC/MS and routine release assay receptor binding by Biacore, indicated a significant correlation to the AA-labelling method across multiple process changes. Taken together, the simplified grouping approach for reporting glycoforms has well-grounded scientific justifications.

Table 23: Risk Categories and Evaluation for Case Study 6

Risk Evaluation Categories	Risk Variants	Case Study Example/Outcome
Attribute(s) Type - Intended use	Quality	Enumeration of terminal sialic acid and M6P residues has limited functional significance as compared to monitoring glycan structures via the SA index and MP index.
Attribute(s) Criticality (Intended use) severity of what you are controlling?	CQA	Under-glycosylation (or occupancy) may negatively affect protein function and stability. Process variables that impact glycosylation of functional glycans are identified and controlled accordingly to ensure batch consistency.
Closeness to Finished Product	DS	Final control point for DS release
Intended Use	Routine Test	Final control point for DS release
Product Lifecycle	Commercial Manufacturing	A comprehensive study to develop new product specifications was submitted as part of the continuous improvement effort on control strategy.

		The request defined the previously generated data set, evaluation procedures, and criteria for establishing new product specifications for the three product attributes.
Replacement Method "Status"	Generic method to commercial process specific method	The three glycosylation methods described in this case study are considered platform analytical procedures for multiple products, with some modifications specific for each product. In addition to providing more robust release data and analytical testing relief for this product, the phased approach strategy described here, providing scientific justification and analysis of previously generated data, will influence future regulatory submissions for all commercially available family products, in addition to all follow-on products, to support improved and enhanced release testing paradigms for future use.

New Specifications

Process data for 92 lots of drug substance material were used to support setting new specifications for the SA and MP indices. This material included lots manufactured by several processes. Extensive comparability studies, performed at the time of manufacturing process changes, indicated no change in the material as a result. The direct correlation between the registered composition methods to the proposed SA and MP grouped peak area percentages using reprocessed historical production data was not statistically significant, however, it demonstrated that the reprocessed data could successfully identify process variables that impacted sialylation and phosphorylation and proved to be a better indicator for monitoring batch consistency.

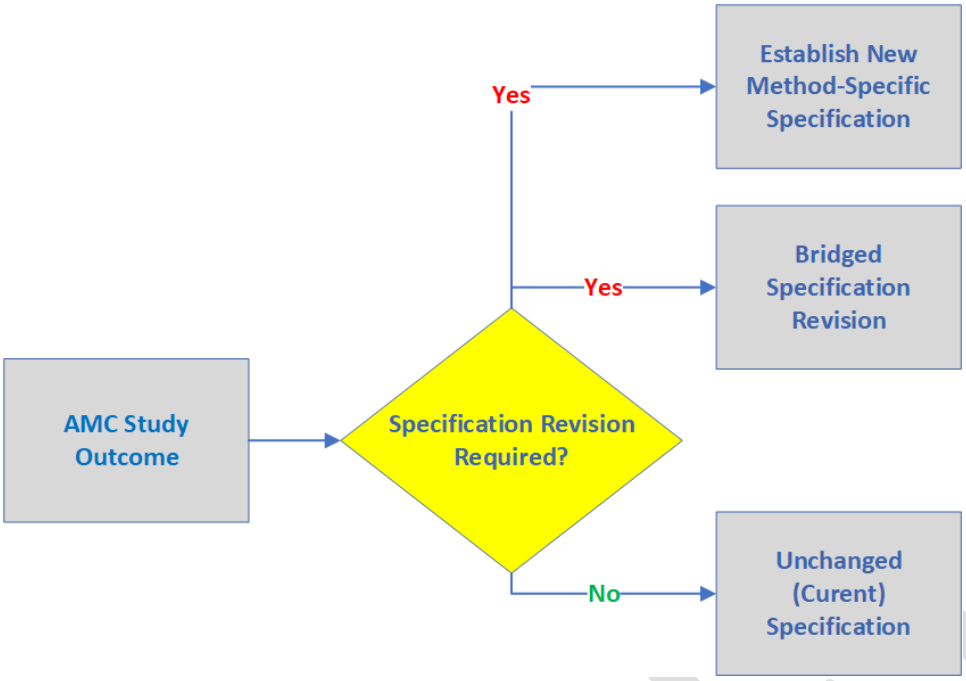
In summary, the three registered methods provided only semi-quantitative results with overlapping information for monitoring glycosylation heterogeneity of a glycoprotein with multiple glycosylation sites. The new single method is more robust and quantitative. Identification of functionally significant peak entities, along with a simplified grouping approach for reporting, allowed for meaningful lot disposition decisions to be based on the functional significance of the product glycoforms, in addition to a reduction in the release testing paradigm.

7.9 Evaluation of Product Specifications

Specifications are one part of an overall control strategy designed to ensure product safety, efficacy, and quality. Specifications are chosen to confirm the quality rather than to characterize the product, so the rationale and justification for including and/or excluding testing for specific quality attributes should be clearly described in the BLA.

When analytical methods used to test against product specifications are replaced, or significantly changed, the existing product specifications should be assessed and/or revised following the decision tree illustrated in **Figure 26** below.

