Workshop on the Use of Statistical Methodologies in the Comparability Assessment of Quality Attributes

Hyungki Park

May 03-04, 2018

CONTENTS

- Overview of Statistical Tools for Comparability Assessment
- Reference Product Selection

SAMSUNG BIOEPIS

Overview of Statistical Tools for Comparability Assessment

Range Approach

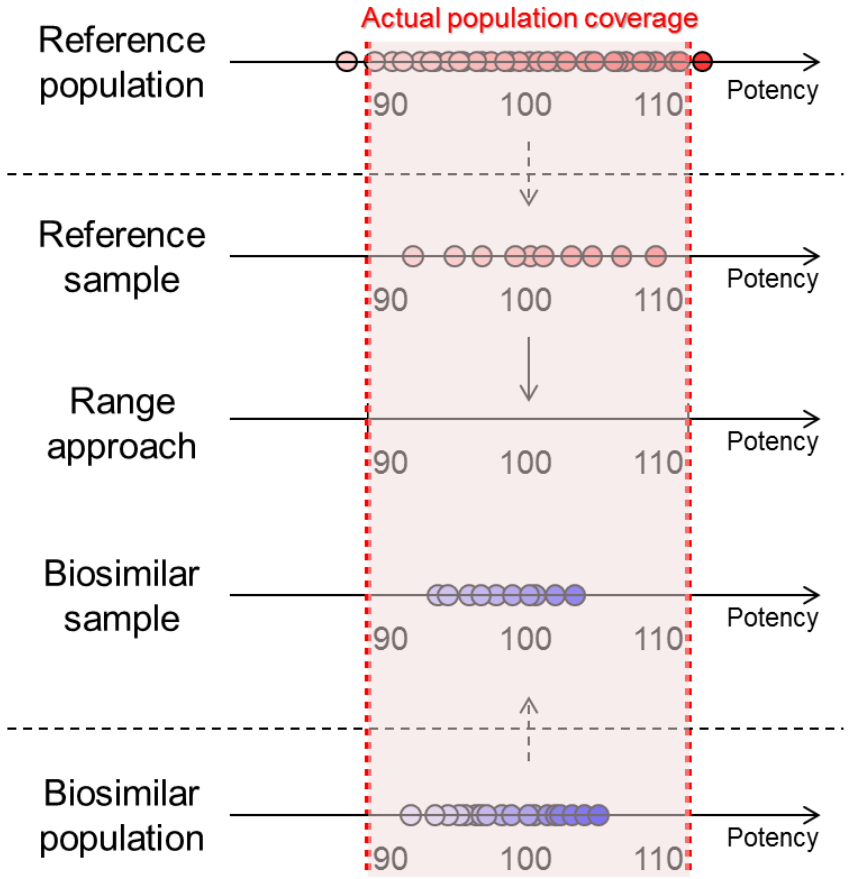
- Range approach utilizes samples of reference products in order to estimate the actual population
- Case studies show that range approaches consisting of min-max, x-sigma and TI can appropriately estimate the actual population

Equivalence Testing

- Equivalence of attributes measured on a continuous scale can be assessed by testing the difference in means between the proposed biosimilar and reference product
- Equivalence testing can be problematic when the mean of the reference product shifts over time and when the selection of relevant reference product batches is difficult

Range Approach

- Range approach utilizes samples of reference products in order to estimate the actual population



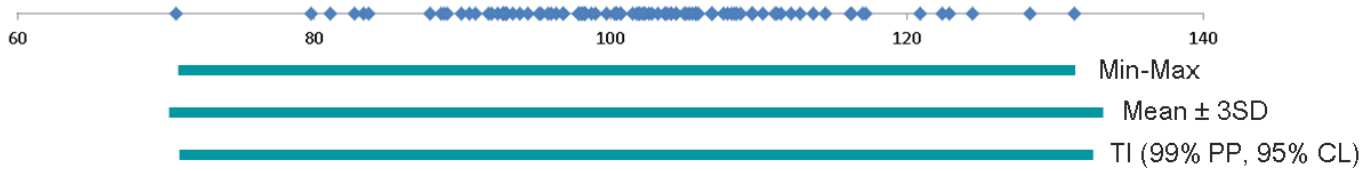
Range Approach Case Studies

- Case Study 1: Case study assuming appropriate sampling
  - 100 randomly sampled data were extracted from normally distributed population with mean of 100 and SD of 10
- Case Study 2: Case study associated with introduction of outliers
  - 100 randomly sampled data were extracted from normally distributed population with mean of 100 and SD of 10 (the same as Case Study 1)
  - Outlier 150.2 was intentionally included to mimic a test error
- Case Study 3: Case study associated with sampling chance
  - When sample size is small, skewed samples could be selected
  - 10 randomly sampled data were extracted from normally distributed population with mean of 100 and SD of 10
  - From those randomly sampled data sets, a skewed data set was selected

