european biopharmaceutical enterprises

ebe

EBE

Concept Paper

Considerations in Setting Specifications

March 28, 2013

European Biopharmaceutical Enterprises (EBE), a specialised group of EFPIA

Leopold Plaza Building
Rue du Trône 108

BE-1050 Brussels

www.ebe-biopharma.org
info@ebe-biopharma.org

Table of Content

Abreviations

AE	Adverse Event
ATA	Anti Therapeutic Antibody
CQA	Critical Quality Attributes
DP	Drug Product
DS	Drug Substance
EMA	European Medicinal Agency
cGMP	current Good Manufacturing Practice
HCP	Host Cell Protein
ICH	International Conference on Harmonization
IPC	In-Process Control
mAb	Monoclonal Antibody
NLT	Not Less Than
PD	Pharmacodynamic
PK	Pharmacokinetic
PQIT	Periodic Quality Indicator Test
QA	Quality Attributes
QbD	Quality by Design
QMS	Quality Management System
QRM	Quality Risk Management
QTPP	Quality Target Product Profile
RTRT	Real Time Release Testing
SOP	Standard Operating Procedures
TPP	Target Product Profile

1 Objective / Scope

The purpose of the concept paper is to discuss considerations for setting of specifications for commercial biotech[1] drug substances (DS) and drug products (DP) as part of an overall control strategy.

The document identifies three relevant steps:

1. Criticality assessment for product attributes.
2. Establishment of specifications as part of an overall control strategy.
3. Definition of acceptance criteria.

Recent developments in the regulatory field have focused on the application of risk-based approaches to the control of product quality based on an enhanced process and product understanding (e.g. QbD). This concept paper is applicable in a "traditional" development setting as well as in cases where such enhanced process understanding has been achieved for a particular product. It is intended that this concept paper will promote a consistent application of the existing guidance on specification setting by providing guidance based on the experience of EBE member companies.

2 Introduction to Topic

Various regulatory guidelines are currently available that deal with the setting of specifications for biotech products, most notably ICH Q6B. Since its publication in 1999, numerous advancements in the field of setting of specifications have been made (e.g. with respect to product characterization, process understanding and applicability of risk based tools) as described in ICH Q2(R1), Q8(R2), Q9, Q10 and Q11. This concept paper is not intended to replace the available regulatory guidelines, but is instead intended to supplement the guidance available by providing expanded discussions on certain topics with the intent to support applicants in meeting the current expectations of regulators. Furthermore, it should be noted that this concept paper focuses on the establishment of specifications as part of an overall control strategy, but does not attempt to describe in full how the complete control strategy may be designed. It should also be noted that the examples provided in Section 6 of this document are cases shared with companies contributing to this document, and are not to be considered as direct applications of the principles discussed.

Despite the availability of the ICH guidelines, specifications remain an area where questions are routinely raised by agencies during the review of marketing applications and amendments thereof. This concept paper reflects the combined experience of the contributors and attempts to address topics that have resulted in agency review questions and divergence of approved specifications between different regulatory authorities. In particular, the concept paper reflects the ongoing dialogue between the EMA, Biotech Working Party and EBE and present global thinking on the topic of specification setting.

[1] The scope of the concept paper is specifically intended to address specifications for products of biotechnology such as therapeutic proteins and monoclonal antibodies, but aspects may be applicable to other product types.

2.1 Overall Control Strategy

A control strategy is defined as a planned set of controls, derived from current product and process understanding that assures process performance and product quality, of which the specifications is one part, to ensure product quality and consistency (ICH Q6B and Q11). The controls can include parameters and attributes related to drug substance and drug product materials and components, facility and equipment operating conditions, in-process controls, finished product specifications, and the associated methods and frequency of monitoring and control (ICH Q10). The primary elements of the control strategy can be divided into the following categories:

- **Control of material attributes**

The quality of materials attributes such as, raw materials, source materials, reagents, primary-packaging materials etc. is critical for the product quality.

- **Process parameter controls**

The control of the process performance should mainly be driven by the critical process parameters. Accordingly the process parameter controls are an essential element of the control strategy as they assure consistent production of a medicinal product with the desired quality controlling process parameters that are functionally related to critical quality attributes is an ideal control to assure product quality in the control strategy.

- **In process testing**

Testing of certain critical quality attributes (CQAs) can be performed for each batch on in-process samples in lieu of end product testing or to monitor and if appropriate to adjust the process. Criteria may be defined as alert and/or action and/or reject according to the internal quality system. This may include defining of the adjustment of the subsequent operating units, e.g. column load, calibration etc.

- **Testing of intermediates**

A product hold time is defined when a product pool is being held during the manufacturing process. In general, a product intermediate is defined as a product pool that is isolated and stored for a prolonged period between process steps. The acceptable "hold time" should be defined based on data from microbial purity studies and, if relevant, degradation testing. Intermediates will have specifications and the storage time will be supported by stability data conducted an in representative storage container.

- **Procedural controls**

Procedural controls are usually documented in SOPs, i.e. as part of the manufacturer's Quality Management System (QMS) and cGMP. There are many different types of procedural controls, which may be used to support commercial manufacture, e.g., procedural controls can be used to ensure that non-critical process parameters such as buffer pH are maintained within the specified ranges, as well as facility/equipment controls

- **End-product controls by release testing**

Testing performed on drug substance and drug product according to a specification that consists of a defined set of test methods with predefined acceptance criteria (ICH Q6B).

- **Stability Monitoring**

The shelf life of a product is determined by stability testing according to protocol, using predefined specifications, including test methods and acceptance criteria. The stability of a product after licensure is monitored as defined in a stability program, to ensure that quality is maintained. The details of the program for establishment of product shelf life and design program for monitoring of product stability is described in detail in ICH Q5C.

- **"Life cycle management"**

During the life cycle of a product, systems should be in place to ensure that the control strategy is reassessed periodically based on emerging knowledge, manufacturing performance and following any changes to the process (e.g. annual product quality review, post approval life cycle management plan, comparability assessments etc.).

Figure 1 below describes in summary how Risk Management is performed to consider the control strategy based on the level of impact on safety and efficacy, on the availability of suitable methods to analyze and quantify specific attributes, and on the availability of data to show that the process is capable of consistently generating desired attributes, and removing impurities. This information is used to design the control strategy for a specific product and indication.

