sanofi

# Evaluation of statistical approaches for comparability and similarity studies

*Recommendations based on simulations*

Birgit Niederhaus – NCS Conference 2022

# Scientific Situations and Literature

- **Comparability pre-post change:**
  ~ 20 pre-change batches (ref)
  ~ 3 post-change batches (test)

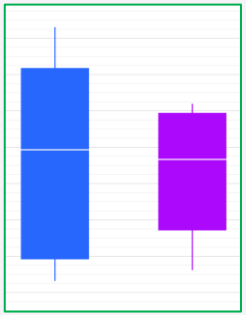
- **Biosimilarity:**
  ~ 15 originator batches (ref)
  ~ 10 biosimilar batches (test)

- **Scale-Down Model (SDM) Qualification :**
  ~ 6 large-scale batches (ref)
  ~ 6 small-scale batches (test)

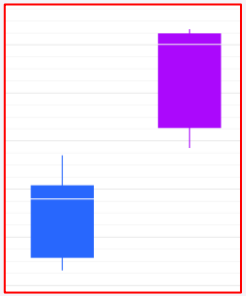
Literature: Draft Guidance from FDA; Reflection Paper and workshop presentations from EMA; ICH Q5E; BioPhorum White Paper, A3P draft Guidance; further papers on Biosimilarity and Scale Down Model Qualification

**Reference data: e.g. before the change**

**Test data: e.g. after the change**



Comparable?

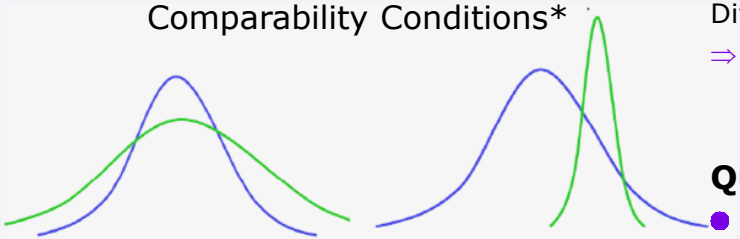


Not Comparable?

# Statistical Approaches

Difference of means
⇒ Equivalence approaches

Comparability Conditions*



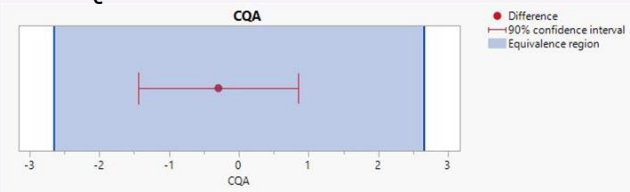
*Figures 2 and 3 from EMA reflection paper

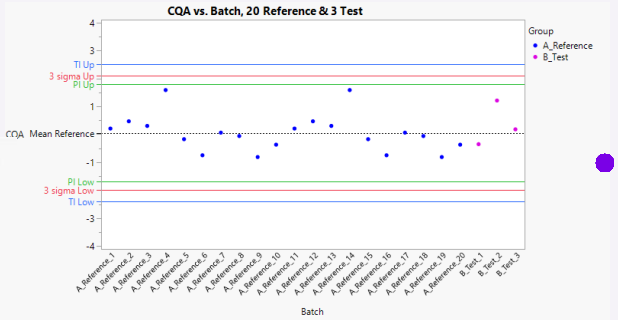
Difference in individual values
⇒ Interval approaches

$$[\hat{\mu}_R - X\hat{\sigma}_R,\ \hat{\mu}_R + X\hat{\sigma}_R]$$

**TOST**

- Parametric
  - Difference of means compared to equivalence margin:
    $H_0: |\mu_T - \mu_R| \geq \delta$ versus $H_1: |\mu_T - \mu_R| < \delta$
    where δ = f*SD(ref)
- Non-parametric
  - Distribution free estimator
  - Median as location parameter
  - IQR or MAD instead of SD




**Quality Ranges for Test batches**

- Parametric
  - Sigma Interval, eg. X=3
  - Tolerance Interval for proportion β with γ-Confidence
  - Prediction Interval for k future individual test batches
  - ‚PI in TI': **P**rediction **I**nterval for 1 future test batch in **T**olerance **I**nterval from Reference
- Non-parametric
  - (Enlarged) min-max
  - smooth curve percentiles

# Simulations (10k simulations for each scenario)

Objective: Evaluate performances of all methods on different scenarios with parameter variations

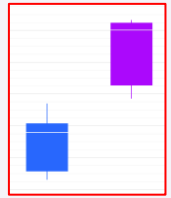
- **Normal** and **non normal** (gamma) distribution:
  Variation of mean shift (scale shift) and variance ratio (shape shift) for normally (gamma) distributed samples
- **Sample size**: 3 to 30, paired test and reference samples; balanced and unbalanced samples with unpaired values