Range Approach Case Studies

SAMSUNG BIOEPIS

- Case Study 1: Case study assuming appropriate sampling



Method	Ranges	Actual Population Coverage*
Min-Max	70.7-131.3	99.7%
Mean ± 3SD	70.2-133.2	99.8%
TI (99% PP, 95% CL)**	70.9-132.5	99.8%

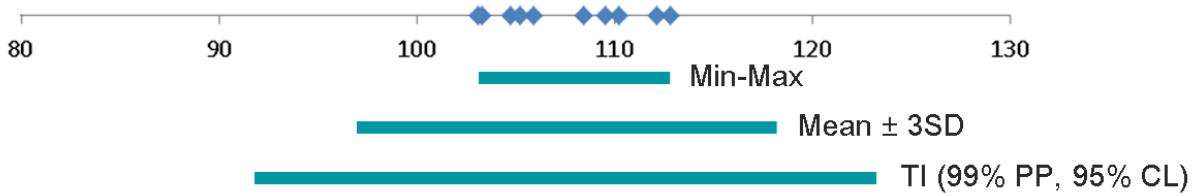
* Actual population coverage of the range estimated from the sampled data
** covers 99% proportion of population with 95% confidence level

- Case Study 2: Case study associated with introduction of outliers

Method	Ranges	Actual Population Coverage*
Min-Max	70.7-150.2	99.8%
Mean ± 3SD	67.6-136.5	99.9%
TI (99% PP, 95% CL)**	68.4-135.7	99.9%

* Actual population coverage of the range estimated from the sampled data
** covers 99% proportion of population with 95% confidence level

- Case Study 3: Case study associated with sampling chance



Method	Ranges	Actual Population Coverage*
Min-Max	103.1-112.8	27.8%
Mean ± 3SD	96.9-118.2	58.8%
TI (99% PP, 95% CL)**	91.8-123.3	78.5%

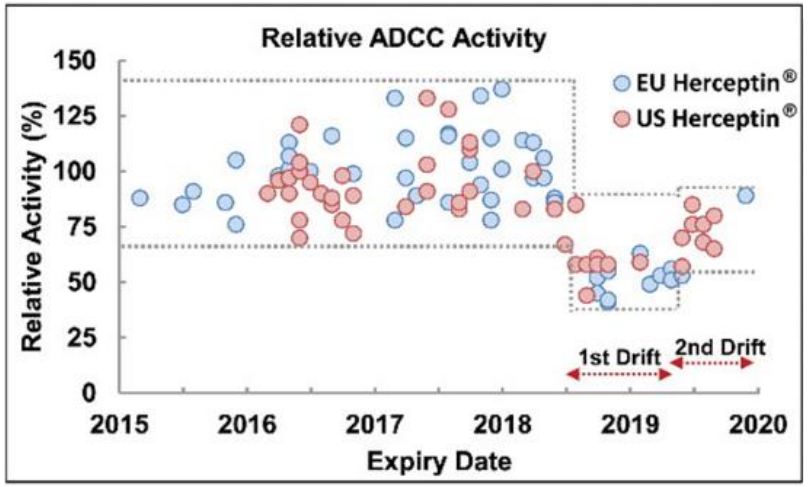
* Actual population coverage of the range estimated from the sampled data
** covers 99% proportion of population with 95% confidence level