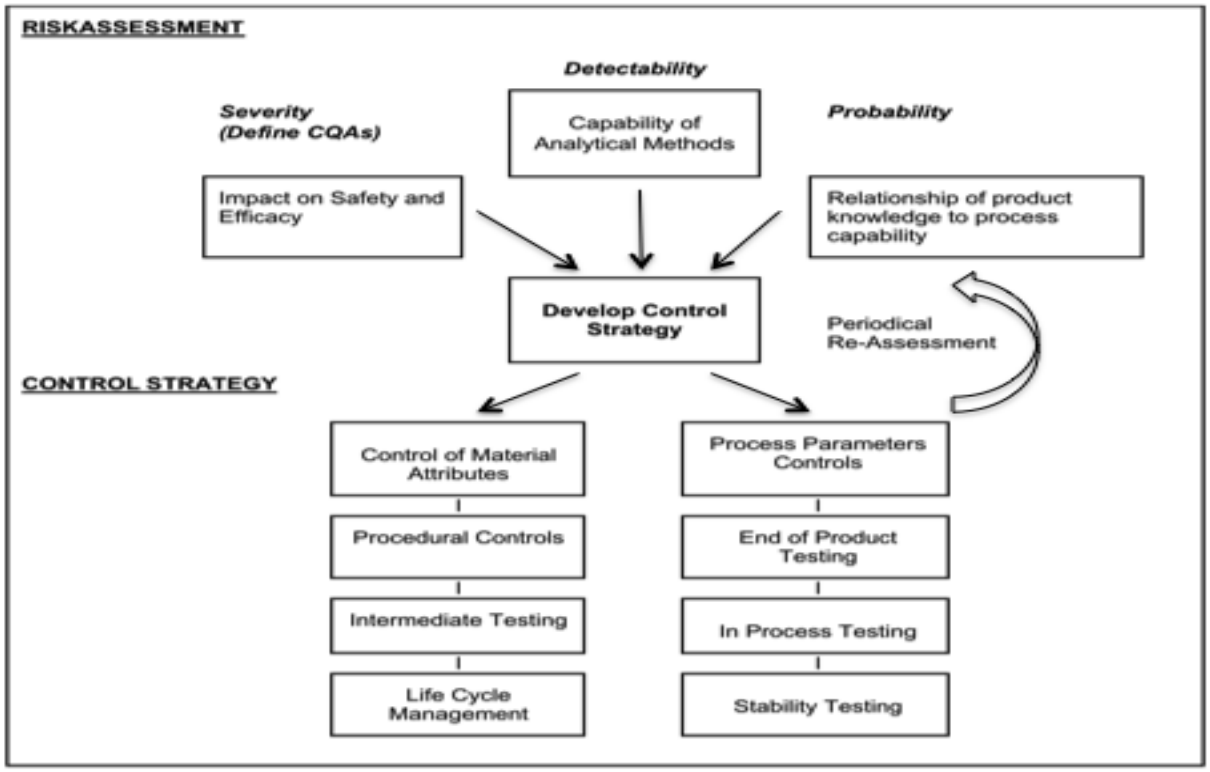

Figure 1 Risk Assessment Model as Applied to Design of the Control Strategy[2]

[2] Reproduced from, Schenerman, MA, Axley, MJ, Oliver, CN, Ram, K, and Wasserman, GF (2009), Using a Risk Assessment Process to Determine Criticality of Product Quality Attributes. In: Rathore, AS, Mhatre, R, Quality by Design for Biopharmaceuticals: Perspectives and Case Studies. Wiley Interscience. pp. 53-84.

The main subject of this document is the establishment of the specifications for a marketed product as part of an overall control strategy that includes all of the above elements. The derivation of the other elements of the control strategy will not be discussed further.

3 Identification of Critical Quality Attributes

The overall development process for a product generally, independent of whether an enhanced or traditional approach is used, starts with the definition of a target product profile (TPP), which defines the desired efficacy and safety of the drug product and key aspects related to the supply and use of the product. The criticality assessment begins with the Quality Target Product Profile (QTPP) as defined in ICH Q11 and Q8 as "a prospective summary of the quality characteristics of a drug product that ideally will be achieved to ensure the desired quality, taking into account safety and efficacy of the drug product." Essentially, the QTPP defines the quality characteristics that should be met in order for the TPP to be achieved. The QTPP is an iterative document that is established early in development and undergoes modification through product development as dosages; strengths, etc. are defined by results from clinical studies. Once the QTPP is established, the detailed process of identifying the critical quality attributes of the product may be performed, as described below:

1. An initial criticality assessment is performed based on the QTPP and prior knowledge of the product type in order to identify a list of **potential** critical quality attributes. This typically takes place early during product development and helps to design experiments for product and process characterization studies.
2. Based on the results of the product characterization studies and further assessment against the QTPP, including available knowledge of the impact of the quality attributes on safety and efficacy, the list of CQAs is confirmed. The list of CQAs is typically established during development and finalized prior to marketing application.
3. As part of the criticality assessment, a criticality rating is typically assigned to each attribute, which will subsequently be assessed together with other knowledge, such as information on how the process impacts the identified CQAs, the detectability of the CQA etc. in order to determine a control strategy. Reassessment of the criticality will continue as part of Life Cycle Management as more data is available, or related to changes of process or methods, and their ability to remove or detect an attribute.

Quality Risk Management (QRM, as described in ICH Q9) can be used to assess the risk and criticality of variability in the identified product quality attributes during manufacture and the resulting analyses form the basis of setting manufacturing processing parameter controls. Knowledge management (as described in ICH Q10) is key to capturing and applying prior knowledge of biological, chemical and manufacturing principles and experience to the establishment of the QTPP and also assessing the criticality of quality attributes and the degree to which a control strategy needs to be employed.

3.1 Criticality Assessment Procedure for product-related variants.

This section describes how to assess the criticality of quality attributes related to product variants. While ICH Q6B describes only critical and non-critical variants - product-related impurities and product-related substances - in this approach criticality of product related QAs is a continuum.

Assessing the criticality of the quality attributes requires knowledge of the product's structure-function relationships, degradation pathways, and potential process-related impurities, as well as information about the intended indication and use of the product.

In assessing the critical quality attributes of a given product, it is important to consider the reliability of the dataset on which conclusions are being made. It will rarely, if ever, be the case that the effect of variability in a particular attribute will have been studied directly in a suitably powered clinical safety/efficacy study which is unequivocally relevant to the product in question and all potential indications and patient populations. As such, a weight of evidence approach is generally applied where the available data from literature, similar products, non-clinical and clinical studies are collectively assessed. With such an approach, it is essential to carefully consider any gaps or inadequacies in the data, in order that the conclusions are appropriately balanced and the uncertainties in the assessment are acknowledged.

In general, criticality should be considered as a continuum, with each attribute having a certain degree of criticality according to its impact on safety and efficacy. The degree of certainty in that assessment should also be considered when making the criticality assessment. Therefore, a different level of criticality can be assigned to each attribute, depending on its assessed impact on safety, efficacy, immunogenicity and PK/PD. The amount of control should be related to the level of criticality. It may be possible to determine at this stage that certain attributes are of sufficiently low criticality and may be excluded from further consideration within the design of the control strategy. Other attributes may be excluded from routine testing as their control through other measures is found to be sufficient, as described in Section 2.1 Overall Control Strategy.

The assessment of criticality lends itself to the use of risk scoring tools, such as the example provided in Table 1 below. Where such tools are employed, it is important to provide both the summarized results and an explanation of the assessment procedure within the application.

The key to criticality assessment is to determine the extent to which prior knowledge can be applied to a particular attribute of a particular molecule, e.g. whether a prior example can be considered to have the same properties with respect to safety and efficacy as the molecule under assessment, taking account of the intended use of the product. Where the prior knowledge is considered insufficient, prospectively designed studies can be performed in order to provide further data and information to support the criticality assessment. For instance, studies based on fractionation of variants and determination of biological activity by in vitro potency assays can be used in order to determine the *in-vivo* relevance of certain variants. However, it should be noted that demonstration of equivalent *in-vitro* potency is not sufficient in itself to claim that a variant has the same properties as the desired product. The possibility that specific variants may impact the pharmacokinetic/pharmacodynamic profile or immunogenicity/safety of the product must also be considered. With respect to PK, this may be assessed through a combination of *in-vitro* (bioassays such as FcRN binding assay) and *in-vivo* (analysis of samples obtained from non-clinical or clinical PK studies) techniques as well as assessment of published literature references. With respect to immunogenicity and safety, this may prove challenging to assess in-vivo, especially in the absence of relevant animal models. As such, greater reliance may need to be placed on published literature to assess whether the type of variant is typically found *in-vivo* and/or contains structural motifs that may be associated with a specific response (e.g. immunogenicity). The fact that a particular variant was present in nonclinical and/or clinical lots is not in itself sufficient to conclude that a variant is an attribute of low criticality without a detailed analysis of the quality of the (non-) clinical lot(s) and the design(s) and outcome(s) of the study(ies). Table 1 provides an example of how these different components of the criticality assessment can be evaluated.