- **Margin δ** for equivalence: 1.5 to 3*SD(ref)
- **Factor X** in the quality-range approach: 2 to 3
- **Confidence level** for TI in 'PI in TI': 90%, 95%, 99%
- **Proportion of test batches** in Smooth Curve: 90%, 100%

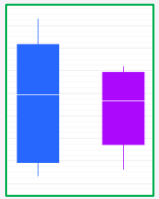
- **Confidence level** for PI and TI: 95%
- **Coverage** for TI: 90%
- **Percentile** for Smooth Curve: 95%

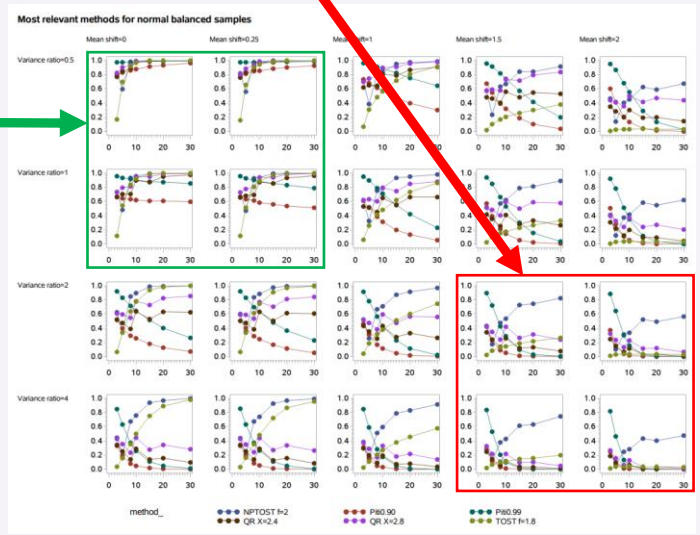
- Quantity of interest for each method:
  - Comparability claim rate plotted against test sample size

*Low probability to conclude comparability*



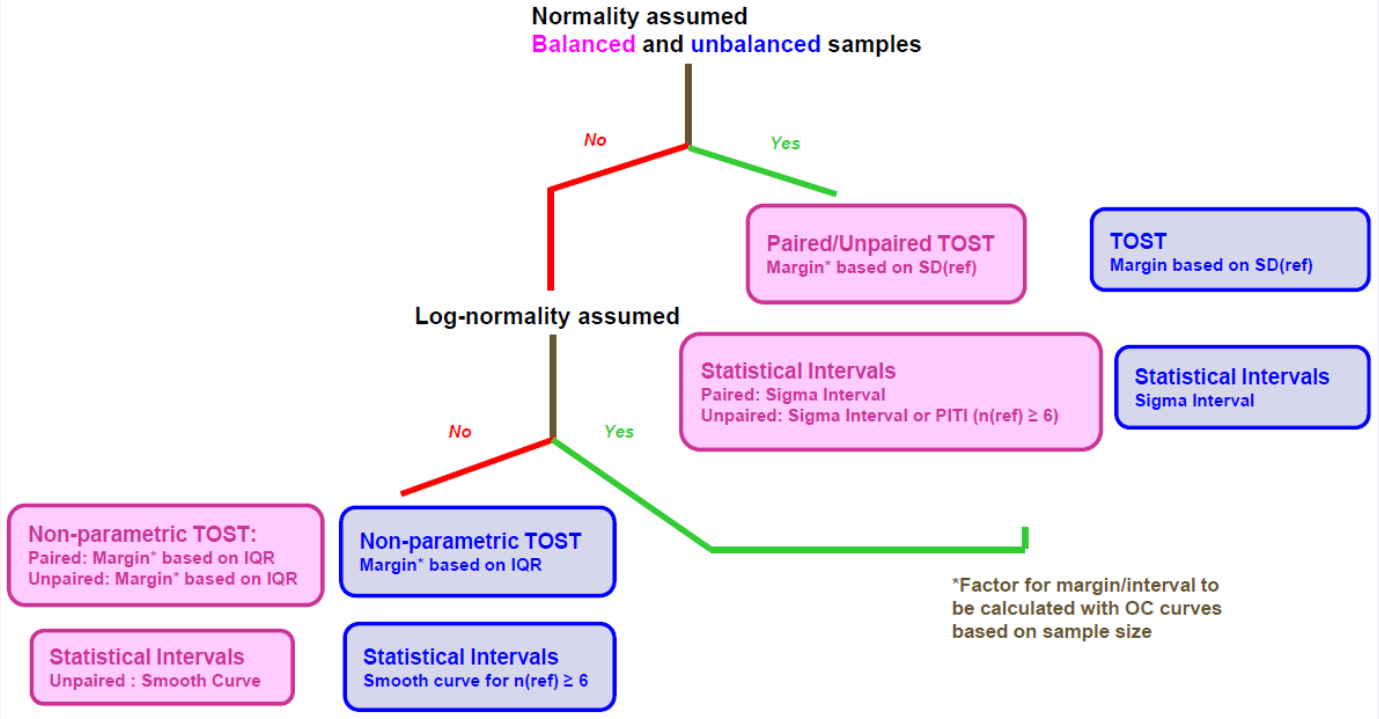
*High probability to conclude comparability*





