It's all about consistency:

Change of Reference Standards for potency assessment in clinical phases

CMC Strategy Forum Europe

Basel, Oct. 2025

Hermann Beck,

Analyt. Dev. Bioassays (PTDA), F. Hoffmann-La Roche Ltd

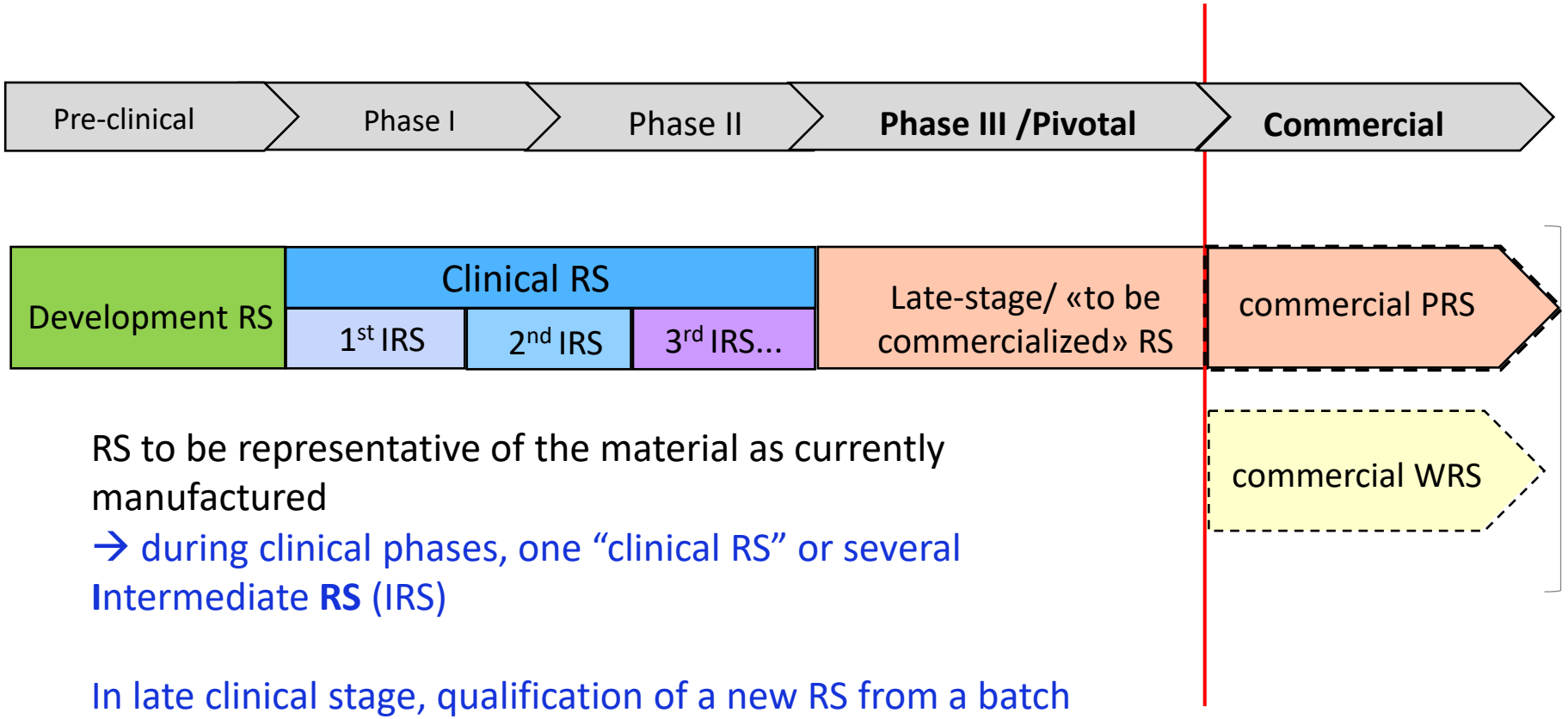
Overview

- Role of **R**eference **S**tandards (**RS**) in potency analysis
- Selection of RS DS lot
- Characterization
- Batch size, container, storage condition,....
- Stability monitoring
- Change of RS

Overview

- Role of **R**eference **S**tandards (**RS**) in potency analysis
- ~~Selection of RS DS lot~~
- ~~Characterization~~
- ~~Batch size, container, storage condition,....~~
- ~~Stability monitoring~~
- Change of RS
  - Testing of equivalence old vs. New RS
  - Determination of potency of new RS
  - Old and new RS with different potencies:
    - Consequences
    - Connecting the potencies of new and old RS