Table 1. Example of Impact Definition and Scale from A-Mab[3]

Impact (Score)	Biological Activity or Efficacy	PK/PDa	Immunogenicity	Safety
Very High (20)	Very significant change	Significant change on PK	ATA detected and confers limits on safety	Irreversible AEs
High (16)	Significant change	Moderate change with impact on PD	ATA detected and confers limits on efficacy	Reversible AEs
Moderate (12)	Moderate change	Moderate change with **no** impact on PD	ATA detected with *in-vivo* effect that can be managed	Manageable AEs
Low (4)	Acceptable change	Acceptable change with **no** impact on PD	ATA detected with minimal *in-vivo* effect	Minor, transient AEs
None (2)	No change	**No** impact on PK or PD	ATA not detected or ATA detected with no relevant *in- vivo* effect	No AEs

AE = adverse event; ATA = anti-therapeutic antibody
A quantitative criteria should be established for biological activity/efficacy and PK/PD. Significance of the change is assessed relative to CQA variability.

As a case in point, a long-standing debate has occurred as to whether glycosylation should be considered as a critical quality attribute for all monoclonal antibodies and needs to be monitored during routine production. As a question, this is too broad to be answered since it ignores the fact that 'glycosylation' is not a single attribute and that variations within the glycan profile of a given product can have differing impacts depending upon the identity of the specific glycans, their relative proportions, the class of the monoclonal antibody, the dose schedule, the route of administration and the intended patient population for the medicinal product. As such, it is necessary to assess the Fc effector-function and its postulated effect on safety and efficacy individually. Jiang et. al[4] propose a three-tiered ranking of the Fc effector-function potential, resulting in a routine monitoring for the highest potential group and a risk and data based approach for antibodies with moderate and low effector-function potential. Examples are given in Section 6.1.

3.2 General quality attributes

In addition to quality attributes that may arise from discrete and measurable variations in the primary, secondary and tertiary structure of a given protein, there are a number of other types of attributes that may be relevant to the safety and efficacy of the product and that must therefore be considered as part of the criticality assessment.

Potency and Quantity

Opinion is divided as to whether potency should be considered to be critical quality attributes since they represent the combined effects of multiple molecular attributes rather than being linked to

[3] A-Mab: a case study in Bioprocess development, CASSS- ISPE.
[4] Jiang, Xu-Rong et al, Nature Reviews Drug Discovery, 10, February 2011, (101-110)

variability in a specific attribute. This distinction is somewhat irrelevant however since potency and quantity will always have to be controlled since the achievement of the desired efficacy and safety profile of any given product must always be assumed to depend to a significant extent upon the potency and quantity of the product administered. Potency should be determined in assays that reflect the proposed mechanism of action for the product, and where a multifunctional mechanism of action is known or suspected for a particular product, product potency in respect of each biological mechanism related to efficacy should be appropriately controlled. A criticality score is assigned taking efficacy and safety into consideration.

Quantity may be expressed as potency or absolute amount of material. In both cases, the dose is based on the measurement and thereby directly related to the efficacy of the product.

Process Related Impurities

Process related impurities are substances introduced to the process during manufacturing. These can be cell-derived such as host cell DNA or proteins (HCP), media components, or leachables from chromatographic resins or buffer components. They should be removed to safe levels. As described in Section 3.1 for process related impurities, a criticality score is assigned in order to define how and where it needs to be controlled. As discussed in later sections of the document, the application of risk-based approaches to the definition of the control strategy may justify that certain impurities are controlled in-process or through demonstrated validation of clearance by the process, rather than on the final specification.

Prior knowledge can be applied on known process related impurities, such as DNA, HCP and media components for the same type of proteins used for similar indications.

Contaminants

Contaminants are substances inadvertently added to the process. These are generally adventitious agents, bioburden and their metabolic products such as endotoxins, or mycoplasmas. Contaminants are unwanted and should generally be considered to be critical quality attributes. It may be acceptable to test for certain contaminants in-process, but due to their adventitious nature and possible consequences for product safety if present at unacceptable levels, it would not generally be appropriate to claim freedom from adventitious agents based on process capability alone. Special requirements do apply for viruses and the confirmation of their absence.

General Tests / Quality

In addition to tests for specific attributes such as purity and impurities (both product and process related), and non-specific or collective molecular attributes such as potency and quantity, the quality of drug substances and products also depends on the physical description, such as pH, osmolality, excipient concentration, sub-visible particles, and appearance. Reference should be made to the monographs of the appropriate pharmacopoeias for information regarding the testing that may be required for specific product types.

A particularly challenging area for criticality assessment is the impact of sub-visible and visible particles. Whilst there are clearly defined compendia requirements related to the control of foreign/adventitious particles in injectable products, the assessment of criticality and definition of appropriate controls for sub-visible and visible intrinsic proteinaceous particles remains an evolving area. Available regulatory guidance and precedents indicate that applicants are expected to take reasonable measures to design their products in order to minimize to the extent possible the presence of protein particles in their products to be practically free from particles when examined under suitable conditions of visibility, as described by the European Pharmacopeia. However, certain

product types (e.g. mAbs) are inherently prone to aggregation and particle formation. Such particles may be transient and form at low levels by uncertain mechanisms, presenting significant challenges as regards detection, identification, characterization, criticality assessment and their potential impact on safety and efficacy for definition of an appropriate control strategy. In general, where a product has been shown to intrinsically contain sub-visible or visible protein particles, it will be necessary to demonstrate what those particles are and that they are appropriately controlled in line with the quality of lots used within clinical trials.

3.3 Weight of Evidence for Assessment of CQAs

When assessing the criticality of a particular attribute, it is important to consider all available sources of information in order to form a risk-based assessment of whether variability in that attribute is liable to impact product safety, efficacy and quality. The level of certainty in this assessment is an important aspect of designing the overall control strategy.

The most important source being the specifications used in the clinical trials and the data generated from the analyses performed on this material.

3.3.1 Published Data

Peer reviewed scientific publications can be useful in the risk assessment of a CQA, particularly when applied to products such as monoclonal antibodies. Literature references, company studies and experience can provide data to support the lack of impact of an attribute such as relative amount of C-terminal lysine on monoclonal antibody heavy chains or typical major glycan forms (G0, G1) on the pharmacokinetics properties of antibodies. However, caution needs to be exercised in extrapolation of published data to a specific product and the potential impact of differences in dose, route of administration, indication and patient populations should be considered when determining the certainty of any assessment.

3.3.2 Product Specific, Nonclinical and Clinical Data

Preclinical studies in animals or cell-based assays can help assess how changes in quality attributes may affect clinical outcome and can also provide a basis for setting preliminary or final limits on those attributes. However, the species specificity of certain products may make the interpretation of non-clinical data challenging since the receptor target and binding characteristics may differ between species and the administration of human proteins in animals may result in significant immune response and the neutralization of the product. Clinical studies are generally designed in order to assess specific clinical endpoints and not with the intent of investigating the potential impact of product heterogeneity on those endpoints. However, useful information regarding the impact of particular attributes on the clinical safety and efficacy can nonetheless be obtained through retrospective analysis. The collection and analysis of plasma samples from clinical trials may be a particularly useful approach to establishing the in-vivo fate of the molecule, and may inform the criticality analysis. For instance, if a molecule is shown to be extensively deamidated within a short time period after administration, then deamidation may be viewed as a less critical attribute since patients have been extensively exposed to deamidated product at levels well above that typically seen in the product prior to administration.