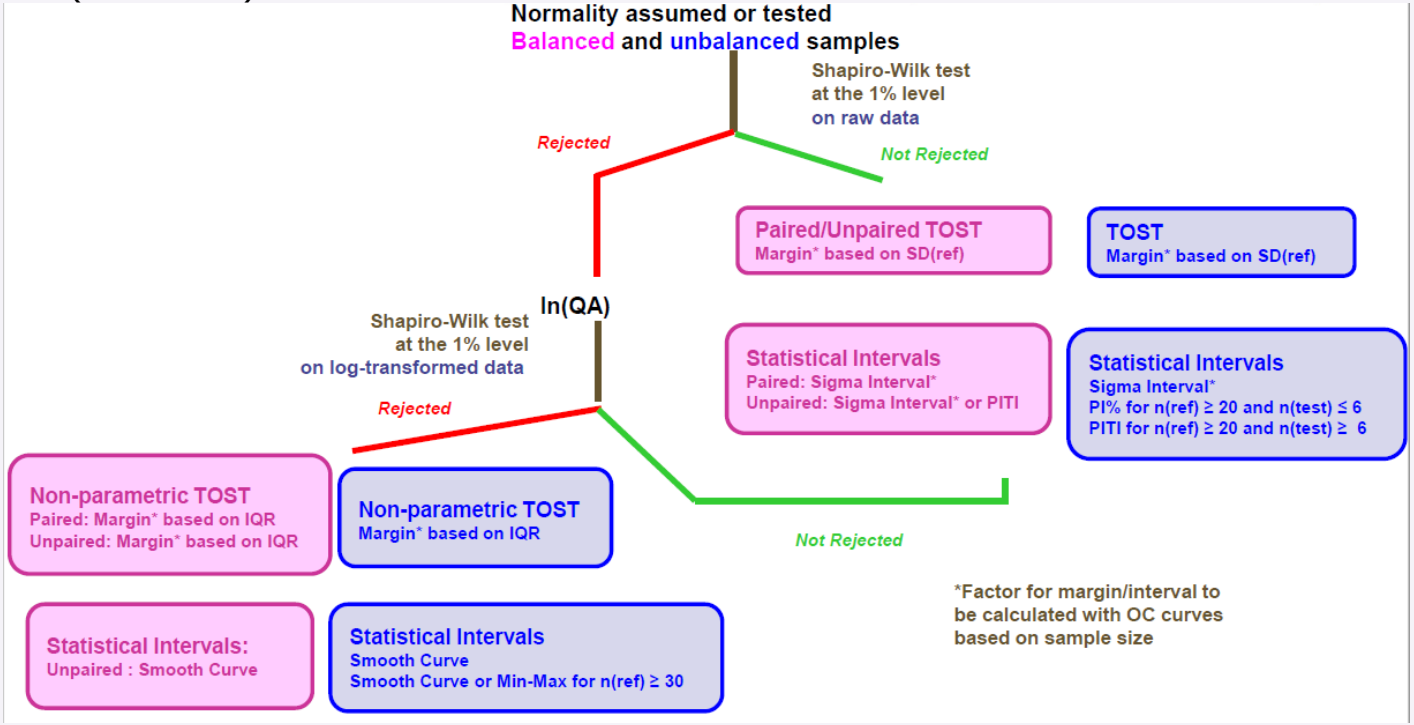
# Results

Decision tree for N (Reference) < 10



# Results

Decision tree for N (Reference) ≥ 10



# Conclusion

Objective: Provide recommendations based on simulation results -> Decision trees

- Recommended minimum sample size:
  - **3** for **normally distributed paired** samples
  - **5 or 6** for **unpaired or non-normal paired** samples
  - **Highly unbalanced: 3 test, 10 reference**

- For new studies the evaluated macros should be used to create **operating characteristics** (highly recommended by EMA) and decide on particular margins
- Important to **confirm comparability with CPV** for pre/post change or SDM qualification (in particular for low sample size)

- **Perspective**:
  - **If specifications are available**, quality range approaches can be combined with **requirements on capability**
  - **Acceptance criteria** for equivalence approaches independent from reference SD/SEM
  - **Multiplicity** correction
  - **Replicates** per batch
  - Normality **transformation**: e.g. Box-Cox

# Thank you

sanofi