Change of RS during clinical phases



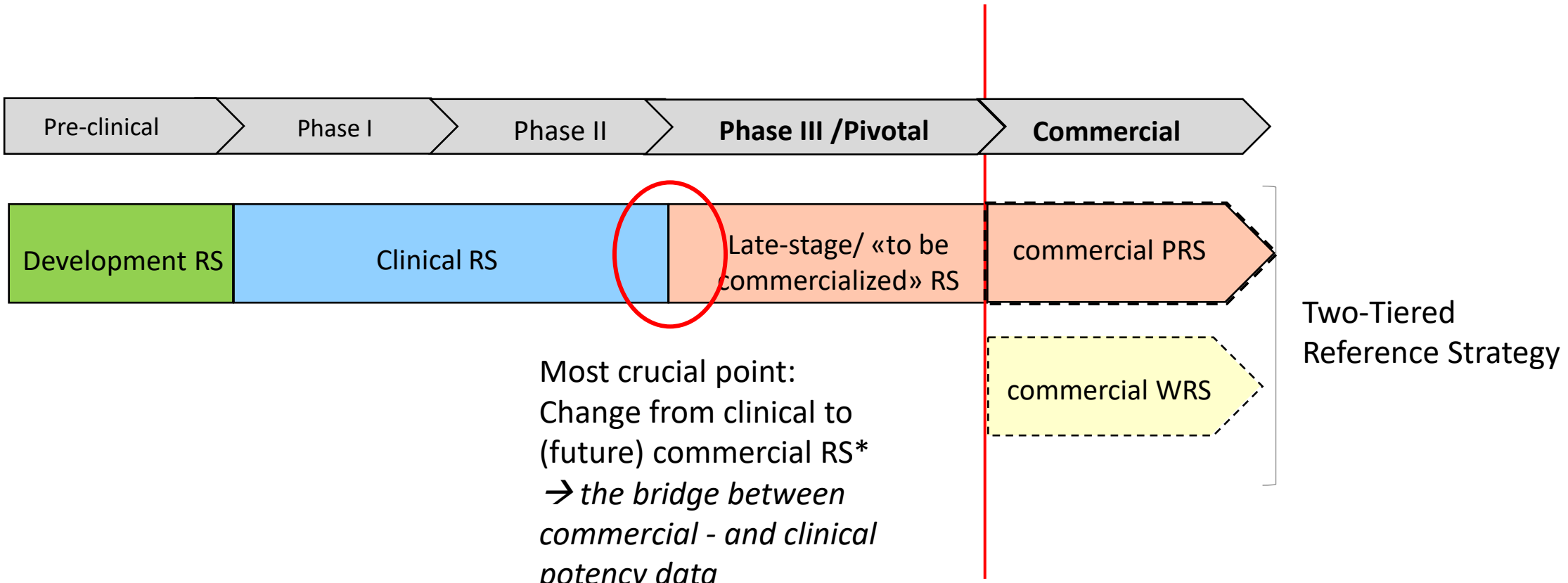
RS to be representative of the material as currently manufactured

→ during clinical phases, one “clinical RS” or several **I**ntermediate **RS** (IRS)

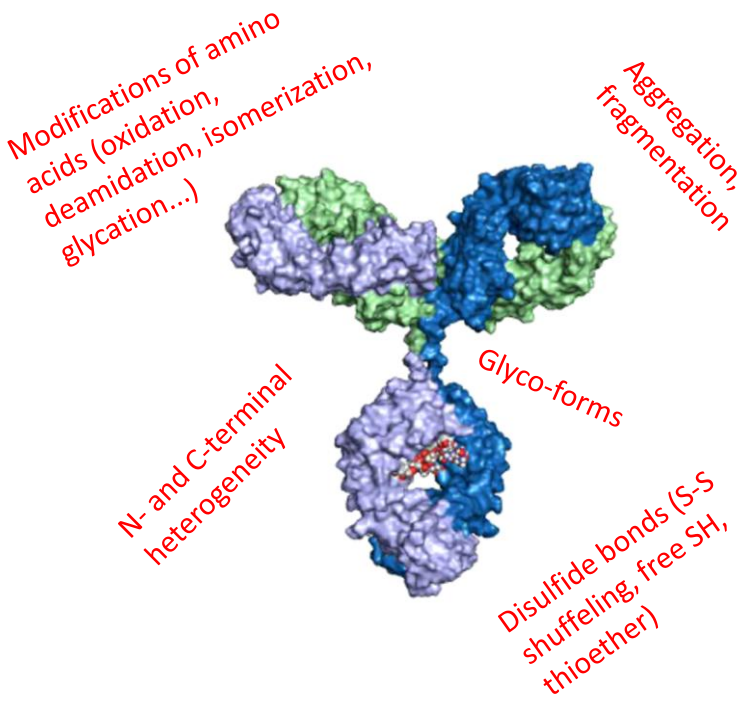
In late clinical stage, qualification of a new RS from a batch which is representative for the commercial material.