3.3.3 Non-Product Specific, Non-Clinical and Clinical Data

Non-product specific nonclinical and clinical data can, in some cases, be used to establish the criticality of attributes such as process-related impurities. These data can provide patient exposure information and toxicology data that support the safety margins for dosing of process-related impurities that may be common to a platform process, for example DNA and Protein A. It may also

be possible to apply non-product specific data to the assessment of criticality for molecular variants of the product, but the extrapolation of knowledge from one product to another would need to be carefully justified. A more acceptable approach may be based on the analysis of experience across a closely related class of products, such as IgG1 or IgG2 monoclonal antibodies, where a significant number of examples of products, with a high degree of sequence and structural homology, have been commercialized. However, even in these circumstances, there is not currently 100% agreement that any of the known quality attributes can be automatically considered to be of low criticality in all instances.

4 Specifications

A complete discussion of all of the factors required to define the control strategy is beyond the scope of this document. Instead, the control strategy is discussed from the perspective of identifying which tests should be routinely conducted (testing strategy) and what criteria should be set to ensure product quality and consistency. The basis for the specifications for a marketed product is the specifications used during clinical development, for release of material for phase 1, phase 2 and phase 3 and the information gathered on the material and the methods used in the testing of this material. A systematic, science and risk-based approach of defining a specification is independent of the extent to which the principles of enhanced approach as described in ICH Q8 and Q11 have been followed in the development of the product. A control strategy is required for every product and the basic considerations for how it is defined are the same irrespective of the approach used to build the knowledge base. However, it logically follows that the application of an enhanced approach leading to improved understanding of the interactions between inputs (production cell line, raw materials etc), process parameters and process outputs (product quality) results in a more targeted approach to testing, and that less final release testing may be required than for a product developed under the 'traditional' approach.

4.1 Testing Strategy

The approach described here defines a control strategy based on a science and risk-based approach in order to obtain a product that ensures safety and efficacy for the patients. The goal of the Testing Strategy is to define a composite set of analytical tests and acceptance criteria to produce drug substance and drug product with the defined product quality attribute profile. The testing strategy is defined to confirm the quality of the drug substance and drug product rather than to establish full characterization, and should focus on those molecular and biological characteristics (CQAs) considered relevant to the safety and efficacy of the product (Q6B). The Testing Strategy defines whether an attribute is controlled by testing or by other means. In addition, based on criticality and process performance the Testing Strategy defines where each attribute is tested, e.g. as an in process control (IPC), as a release test on drug substance and/or drug product or on stability testing, or if no routine testing is required. Furthermore, it is relevant to draw a distinction between those parameters and acceptance criteria that are registered within the marketing authorization and those that the manufacturer may apply within the context of the QMS for the monitoring of process performance through trending, or as part of comparability analysis following process changes. In general, each identified CQA must have predefined acceptance range and it has to be ensured that each CQA is within this range. Depending on its criticality and on the capability of the validated process to control the respective CQA the following decision process can be applied to identify the testing strategy. In consequence, this approach differs from the approach described in ICH Q6B, where specifications are predominantly set based on process experience and focus mainly on ensuring process consistency to obtain a product of sufficiently high quality.

4.2 Selection of Analytical Methods

In general, each attribute to be tested as defined in the testing strategy has to be controlled by an adequate analytical method. Methods should be selected to control all types of attributes described in Section 3 of this document and detailed in ICH Q6B. A strategic approach should be applied in the selection of analytical methods to ensure that the quality attributes can correctly be measured. It should be recognized that one method might measure a collective impact, e.g. charge heterogeneity can be the result of a change in deamidation/ iso aspartic acid content; level of sialic acid; as well as from the level of Lysine processed for an antibody.

The methods selected for routine control must be robust enough so that a change in a result can be related to a change in the level of an attribute, distinguished from a variation in the analytical method. Ideally the method should be stability indicating to detect those changes in the molecule over time. The methods can be built on the technologies used for characterization of molecular and biological attributes of the product that are further developed to methods suitable for routine analysis.

In a traditional approach, as defined in ICH Q8(R2), the methods used in the control strategy are typically identified empirically and are based on the development experience, platform knowledge, the production history and the process validation. The methods cover the most relevant quality attributes linked to safety and efficacy and are supplemented by a set of methods demonstrating process consistency. In the course of method validation it has to be demonstrated, that the methods are capable of detecting the respective attribute within its predefined acceptance criteria. A set of methods should also be developed to support extended characterization of DS and DP processes.

In an enhanced approach, particular attention has to be given to the selection and development of the methods. During extended characterization and validation of the final DS and DP processes, methods have to be used that are capable of assessing the individual CQAs in order to be able to evaluate multivariate experiments and link the CQAs to the respective process input parameters. These methods do not require full assay validation but should be scientifically sound with specificity and accuracy being of particular interest, to ensure accurate measurement without interference by the sample matrix. Often the methods employed are developed in parallel with supporting process development and design studies.

The methods applied for routine testing should be validated according to relevant guidelines (ICH Q2R) once the respective testing strategies (in process-, end product-, and stability testing) have been defined, and the acceptance criteria for method validation can be assigned. Compendial methods, as described in relevant pharmacopoeias, must be assessed as to whether they are suitable for a specific product.

4.3 Testing Frequency

The inclusion of a test on a specification is one approach to ensure that each lot manufactured of the product will meet that specification. However, depending on the extent of process understanding, by implementation of procedural controls and in-process controls etc., it may be possible to justify alternate approaches to end product testing for certain attributes on every lot. Such approaches have been commonly applied in the field of small molecules for some time, e.g. parametric release for sterility of terminally sterilized products, or skip lot testing for heavy metals. Such approaches have generally been applied less frequently to biotech products in the past, but the application of enhanced process and product development approaches may support their use more commonly in the future. Different alternatives of testing are listed below:

- **End Product Controls / Release testing**

Conformance testing of each batch against predefined specifications for release as a part of the CoA.

- **Real Time Release Testing (RTRT)**

The ability to evaluate and ensure the quality of in-process and/or final product based on process data, which typically include a valid combination of measured material attributes and process controls (ICH Q8(R2)), such as safety testing on harvested bulk, testing for product impurities prior to addition of stabilizers, etc. The result for testing that is performed as in process controls in lieu of end product testing can be reported at the certificate of analysis with reference to the place of testing, or be reported in the executed batch record.

- **Parametric release**

"A system of release that gives the assurance that the product is of the intended quality based on information collected during the manufacturing process and on the compliance with specific GMP requirements related to Parametric release", as defined by European Organization for Quality. It is recognized that a comprehensive set of in-process tests and controls may provide greater assurance of the finished product meeting specification than finished product testing.

- **"Skip lot" (periodic) testing**

Performance of specified tests at release on preselected batches and/or at predetermined intervals, rather than on a batch-to-batch basis, with the understanding that batches not tested would meet all acceptance criteria established for that product if tested. Such tests are also referred to as Periodic Quality Indicator Tests (PQIT).

- **Testing as part of post approval commitment (no fixed intervals)**

In addition, testing of certain CQAs might also be performed as part of life cycle management. During filing events should be defined which require testing of certain CQAs that are not normally tested for: e.g. testing as part of an investigation or during troubleshooting following an unplanned deviation during manufacturing; as part of comparability after planned process changes or as part of a process transfer. This testing strategy is typically defined for CQAs which are not tested for as the process has been proven to be robust in removing them or which do not occur when the process is run within its approved operating ranges.

4.4 Acceptance Criteria

The control strategy should ensure that each CQA is maintained within a defined range that yields a product of acceptable quality from the perspective of clinical safety and efficacy. However, it may prove challenging to establish a direct quantitative relationship between the level of a specific attribute and the clinical safety and efficacy of the product. A more holistic approach is generally indicated in which the available information regarding the impact of the attribute on safety and efficacy is considered.