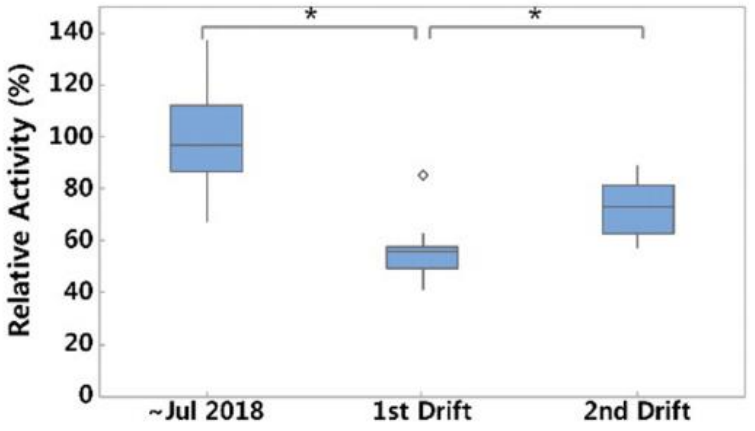
SB_2018_04_20_018_S

Equivalence Testing Case Study

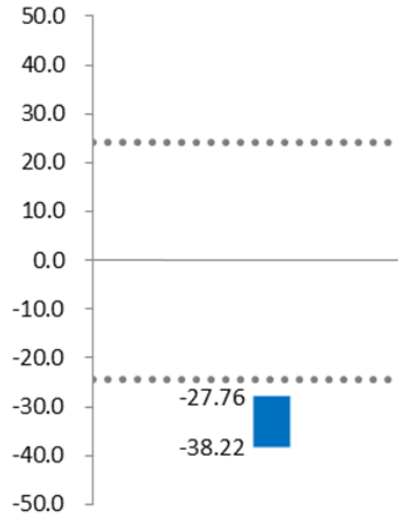
SAMSUNG BIOEPIS

- Equivalence testing comparing the means can be problematic since the mean of the reference products can change over time
- Variability of ADCC potency in Herceptin® reference product over time indicating a drift in mean ADCC potency
- Comparability within the originator products using equivalence testing can not be demonstrated due to the drift in mean





E.T. (Pre-drift vs. Post-drift)



Kim et al. mAb. (2017)

SAMSUNG BIOEPIS

Reference Product Selection

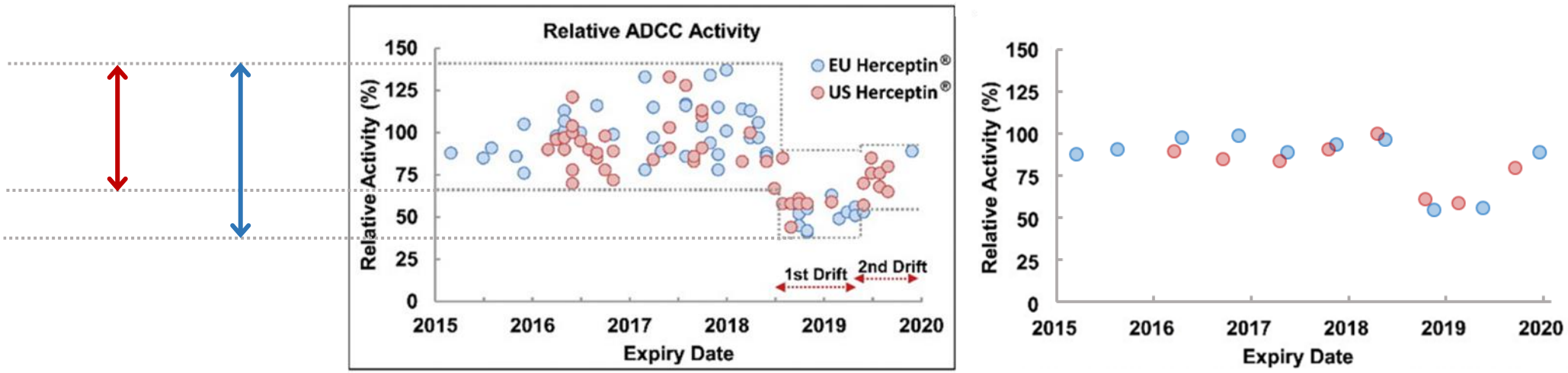
Agency's Recommendation on Reference Product Selection

- Age of the reference product should be taken into account for comparative assessment
  - It is not recommended to compare fresh biosimilar product with reference medicinal product at the end of the shelf life
- The reference product lots should be selected across the shelf-life of the approved product shelf-life
- The age of the reference product lots selected for the similarity assessment should be similar to that of biosimilar batches to minimize the impact of age on product quality

Reference Product Selection: Case Study 1

SAMSUNG BIOEPIS

- Results of statistical assessment, especially that of equivalence testing, could vary significantly depending on the selected reference product
- Selection of appropriate reference product for comparability assessment is difficult since the trend in quality can not be easily detected without extensive monitoring of the reference products