This «to be commercialized RS» will later become the commercial primary RS

Change of RS during clinical phases



The role of Reference Standard in potency determination



Secondary, tertiary, quarternary structure; PTMs, glyco-structure…

→ Biologicals are complex molecules

→ Biological Activity depends on multiple attributes

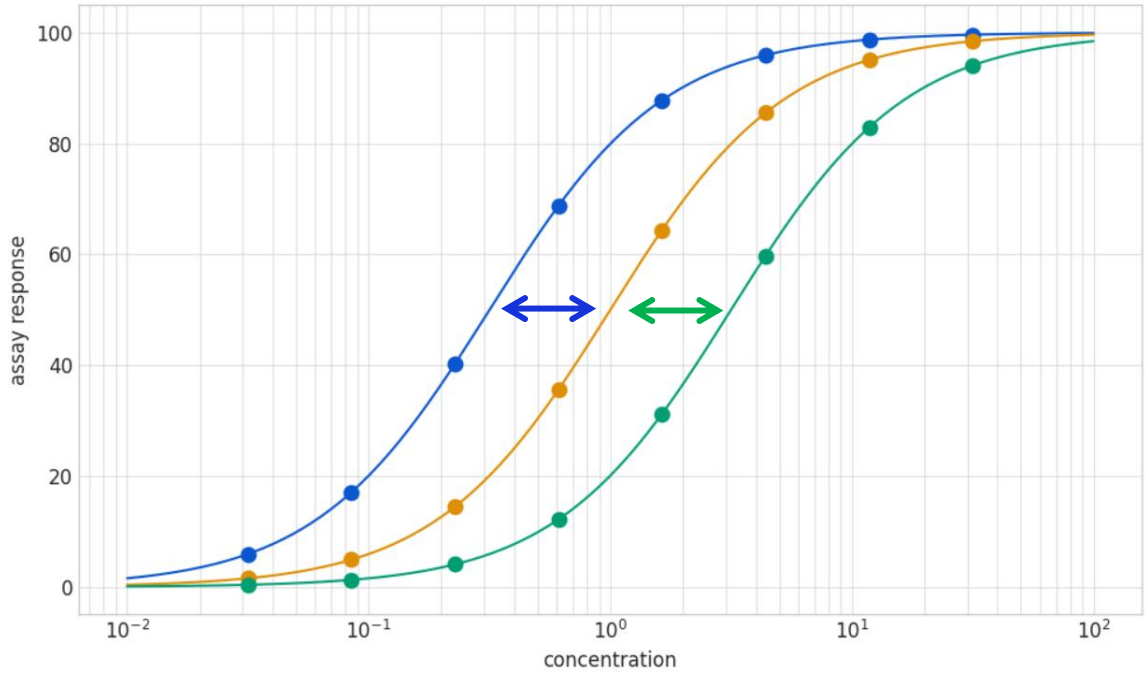
➔ **Potency Assays can monitor the interplay of multiple attributes simultaneously**

Such attributes vary depending on the manufacturing processes and thereby may affect a molecule's biological activity

Relative Potency Determination

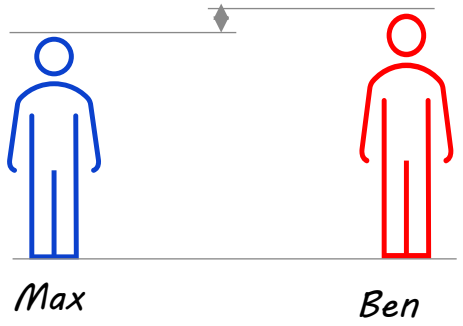
FDA, 21 CFR 600.3(s): potency/*"biological activity is the specific ability or capacity of the product to achieve a defined biological effect"*

→ the potency of a sample is manifested in the position of its dose-response curve on the concentration axis

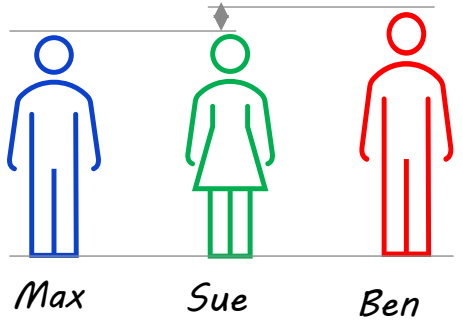


But there is no material with «100% ideal structure» available, and hence no expected absolut potency value exists.

→ For Biologics, Potency is always determined *Relative to* a *Reference* analyzed in the same assay



Ben is 10% taler than Max



Sue is as tal as Max → Sue has

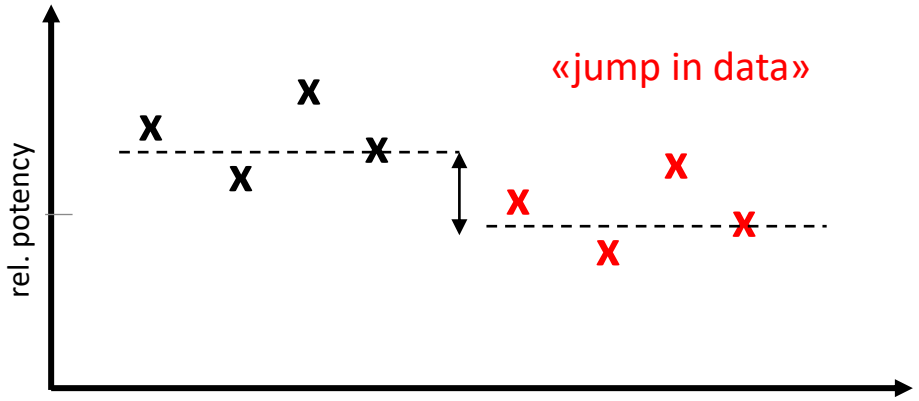
- 100% height rel. Max
- 90% height rel. Ben