4.4.1 Clinically Qualified and Clinically Acceptable Ranges

According to ICH Q6B, emphasis is placed on linking the quality of the commercial lots to the quality of lots used in non-clinical and/or clinical studies. However, this should not be misinterpreted as using solely the range of the attribute that was present in clinical batches (generally defined as the range of the attribute that was present in clinical batches). Whilst some extrapolation may be appropriate and clinical experience should always be considered in the context of indication, patient

population and dose, and be scientifically sound. The range of CQA present in materials used in clinical trials is generally taken as the starting point for defining specification acceptance criteria. However, considering that it is not typically the case that pivotal trials include batches of widely varying product quality, it must be concluded that the, 'clinically acceptable range', i.e. the range of the attribute that results in acceptable safety and efficacy, may be different from the 'clinically qualified range'. The acceptance criteria are therefore often supported by a statistical analysis of the range of the materials used in clinical studies.

In addition to the batches used in clinical studies, the following should be considered in order to define the acceptance ranges:

- Data from on-clinical studies
- Magnitude of impact to PK and bioactivity
- Excipient ranges; justified by studies at range extremes
- Leverage "platform" and prior knowledge

To define a clinically acceptable range that is wider than the clinically qualified range, a number of approaches can be applied. The first is to conduct an assessment of the criticality of the attribute and eventual correlation of safety/efficacy and the CQA level to justify wider numerical acceptance ranges than the clinical tested ranges. In this is feasible for CQAs impacting efficacy and PK, where non-clinical or in-vitro data can be generated e.g. by showing that CQA does not affect for example the FcRN binding or by showing that the potency assay is not effected in the desired range, or by an *in vivo* study in an animal model to show comparable PK data. However, for CQAs impacting safety in a dose depending way, it will be harder to justify wider clinical acceptance ranges. One option can be to relate further investigations of the CQA the range of values determined for the lots used in the clinical program. If patients have been treated with higher doses in the phase 2 program, resulting in an exposure of a higher level of this attribute than that intended for the final product, an expansion of the clinically acceptable range based on this scaling of the doses may be justified, and a higher range for a specific CQA would be accepted. It is expected that a comprehensive criticality assessment has already been performed and that available knowledge can be re-considered.

Finally, prior knowledge or "platform knowledge" can be utilized under some circumstances. By prior knowledge, is acceptance criteria extended beyond ranges proven safe and efficacious in clinical and non-clinical studies of one product, based on experience from other similar products. For example if related molecules with similar mechanism of action can be applied as precedents, e.g. host cell proteins, aggregates, c-terminal Lysine, FC-receptor activity, (depending on whether or not FC activity of related to mechanism of action etc).

The process capability must also be considered in the justification of the clinically qualified ranges. In the case the clinically qualified range for an attribute related to a molecular characteristic is wider than the range of predicted process performance the specifications would normally be set at the clinically qualified range, e.g. for aggregates. However, in other circumstances the quality attributes, e.g. process related impurities, and in particular those that may be considered generally undesirable to be present in the product (e.g. arising from the process or raw materials), it may be expected that the acceptance criteria are set based on process capability. When following the enhanced approach, a design space can be defined also for certain process related impurities (e.g. HCP) and in this case the acceptable range may be wider than the expected process capability to accommodate movement within the design space.

The following Table 2 assembles some typical situations that can arise in defining the clinical acceptable ranges.

Table 2. Example of utilization of prior knowledge to define clinically acceptable ranges.

Experience	Aggregates	HCP	Deamidation
- Clinical experience	4 %	20 ppm	5 %
- Predicted process capability	3 %	5 ppm	10 %
- Prior knowledge -Phase 2 material* -Similar products/indications	 5% NA	 20 ppm 10 ppm	In vitro and animal data show: no impact PK
Proposed specifications	**< 5 %**	**< 10 ppm****	**10 %**

* Phase 2 dose was 2x to be marketed dose, ** HCP level will be set lower than tested during clinical trials.

4.4.2 Process Capability

In addition to clinical experience, process capability and the availability of data on process performance should be assessed as part of establishing the intended commercial control strategy. At the time of submitting the authorization application, in particular for orphan products, or when products filed on the basis of Phase 2 data, very limited amount of data may be available for establishing specification criteria. The clinical range may be very narrow, driven by the fact that only a small number of lots supported the clinical program with limited variation in normal process and raw material etc. In addition, the predicted variability of the commercial manufacturing process may be wide, driven by the inherent uncertainty of a small dataset. As a consequence, the clinically qualified range may not encompass the range of predicted process performance. In such circumstances, restricting the specification acceptance criteria to the clinically qualified range may result in a high probability of rejecting lots. Acceptance criteria wider than the clinically qualified range may be justified as the basis for the specification, to be adjusted if needed when more data is available.

Various statistical approaches can be used in order to assess the variability in the data set for particular attributes and to predict the potential future variability of the process. However, as noted above, the precise statistical approach is not of primary importance since the starting point for specification setting is more generally taken to be the clinically qualified range. The process capability however, should be used to determine internal action limits, i.e. in order to monitor the process performance as part of the internal QMS. For this purpose the 95/99% tolerance intervals might be used. See example in Section 6.3.

4.5 Types of Acceptance Criteria

The types of acceptance criteria are defined as qualitative or quantitative, and then further into subtypes according to ICH Q6B. The type of acceptance criteria chosen should be suitable for the quality attribute(s) they control. Quantitative reporting by numerical limits or ranges should be used whenever possible. When a qualitative acceptance criterion e.g. "Comparable to Reference Material" is applied, it is important to specify the criteria on which parameters that should be compared. For certain product specific analytical procedures it may be appropriate to set both a qualitative and a quantitative acceptance criterion. This could be relevant for an impurity assay where the batch data and analytical validation justify a quantitative limit such as "NLT 85%". To ensure the control of the remaining 15 a qualitative criterion such as "Comparable to Reference Material" can be included.

4.6 Drug Substance vs. Drug Product and Shelf-life Specifications

The acceptance criteria for a specific quality attribute may be different for drug product and for drug substance. If storage of drug substance and / or the handling during fill and finish of a drug product

result in a change in a level of an attribute, the acceptance criteria for drug product acceptance criteria should reflect this change.

The acceptance criteria for shelf-life specifications are defining the acceptable levels of specific attributes at the time of expiry. Shelf-life specifications may be necessary in addition to release specifications for some parameters, such as for degradation products and potency if a change is observed for these parameters during storage. Shelf-life acceptance criteria may apply to either or both, drug substance and drug product specifications.

For products, which are reconstituted for use, stability data should be generated during the in-use period to ensure that the reconstituted product meets the limits of the shelf-life specifications. In-use specifications may be justified, e.g. for products in multi-dose containers, if a slightly increased level of degradation is observed.

4.7 Pharmacopoeial Requirements

The European Pharmacopoeia contain requirements for acceptance criteria for general tests such as for sterility, bacterial endotoxins, microbial limits, volume in container, uniformity of dosage units and particulate matter. The pharmacopoeia may also contain monographs for specified biotechnological and biological products that set the acceptance criteria for analyses of these products. Whenever a pharmacopoeia-defined acceptance criterion is available this should in general be adhered to, unless otherwise justified. However, the principles for setting specifications to reflect process capability for internal control, and for patient safety and exposure, also apply for compendial methods.

5 Acknowledgements

Concept paper focus group and main authors are:

Karin Sewerin, consultant for MedImmune
Christoph Lindenthal, Roche
Simon Hotchin, Amgen

The paper was written in collaboration with experts in setting specifications, reviewers and others from the EBE Biological Manufacturing Group member companies that contributed and supported the preparation of this document.

6 Case Descriptions

These cases are example provided by companies contributing to this paper, and are not meant to represent direct applications of the principles described herein.