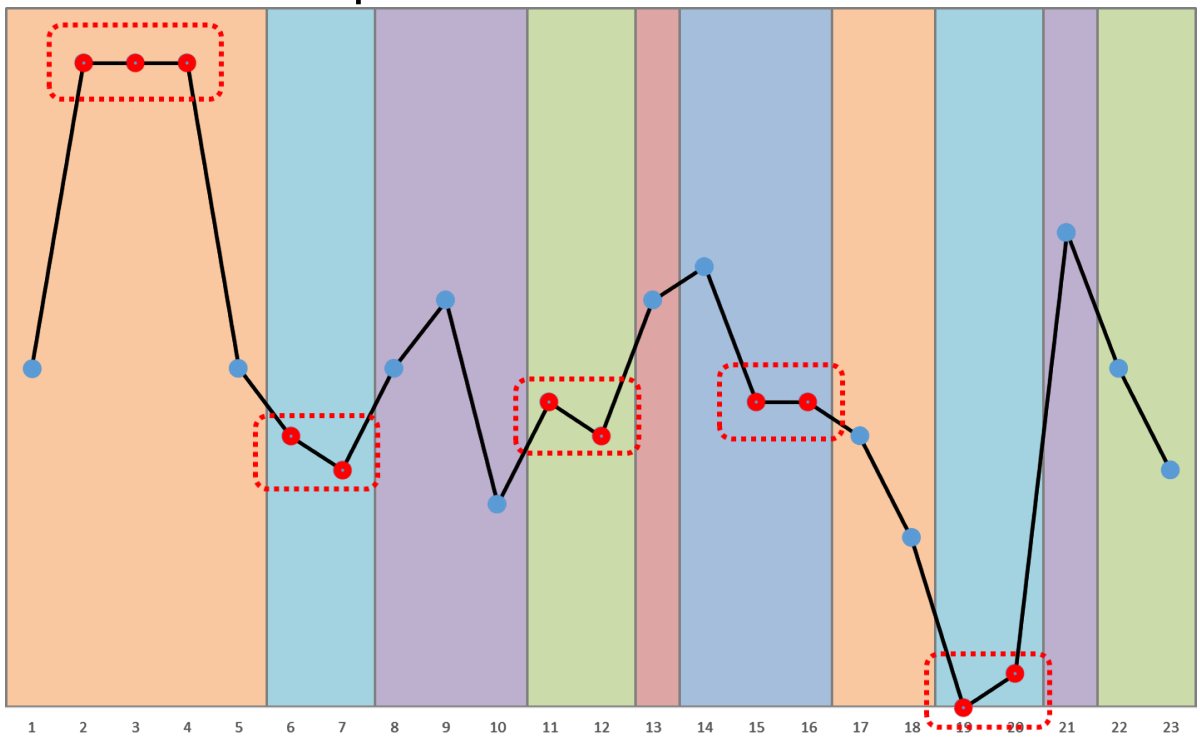


Kim et al. mAb. (2017)