→ Sues' relative body height depends on the reference

Why do we need to know new RS' potency?
Consequences of differences in potency between old and new RS

If old and new RS have a real or apparent difference in potency
→ material analyzed against both RS has diffferent potencies,- relative to reference

Example: new RS has a higher potency than old RS

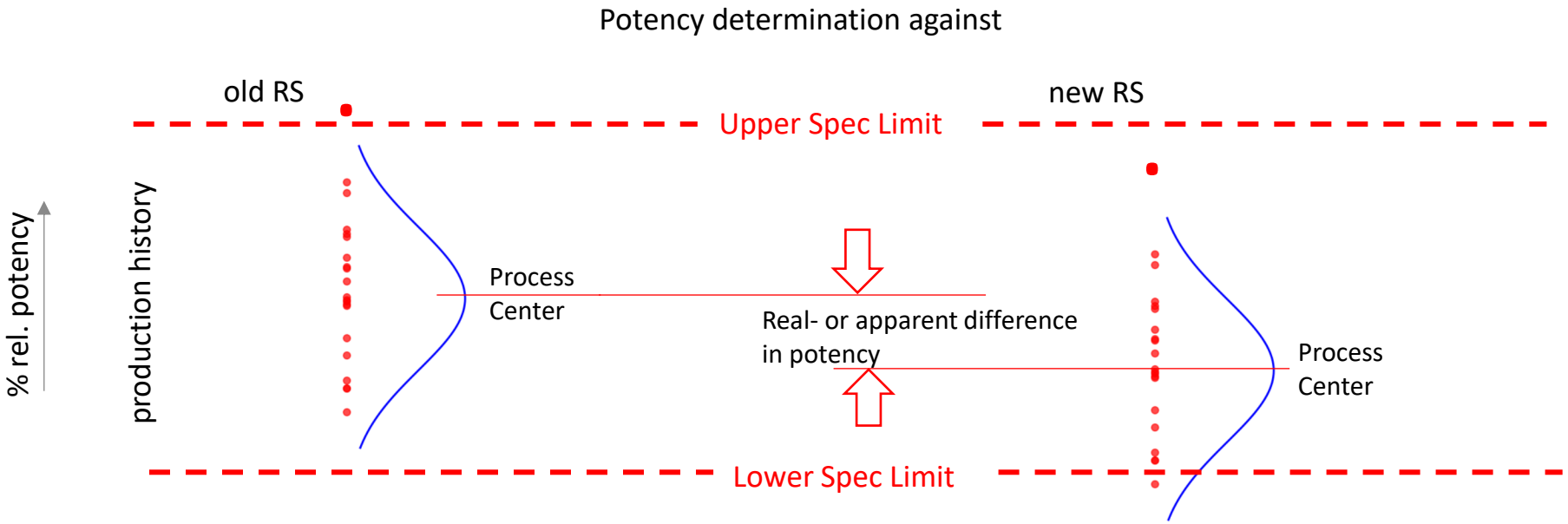


X: analyzed against «old» RS-potency
X analyzed agains «new» RS-potency

This may challenge release- and stab,-data evaluation; could indicate differences between samples where in reality there are none;….

To neutralize the «jump in data» as much as ~~possible~~ necessary, the potency of the new RS should be known ~~as good as possible~~ good-enough

How good do we need to know new RS' potency?
or: what could happen if the determination is too inaccurate?



During clinical phases, specification range and it's centerpoint* was defined based on results relative to the "old" clinical RS

** as average of batch-release results*

if the new RS actually or apparently has a different potency
→ Specs and analysis results become "decoupled":
- → the business-risk of iOOS increases
- → the patient-risk that "bad material comes into specs" increases

Determination of new RS' potency

There is not much on specific methodology, neither in HA guidances nor in the literature*.

«common sense»

Most potency assays show relatively high variabilty (5-15% SD)

- → analyze old and new RS side-by-side, i.e. in the same assay
- → enough replicates to approximate the true value as good as possible/necessary

Some approaches that we know are used are briefly presented in the following

**examples*

FDA Biosimilars Guidance

Faya et.al. *Potency assignment of biotherapeutic reference standards*. J. Pharm. Biomed. Anal. 191 (2020)

BEBPA White Paper: Reference Standards for Potency Assays

Determination of new RS' potency
A simple approach

Old and new RS were analyzed side-by-side, i.e. on same assay-plates, with fixed n *(e.g. 20)*

Average of n results = new RS' potency ≠ old RS

statement *'old and new RS have different potencies'* has far-reaching implications (shown later).
→ We only want to accept these implications if the difference is truly meaningful.