6.1 Example for defining a control strategy

In order to determine an overall control strategy for a certain Quality Attribute (CQA) the following simplified approach was used:

1. The criticality of the Attribute is assessed using a risk ranking and filtering tool based on prior knowledge, literature data and in vitro studies (e.g. forced degradation, structure function characterization, PK Studies, binding studies to different receptors, Bioassay data)[5]. See Section 3 of this document.
2. Acceptance ranges for each CQA are defined based on relevant in vitro, preclinical and clinical studies including those reported in literature. In order to evaluate potential interactions among the CQAs as well as additive effects a cumulative approach was used for bioactivity and pharmacokinetic.
3. The testing strategy was defined using a risk ranking and filtering tool, that takes into account the criticality of each attribute, the abundance of a CQA, the likelihood of a CQA to be formed when running the production process within its proven acceptable ranges (PAR) and during storage (stability impact) as well as process understanding (e.g. existence of a predictive Process models).

6.1.1 C-terminal Lysine Truncation

The example molecule comprises an Fc domain grafted from a human antibody heavy chain that is prone to C-terminal lysine truncation. Very high levels of truncation (91-99%) were observed in drug substance during manufacturing for Phase I, II and III clinical.

The terminal lysine truncation results in charge heterogeneity since the number of charged terminal lysine per molecule may be 0, 1 or 2.

Table 3 Clinical experience with C-terminal truncation.

Clinical Phase	n	Mean ± SD (%)	Min-Max (%)
I	14	93 ± 1	91-94
IIa	85	98 ± 1	95-99
IIb	10	97 ± 1	96-98

Truncation of the C-terminal lysine of antibodies due to carboxypeptidase[6] activity during cell culture is well known, sometimes leading to completely truncated molecules in the drug substance.

[5] A-Mab: a case study in Bioprocess development, CASSS- ISPE

[6] Harris R. J. Heterogeneity of recombinant antibodies : linking structure to function. Dev. Biol. (Basel) 2005; 122: 117-127

A number of scientific publications suggest that C-terminal lysine truncation has no impact on biological activity, PK/PD, immunogenicity and safety:

- Natural plasma-derived immunoglobulins lack the C-terminal lysine residues[7]. Their long half-life and activity of extracellular human carboxypeptidases in plasma strongly support that C-terminal truncation does not induce efficacy, PK/PD, immunogenicity or safety issues. It is published[8] that monoclonal antibodies and Fc-fusion proteins charge variants K2 and K1 are promptly processed to truncated form K0 when the product is injected due to the presence of several plasma endogenous carboxypeptidases.

- The presence or absence of C-terminal lysine residues on antibody heavy chains is not likely to influence antigen binding, which is mediated by the distant complementarity-determining regions[9], nor are binding to the Fcγ receptor or complement C1q likely to be affected, as these involve residues in the hinge-CH2 region and CH2 domains, respectively.

- The CMC strategy forum held in January 2003 to discuss lot release and characterization test issues specific to MAbs concluded that C-terminal lysine heterogeneity of the heavy chain is a common post-translational modification present in IgGs which has been the subject of many discussions. Lysine truncation does not appear to adversely affect product potency or safety. However, taking a conservative approach potential C-terminal lysine effects on all antibodies cannot be ruled out. Thus, lysine truncation should be characterized, and process consistency should be demonstrated during product development. This is in line with the agency input that emphasized the importance of reporting the ranges of C-terminal lysine heterogeneity observed during the characterization and development phases[10].

In addition, internal *in vitro* studies have demonstrated that two batches of the drug substance with respectively 95% and 80% terminal lysine truncation had comparable biological activity.

The risk assessment of the quality attribute "C-terminal lysine truncation" based on *in vitro* preclinical studies and the published scientific literature is as follows:

A very low impact score (2) has been assigned for biological activity, PK/PD, immunogenicity and safety. An uncertainty score of 3 was assigned to biological activity as *in vitro* studies with the molecule demonstrated no impact on biological activity. As information on the impact of C-terminal lysine truncation on PK/PD, immunogenicity and safety comes from the published scientific literature, an uncertainty score of 5 was assigned.

The highest single risk score is 10 for PK/PD, immunogenicity and safety and this is therefore the overall risk score attributed to the quality attribute "C-terminal lysine truncation".

---

[7] Edelman GM, Cunningham BA, Gall WE, Gottlieb PD, Rutishauser U, Waxdal MJ. The covalent structure of an entire γG immunoglobulin molecule. PNAS 1969; (63): 78–85.
[8] Cai B. et al. C-terminal lysine processing of human immunoglobulin G2 heavy chain in vivo. Biotechnol. Bioeng. 2011; 108(2): 404-412
[9] Sondermann P, Oosthuizen V. Mediation and modulation of antibody function. Biochem Soc Trans 2002;30:481–6
[10] Schenerman M.A. et al. CMC strategy forum report. BioProcess International. February 2004, p42-52

6.1.2 CHO Host Cell Protein (CHOP)

Host cell proteins derived from the expression system used for manufacturing are unlikely to have any impact on biological activity and PK, based on the experience and wide usage of CHO cells as expression system. The risk of host cell proteins acting as an adjuvant to produce a significant anti-therapeutic immune response is considered to be low, given the route of administration (IV). This is supported by in house data with a similar molecule, containing a CHOP content of up to 250 ppm and showing no relevant immune reactions in patients. Therefore, criticality scores for bioactivity, pharmacokinetics and immunogenicity are low. However, a biologically active host cell protein could present a safety risk[11,12]. The expectation is that possible adverse events (such as anaphylactic reactions) would be manageable, and thus a moderate impact score for safety was assigned. Therefore, host cell proteins are considered a CQA, based on the risk score for safety, e.g. cytokine release giving rise to IRRs (infusion related reactions).

An acceptance limit for CHOP was set to < 60 ppm, based on clinical history (a 2 fold higher dose was used in pivotal trials) and prior knowledge with other CHO produced molecules.

Formal process validation showed that residual CHOP is consistently removed to < 5ppm when producing within the PAR.

Hence, no routine testing is required, but it is committed, that testing is performed in the course of a comparability exercise (e.g. process transfer, process changes) and to perform an assessment following a deviation during manufacturing, to decide whether additional CHOP testing needs to be performed.

6.1.3 High Molecular Weight Species (HMWS)

HMWS were shown to have lower (antibody dependent cellular cytotoxicity) ADCC activity in vitro potency assay and a significant higher binding to FcRN in binding studies and SPR studies. Therefore, they were assigned high criticality scores for bioactivity and pharmacokinetic.
Since they have been present up to 1.5 area% in clinical trials and no significant levels of ADA formation or IRR have been observed, the criticality score for immunogenicity and safety are rather low. However, based on Bioactivity and Pharmacokinetics HMWS are considered a CQA.

An acceptance range for DP was calculated to be 2.5 area%, extending the range that has been used in the clinical trials based on safety data from another similar molecule, taking into account the dosing and applying a safety factor.

Since HMWS can be formed at any time they were include in the routine testing strategy (Size Exclusion HPLC), and in the stability protocol (a significant increase of 0.2 area%/year in a liquid formulation is seen over time). Based on these data a shelf life limit of 2.5 area% was set, and release limit for DS and DP of 1.7 area%, taking the DP storage of 3 years into account. DS and DP release limit are identical, since process validation has shown no impact of the DP process and DS storage on HMW formation.

---

[11] Champion K, Madden H, Dougherty J, Shacter E. Defining your product profiles and maintaining control over it, Part 2. BioProcess Int. 2005: 52-57.
[12] Lupker J H. Residual host cell protein from continuous cell lines. Effect on the safety of protein pharmaceuticals. Developments in biological standardization 1998; (93): 61.