Figure 26: Evaluating Product Specifications



Three possible AMC study outcomes for any of the five change scenarios, illustrated and described in **Section 7.2**, *Method Change Scenarios*, are shown in **Figure 26**. When the expected test result shift and/or variation is insignificant, as indicated by comparing the manufacturing capability (CpK) pre- and post-change, a specification revision may not be required. When the shift and/or variation is significant, a revision becomes necessary. The required specification revision is established based on the scenarios described in **Section 7.2** and should consider the following:

- Specifications limit the level and/or variation of critical quality attributes and are linked to the manufacturing capability. The significance of alterations in product-related substances, product-related impurities, and process-related impurities, which may differ from those observed in the material used during clinical development, should be evaluated.
- Specifications should account for the stability of drug substances and drug products. For analytical methods related to the stability profile, the new or revised method should be evaluated so that changes in the quality of the product will be detected, and product stability assured.
- Specifications are risk-based linked to clinical studies. Data from the new or revised method should be compared to data obtained for lots used in clinical studies, to ensure that the quality of the material made later-stage product development and/or commercial scale can be compared and considered to be representative of the lots used in clinical studies.
- Specifications are linked to analytical procedures. It is important to confirm that data generated by the new or revised method correlates with data generated by the previous method.
- A proper understanding of the bias between the methods (current and new) can help correct the current specification into the new one. It is acknowledged that this might seem to be a pure mathematical correction of a limit, and therefore attention should be paid to the scientific soundness of the corrected limit. This could be checked with, e.g., samples with clinical coverage, so that there is some extra assurance that the same level of control (in specification or out of specification) is unchanged between the current and new methods.
- The evaluation of product specification, especially in the early phases of a project, could be a sound verification around the specification limit. For example, it could be checked that a result declared OOS by the former method would still be OOS using the new method and, vice-versa, that the two methods declare a sample consistently within specification (without focusing on the numerical value obtained with

the two methods). The effort can be increased, especially in the case of a commercial specification, to understand the potential correlation between the methods on the whole analytical method range. For specification limits that are not data-driven (which is typical for early clinical phases), the verification on "within specifications" / "OOS" using the 2 methods might suffice.

- Certain methods are part of a non-routine testing, without specification (e.g., PPQ supportive methods, used to demonstrate clearance). demonstration that the clearance profile is comparable between the 2 methods might be sufficient, without exploring the correlation between the analytical values at each process step.

8 Bibliography

1. Krause, S., *PCMO L01-Setting Specifications for Biological Investigational Medicinal Products*, PDA J Pharm Sci Technol., 2015 Sep-Oct;69(5):569-89.
2. Parenteral Drug Association, Technical Report No. 57: *Analytical Method Validation and Transfer for Biotechnology Products.* 2012, PDA: Bethesda, Md. p. 75.

3. International Conference on Harmonisation., Quality Guideline, *ICH Q2(R2): Validation of Analytical Procedures - Scientific guideline*. 2022, ICH: Geneva.
4. International Council for Harmonisation, Final Concept Paper, *ICH Q14: Analytical Procedure Development and Revision of Q2(R1) Analytical Validation*. 2018, ICH: Geneva. p. 1-3.
5. U.S. Food and Drug Administration, *Guidance for Industry: Analytical Procedures and Methods Validation for Drugs and Biologics*. 2015, U.S. Department of Health and Human Services: Silver Spring, Md. p. 18.
6. Krause, S., *Using Analytical Platform Technologies to Support Accelerated Product Development—Concept Review and Case Study*.; PDA Journal of Pharmaceutical Science and Technology PDA Journal of Pharmaceutical Science and Technology 2022. 76 (3): p. 248-262.
7. U.S. Pharmacopeial Convention. General Chapter <1033> *Biological Assay Validation. 40–NF 35*, 2015; USP: Rockville, MD.
8. Parenteral Drug Association, Technical Report No. 57-2: *Analytical Method Development and Qualification for Biotechnology Products*. 2015, PDA: Bethesda, Md. p. 60.
9. Duval, A.E., Broly, H., Gleixner, R., *Quality attributes of recombinant therapeutic proteins: An assessment of impact on safety and efficacy as part of a quality by design development approach.* Biotechnolegy Progress; May/ June 2012; Volume 28, Issue 3; Pages: 595-892.
10. Alt, N.,Zhang, T.,Motchnik, P.,Taticek, R., Quarmby, V.,Schlothauer, T., Beck, H., Emrich, T., Harris, R. J.. *Determination of critical quality attributes for monoclonal antibodies using quality by design principles*. Biologicals; 2016; Vol. 44, Issue 5, pages 291-305.
11. The CMC Vaccine Working Group, *A-VAX: Applying Quality by Design to Vaccines*. May 2012. https://qbdworks.com/wp-content/uploads/2014/06/a-vax-applying-qbd-to-vaccines.pdf
12. The CMC Biotech Working Group, *A–Mab: A Case Study in Bioprocess Development*. October 2009. https://ispe.org/publications/guidance-documents/a-mab-case-study-in-bioprocess-development
13. Krause, S., *Strategies for Analytical Method Replacements for Approved Products, Presentation at WCBP Meeting*, Washington, DC, January 2017; pages 24-26.
14. U.S. Pharmacopeial Convention. General Chapter <1010>*Analytical Data – Interpretation and Treatment*. 2021, USP: Rockville, MD.
15. Ph. Eur. *Substitution of in vivo method(s) by in vitro method(s) for the quality control of vaccines (general chapter 5.2.14).*; January 2018.
16. Dierick, J.F., Halder, M., Jungbaeck, C., Lorenz, J., Preaud, J.M., Riou, P., Tesolin, L., Uhlrich, S., Molle, W.V.,Vandeputte, J. *The consistency approach for the substitution of in vivo testing for the quality control of established vaccines: practical considerations and progressive vision*. Open Research Europe, Version 2, December 2022. The consistency approach for the ... | Open Research Europe (europa.eu) .
17. World Health Organization. *Recommendations for inactivated rabies vaccine for human use produced in cell substrates and embryonated eggs, Annex 2, TRS No 941*; January 2007. https://www.who.int/publications/m/item/inactivated-rabies-vaccine-for-human-use-annex-2-trs-no-941.