Determination of new RS' potency
A simple approach **with simple equivalence assessment**

Old and new RS were analyzed side-by-side, i.e. on same assay-plates, with fixed n *(e.g. 20)*

Difference of potencies new vs old RS is

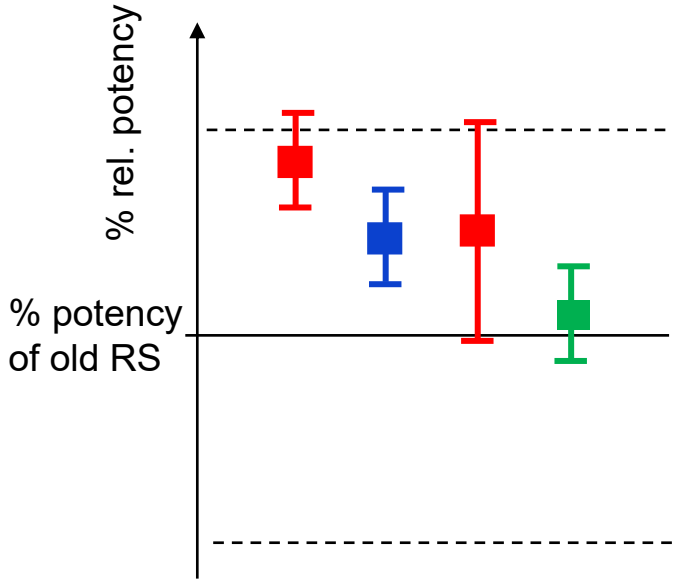
< 5%

"similar enough"
Old and new RS considered equi-potent

> 5%

meaningful different
New RS' potency = mean of the n = 20

Determination of new RS' potency
Approach based on confidence intervals*



– – – phase-appropriate acceptance range
(e.g. 80-120%, 90-110%; 95-105%)

average and CI of new RS' measurement results

Potency new RS = it's average or considered sufficiently similar → Potency new = potency old RS

i.e.: an equivalence consideration!

**Refer to e.g.*
FDA Biosimilars Guidance
Faya et.al. *Potency assignment of biotherapeutic reference standards*. J. Pharm. Biomed. Anal. 191 (2020)

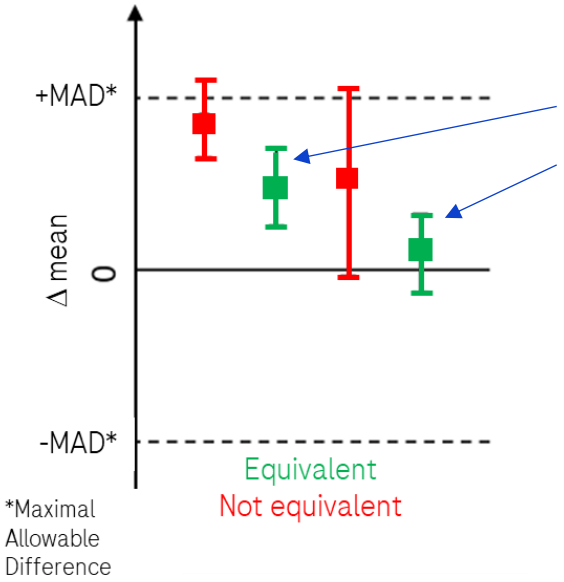
Statistical 2-step Approach

1) Test on equivalence of old and new RS

Approach

- Old and new RS were analyzed side-by-side, i.e. on same assay-plates ← *Number of required n (results) calculated in advance*
- Outlier test, CL 90%
- **TOST*** ← *CI 90%, MAD = 2SD (from PC-trending)*

Tuning the rigor of the test



Difference in mean with corresponding 90% CI is within acceptance range
→ same potency value as old RS can be assigned to new RS

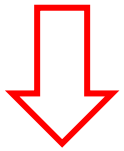
*TOST (Two One-Sided Tests) is a statistical procedure used to test for equivalence by determining if the observed effect falls within a predefined equivalence margin.