6.1.4 Level of Fucosylation

Since a linear correlation between level of non-fucosylated glycoforms and the ADCC activity (antibody dependent cytotoxicity) was shown, a high impact score for Bioactivity was assigned. A low impact score for Pharmacokinetic and Immunogenicity was assigned, based on in vitro studies (FcRN binding) and the absence of ADAs in clinical trials. A moderate score for safety was assigned, due to a potential correlation between high non-fucosylated species and AE could be seen.

The highest level of non-fucosylated species that has been used in clinical trials was used as upper limit in order to account for the safety aspect. In order to determine the lower acceptance limit a cumulative approach was taken. The acceptance limits for Bioactivity were set to 70 -130 % relative potency. Assuming that all CQAs affecting potency would be at their respective lower acceptance limit, the lower acceptance limit for potency of 70% should still be met. The lower acceptance limit for non-fucosylated species was therefore set to 40% representing a 20% decrease in potency and allowing for addition 10% potency decrease by other CQAs.

Non–fucosylated species will be monitored during DS release by HPLC, due to its high criticality, and lack of process understanding with respect to formation of this CQA. However, no testing at DP level or during stability is required, since the level of fucosylation does not change significantly.

6.2 WHO HACCP Risk Assessment Procedure

An approach in determination of the Critical Control Points using a HACCP risk analysis is provided in the flow chart below. Combined with an FMEA performed at the end of the process characterization and robustness validation that – for an enhanced approach – will generally lead to the establishment of a global, multistep or individual step design space and determines the process-related elements of the control strategy and the evaluation of the appropriate acceptance criteria will determine the overall control strategy. See Figure 2 below.

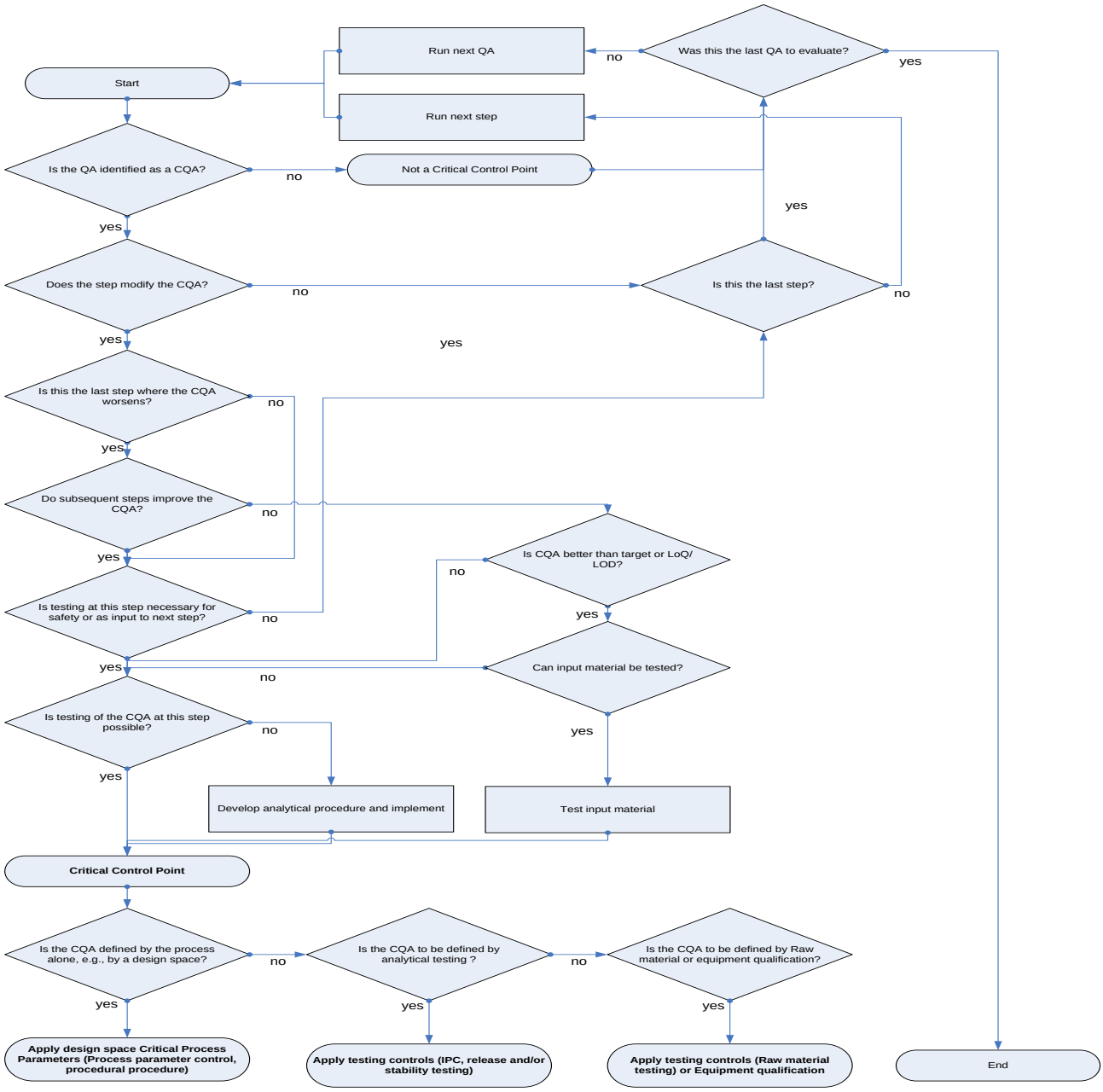

Figure 2 WHO HACCP for food products – adapted to pharmaceutical products

6.3 Statistical Approaches in setting Acceptance Criteria

For some acceptance criteria one can use a statistical approach to define the acceptable range. The statistical methods can be used to analyze release and stability data and thus assess how well a company will be able to fulfill given requirements.

Below are examples of statistical application on defining the acceptance criteria.

Process Capability

Process capability is the fundamental concept in the statistical approach to evaluating specification limits. The process performance index, Cpk, expresses the extent to which the actual output of a process meets the pre-determined specification limits (i.e. how well the process performs). The aim is to have Cpk-values larger than 1.0 and preferably larger than 1.33[13]. Cpk is calculated by the following formula:

$$C_{pk} = \min\left(\frac{URL - \bar{x}}{3s}, \frac{\bar{x} - LRL}{3s}\right),$$

where $\bar{x}$ is the mean of the process, s is the total standard deviation, including both process variation and analytical uncertainty. An illustration of different Cpk values can be seen in Figure 4.

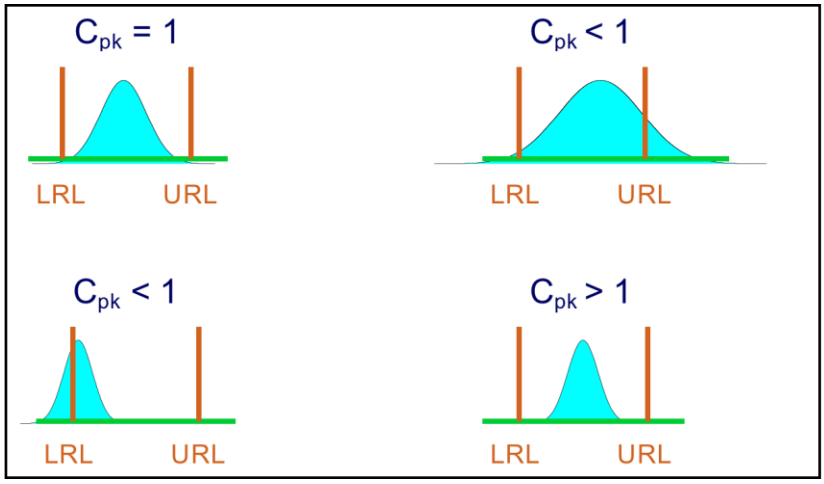

Figure 4 Illustration of Cpk values

The idea of finding a capability interval is the other way around, that is, to present a set of release limits (lower and upper), which correspond to a given value of Cpk, e.g. Cpk = 1.33 gives:

$$[LRL; URL] = [\bar{x} - 1.33*3s; \bar{x} + 1.33*3s]$$

In this case the release limits will correspond to ±4 standard deviations from the mean.