SB_2018_04_20_018_S

Reference Product Selection: Case Study 2

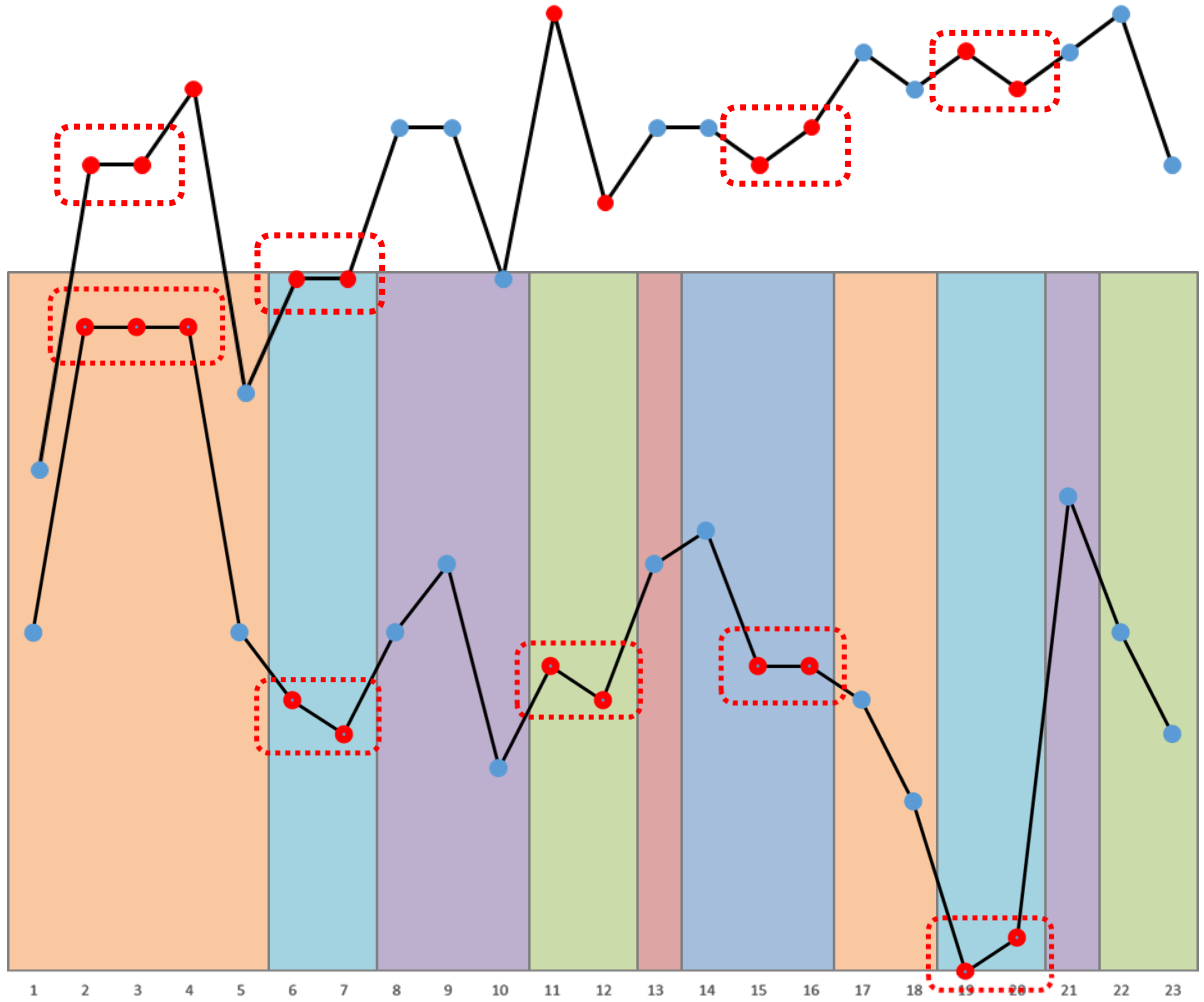
- When reference products made with the same drug substance are chosen, equivalence testing becomes problematic since it does not reflect the variability of the reference product accurately
  - Selection of reference product that represent the variability of the reference may be difficult since information of each reference products are not available to the biosimilar developers



Reference Product Selection: Case Study 2

SAMSUNG BIOEPIS

- Similar trend was observed for another quality attributes



Reference Product Selection: Case Study 2

SAMSUNG BIOEPIS

- Biosimilarity assessment results can differ significantly depending on the selection of reference products
  - Two samples (A and B) having 10 data each are all randomly sampled from normally distributed population with mean of 100 and SD of 10
  - The pass rate of sample B to fall within the comparability range established using sample A are calculated using simulations
  - When data clusters were intentionally introduced into sample A, the pass rate of each statistical tools were altered
  - The impact is largest for equivalence test

Method	Random	Adjusted*	Gap
Equivalence Testing	81.2%	68.1%	13.1%
Min-Max	23.5%	15.5%	8.0%
Mean ± 3SD	86.3%	76.0%	10.3%
TI (99% PP, 95% CL)**	98.2%	94.2%	4.0%

*adjusted to include data cluster to mimic data of DPs manufactured from the same DS
**covers 99% proportion of population with 95% confidence level

Conclusion

- All statistical tools have their own pros and cons, which are associated with the condition of the data sets
- EMA's draft reflection paper focuses on the limitation of statistical tools focusing on intervals, such as the min-max, x-sigma and TI
- Case studies show that equivalence testing can be problematic when the mean shifts over time and when the selection of relevant reference product batches is difficult
- Case studies show that range approaches consisting of min-max, x-sigma and TI, may be selected and used appropriately depending on the data sets
- Reflection paper and the subsequent guideline should provide flexibility in statistical tools used for comparability assessment

SAMSUNG BIOEPIS

Thank You