Comparison of old and new ReferenceStandards for clinical phases

A) Statistical 2-step Approach

1) Test on equivalence of old and new RS

old and new RS
have same potency



2) Determination of the potency of the new RS

By a statistical approach to calculate the potency of the RS with a targeted accuracy

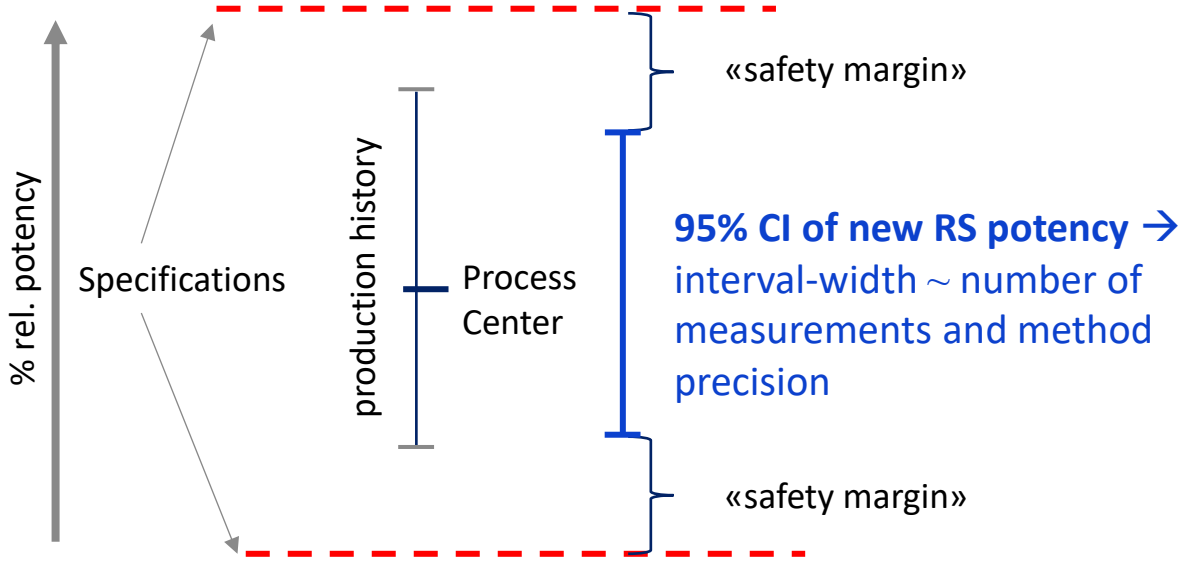
- Define a CI95 for the new RS potency considering the (anticipated) specification range and method variability
- **Calculate number of measurements n** that are required to reach the targeted CI95

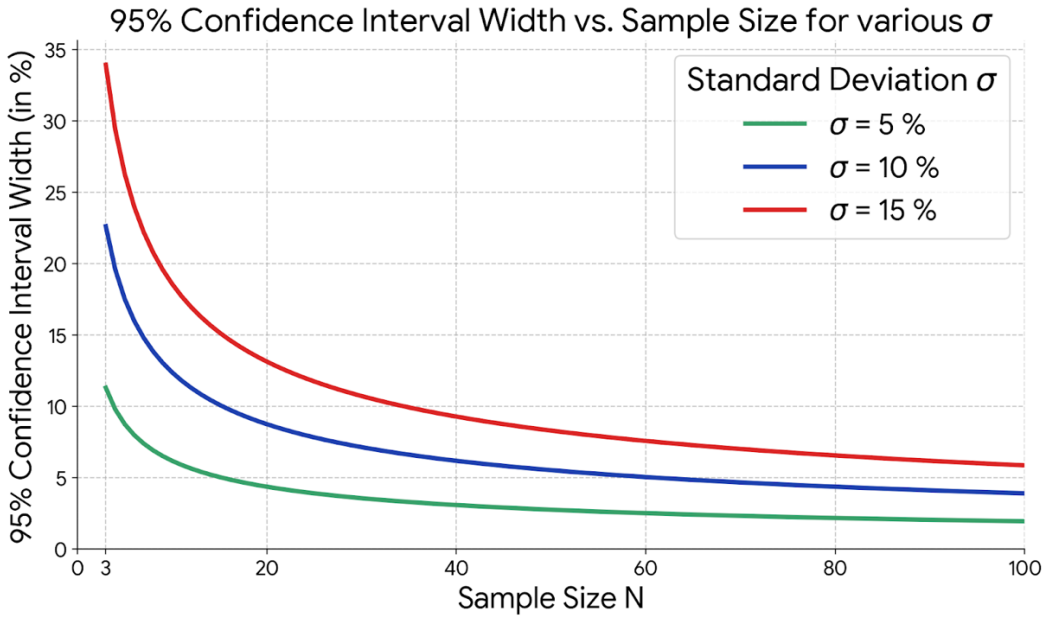
width of CI95 (n) ~ n and method precision

potency interval which contains the true RS potency with 95% certainty

Current strategy of Roche late-stage development Bioassay Europe

2) Determination of the potency of the new RS
Required accuracy





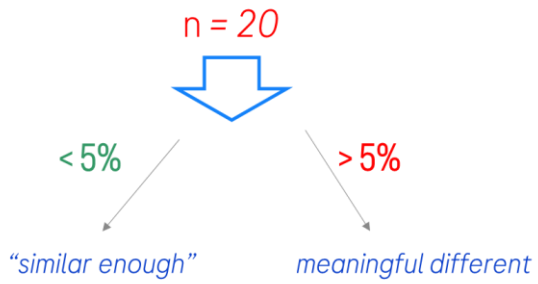
project-specific study design, reflecting