The process performance index is based on the assumption of normally distributed data. If data are described by other distributions, one should similarly ensure that e.g. at least 99.7% of the distribution is within the limits.

Statistical Uncertainty

[13] Montgomery, D. C. (2001) Introduction to Statistical Quality Control, 4th edition, Wiley, New York.

When limited representative data is available, the uncertainty of the estimate of mean and variance is not negligible. This can be accounted for by using a tolerance interval. A tolerance interval ensures a given process performance, say Cpk=1.0, with a high degree of confidence, typically 95%.

The tolerance interval may also be interpreted, as an interval that ensures with the given confidence, that a certain percentage of batches, say 99.7 %, will be within the limits.

The tolerance interval has the form

$$\bar{x} \pm ks$$

where the factor k depends on the desired process performance and the confidence level[14].

An alternative, but related concept is a prediction interval, which gives a specified degree of confidence, say 95%, that either a single observation or the mean of the next randomly selected sample is within the limits. The prediction interval does not ensure a given process performance, however.

Release limit

Batch release data are available for 13 batches at the time of filling. The mean impurity level is 0.21% and the standard deviation 0.05%. Based on these production data, the upper release is calculated as a statistical tolerance limit that provides high confidence (95%) that 99.7% of the future batches will be released; this gives an upper release limit of 0.42%.

Shelf life limits

For parameters that change over time different limits are needed for release and shelf life. The difference between the release and shelf life limits should account for the degradation during shipping, the degradation during long-term storage and a margin to account for the uncertainty in the estimated degradation, the analytical uncertainty and if relevant also batch-to-batch variation. These terms may be estimated from long-term and accelerated stability data using statistical methods.

When the shelf life limits for a given parameter is given, and the shelf life needs to be established, one can follow the approach outlined in ICH Q1E. This approach is well suited for situations where data are sparse, but it does not provide a high degree of assurance that future batches will remain within the shelf life limits.

Instead the principles from Allen[15] can be used to ensure with high confidence that a batch remains within the shelf life limits throughout the shelf life. The method links the release limit, the shelf life limit and the shelf life, and can be used to determine for instance the shelf life based on given release and shelf life limits, or the release limit for a given shelf life and shelf life limit. Notice that a shelf life granted based on extrapolation should be verified by additional long-term stability data.

Example

[14] ISO 16269-6:2005(E) Statistical Interpretation of data – Part 6: Determination of statistical tolerance intervals
[15] Allen, P. V. (1991) Determination of Release Limits: A General Methodology. Pharmaceutical Research, Vol 8, No. 9

The example below illustrates the above principles for establishing specification limits for an impurity in a new insulin product based on statistical considerations. The principle ensures with high confidence (95%) that all released batches will remain within the shelf life limits until expiry.

The specification limits are calculated "from the inside and out" starting from the production capability, and adding contributions from handling and storage as well as margins to account for analytical variability and statistical uncertainty in the estimated degradation rate.

In this derivation, the shelf life limits are based entirely on statistical considerations, and the final limits should be compared with the clinical requirements for the product to be safe and efficacious.

Allen's formula that links the shelf life and release limit is:

$$USL = URL + Handling + EAC_T + t_{0.95}(df)\sqrt{s_T^2 + s_A^2 / n}$$

Based on this the USL is calculated to 1.30 %. The different terms in the calculation are summarized below and illustrated graphically in Figure 5.

Term	*Explanation*	**Value**
URL	Upper Release Limit	0.42%
Handling	Degradation during handling	0.08%
EAC_T	Degradation during storage	0.66%
s_T	Standard error of EAC_T	0.05%
s_A	Intermediate Precision of Analytical method (standard deviation)	0.07%
n	Number of independent determinations at release	1
df	Degrees of freedom of combined variance estimate	67.7
$t_{0.95}$(df)	95%-quantile in the t-distribution	1.67
USL	Upper specification limit	1.30%

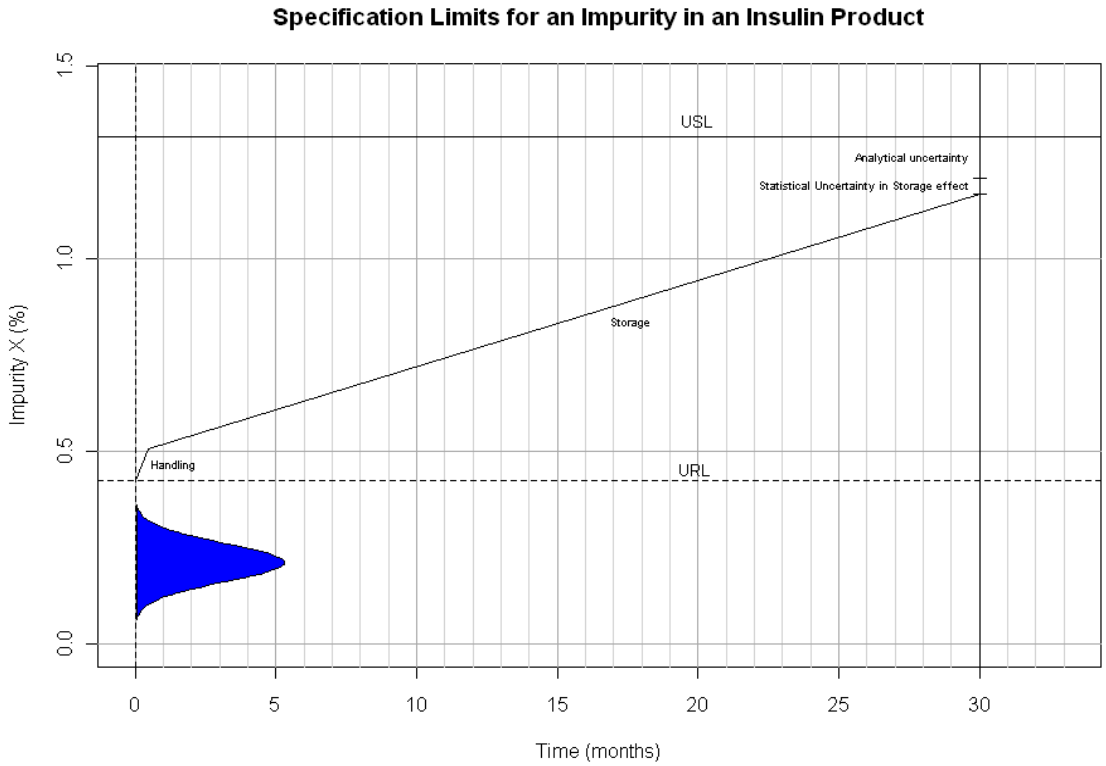

Figure 5 Illustration of the construction of release and shelf life limits based on release and stability data.

Difference between release limit and shelf life limit

After being released, the batch may experience higher temperatures during handling, e.g. shipping of the product. The degradation that will occur at different temperatures is estimated based on accelerated stability studies and Arrhenius calculations. Based on the allowed transportation time at different temperatures, the maximal degradation during handling is estimated to 0.08%.

The product may be stored for up to 30 months in refrigerator, during which a further 0.66% impurity is formed. This is estimated based on long-term stability data; the uncertainty of the estimate is 0.05% (std. err.).

The intermediate precision of the assay has a standard deviation of 0.07% (absolute).