- Process- and analytical variability
- risk tolerance
- regulatory commitments

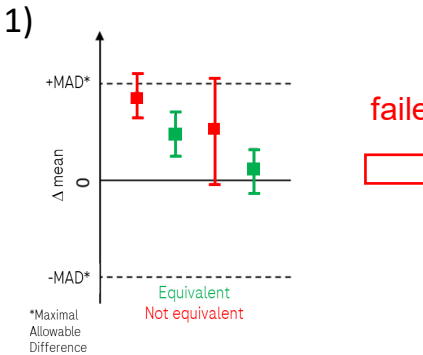
Comparison of old and new RS for clinical phases

Comparing the approaches

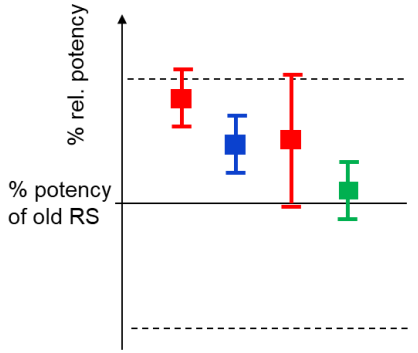
Minimal approach
e.g. fixed sample size (n), and acceptance range
→ no statistics; one fits all



Extensive approach: e.g. two-step process involving a statistical equivalence test (TOST) and project specific customized.



Everything in between
e.g. based on CIs



"similar enough"
→ Potency old RS = potency new RS

"meaningful different"
→ potency new RS = av. (n)

The choice of approach must, of course, be scientifically justified, but it also reflects a balance:
how much effort seems reasonable versus how much uncertainty is acceptable.

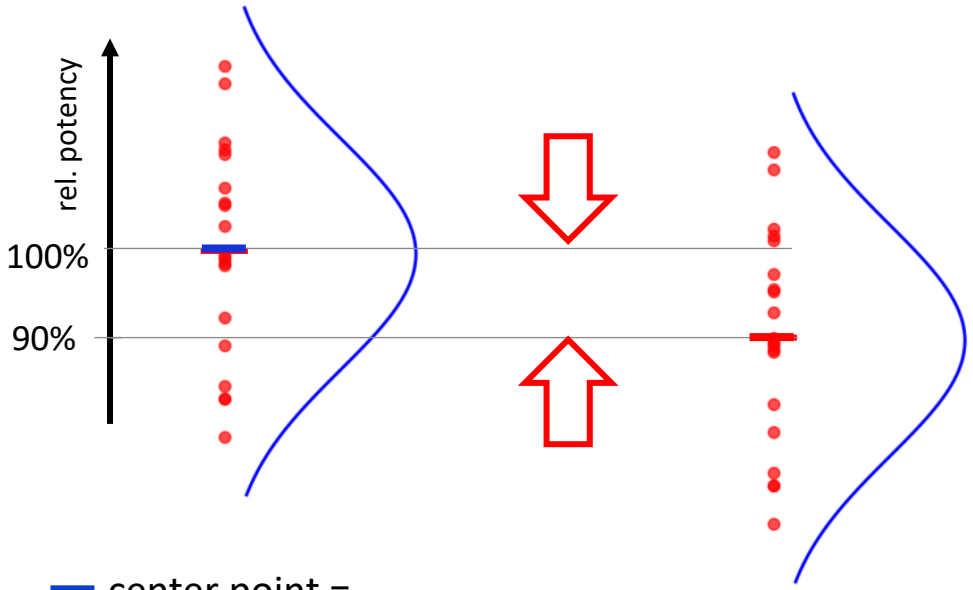
Old and new RS with different potencies
Consequences

both RS have their own potencies → material analyzed against both RS has diffferent potencies,- relative to reference

Example: average of batch release results

Material analyzed against clinical RS with 100% potency

Material analyzed against late-stage/ commercial RS with 110% potency



center point = 100% potency relative to clinical RS

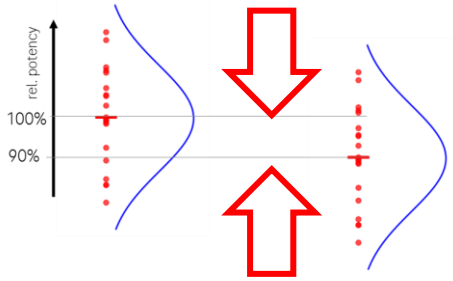
center point = 90% potency relative to late-stage/ commercial RS

Material with same potency has, in this example, 10% difference in potency depending on the RS used.
→ But it's just apparent,- the real, absolute potency didn't change!

Connecting the potencies of new and old RS

an RS potency of 100% is well suited as an idealized baseline.
Therefore

100% potency is assigned
to the clinical RS,
used in clinical phases



100% potency is assigned
to the PRS,
used in late-stage/
commercial phases

But if both RS do not have the same absolute potency:
gap between clinical and commercial data

Bridge the gap by a normalization factor

Recalculate historical data *

* of analysis against old RS

In our example:
normalization factor
= 100%/110% = 0.91

potency relative to new standard (%) = potency relative to old standard (%) x factor

FDA Biosimilars Guidance

"*A sponsor generally should not use a correction factor to account for any differences in, for example, potency or biological activity between reference materials. Under certain situations, the use of a small correction factor or factors may be considered if proposed and scientifically justified by the sponsor. If a sponsor intends to propose the use of a correction factor, discussion with the Agency during product development is recommended.*"

Connecting the potencies of new and old RS

Recalculate historical data *

* of analysis against old RS

potency relative to new standard (%) = potency relative to old standard (%) x factor

Mathematically very simple. The difficulty lies in adequate documentation!

2 options:

1) actual, «real» change of the data in (electronic) reporting systems (LIMS, GSMP...)

Pro: only one set of data existing
But: release and stability decisions were made based on data relative to the RS used at that time

→ approach would compromise GMP compliance & ALCOA(++) requirements

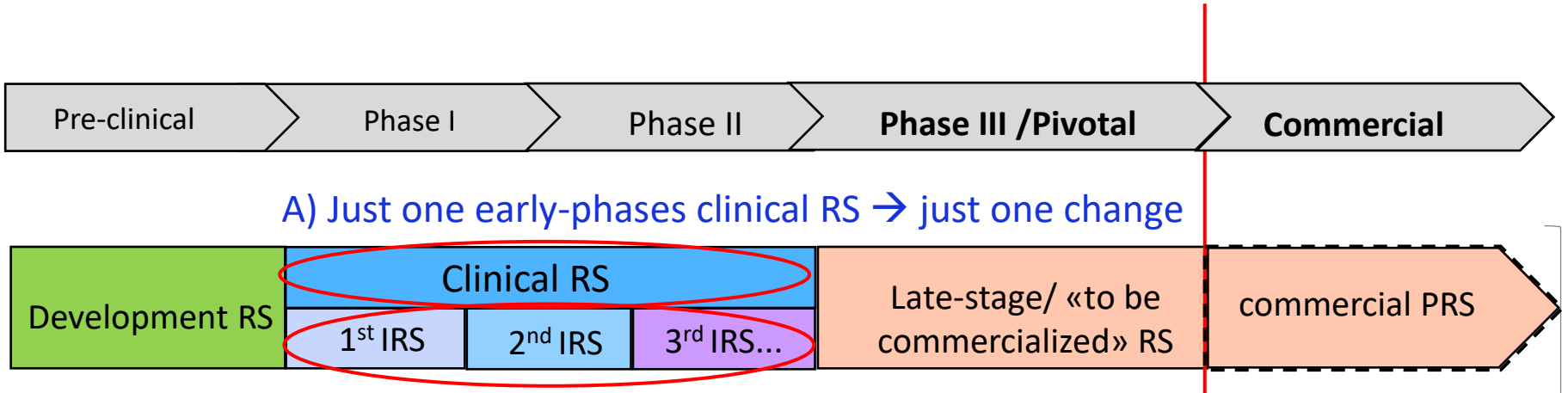
→ Not recommended

2) «the paper-solution»: leave the data as-is in (electronic) reporting systems, summarize and report recaluclated database in a document which is also used for HA submissions

Implications tbc:
2 sets of data,- 2 results for each sample measured during development

- Risks and potential compliance issues arising from having 2 data-sets (“QA-question”)
- Potential impact regarding already submitted data (“Reg.-question”)
- Risk of confusion: you can't tell from the % potency value whether it refers to the old or new RS

Change of clinical RS
How often?



B) establishment of Intermediate RS (IRS) for each clinical Phase and/or after each significant process change (e.g., a cell line switch)

→ RS most representative for current manufacturing process

but

- Many changes = many RS qualifications → if RS differ in potencies, data have to be connected many times
- If “chain qualifications“ applied: small differences which lead to an overall drift can still be overlooked

Summary

- Limited guidance on specific methods for RS qualification in regard to potency from HAs or in literature.
- Approach must balance scientific justification, reasonable effort, and acceptable uncertainty on RS potency.
- Differences between old and new RS potencies present unique challenges, often related to GDP/GMP rather than scientific aspects, which are neither addressed in HA guidance nor literature.

<u>Recommendations:</u>

- Conclude on differences in potency between RS only if they are relevant.
- Limit RS changes during clinical phases to essential cases (e.g., inability to demonstrate comparability pre-/post-process change), avoiding default changes at every phase or process change.

  Because

  - Frequent RS changes increase uncertainty in potency determination and complicate data linkage for release and stability datasets. While representativeness of RS improves with more changes, data quality does not necessarily benefit, but complexity could rise strongly.

Acknowledgements

Contributors

- *Adelheid Rohde*
- *Alex Briguet*
- *Felix Kepert*
- *Manuel Stoll*
- *Susanne Baroth*

Doing now what patients need next