Critical Remarks on Reference-Scaled Average Bioequivalence

Yaroslavl, 22 October 2021
Highly Variable Drugs / Drug Products

- Clinically not relevant difference $\Delta$?
  - Based on pharmacokinetics but extrapolated to similarity of safety and efficacy in the patient population
    - $\Delta$ depends on the dose-response curves: HVD (flat), NTID (steep)
Statistical Hypotheses

• **Average Bioequivalence (ABE)**

\[ H_0 : \frac{\mu_T}{\mu_R} e \{ \theta_1, \theta_2 \} \ vs \ H_1 : \theta_1 < \frac{\mu_T}{\mu_R} < \theta_2, \]

where the – fixed – limits \{ \theta_1, \theta_2 \} of the acceptance range depend on the clinically not relevant difference \( \Delta \) by

\[ \theta_1 = 1 - \Delta, \ \theta_2 = (1 - \Delta)^{-1} \]

• **Scaled Average Bioequivalence (SABE)**

\[ H_0 : \frac{\mu_T}{\mu_R} / \sigma_{wR} e \{ \theta_{s1}, \theta_{s2} \} \ vs \ H_1 : \theta_{s1} < \frac{\mu_T}{\mu_R} / \sigma_{wR} < \theta_{s2}, \]

where \( \sigma_{wR} \) is the standard deviation of the reference and the scaled limits \{ \theta_{s1}, \theta_{s2} \} of the acceptance range depend on conditions given by the agency.
Frameworks

- Implemented

**Average Bioequivalence with Expanding Limits** «ABEL» (EMA, EEU, ...)

**Reference-Scaled Average Bioequivalence** «RSABE» (FDA, CDE)
Lack of Harmonization

- $\Delta > 20\%$
  - GCC $25\% \rightarrow$ BE-limits $75.00 - 133.33\%$ ($C_{\text{max}}$ only)
  - EMA, EEU Scaled based on $CV_{\text{wR}}$ ($C_{\text{max}}$ only)
  - WHO Like EMA (if justified, also $AUC$)
  - HC Like EMA ($AUC$ only)
  - FDA Scaled based on $CV_{\text{wR}}$ ($AUC$ and $C_{\text{max}}$)

<table>
<thead>
<tr>
<th>EMA, EEU, WHO, …</th>
<th>Health Canada</th>
<th>FDA, CDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CV_{\text{wR}}$</td>
<td>$BE$ limits (%)</td>
<td>$CV_{\text{wR}}$</td>
</tr>
<tr>
<td>$\leq 30$</td>
<td>$80.00 - 125.00$</td>
<td>$\leq 30$</td>
</tr>
<tr>
<td>$35$</td>
<td>$77.23 - 129.48$</td>
<td>$35$</td>
</tr>
<tr>
<td>$40$</td>
<td>$74.62 - 143.02$</td>
<td>$40$</td>
</tr>
<tr>
<td>$45$</td>
<td>$72.15 - 138.59$</td>
<td>$45$</td>
</tr>
<tr>
<td>$\geq 50$</td>
<td>$69.84 - 143.19$</td>
<td>$\geq 50$</td>
</tr>
<tr>
<td>$\geq 57.4$</td>
<td>$66.7 - 150.0$</td>
<td>$\geq 57.4$</td>
</tr>
</tbody>
</table>

$100 \exp(\mp 0.760 \cdot s_{\text{wR}})$

$100 \exp(\mp 0.8925742 \cdot s_{\text{wR}})$
Sample Sizes

- **Example TRTR | RTRT**
  - Assumed T/R-ratio = 0.90
  - Target power ≥ 80%
- **RSABE requires smaller sample sizes for target (desired) power than the ABEL variants**
Power

- **Example TRTR | RTRT**
  - Designed for ABEL
    - Assumed $CV_{WR} = 40\%$
    - Assumed T/R-ratio = 0.90
    - Target power $\geq 80\%$
    - $n = 30$ (80.7% power)
- **For any given sample size** the ABEL variants are less powerful than RSABE
- **Hypothetical situation**
  - The *same* study is submitted to *different* agencies
  - Might *pass* for one and *fail* for another
Inflation of the Type I Error

- SABE as implemented by agencies in ...
  - ABEL
  - RSABE
- ... are frameworks, where the acceptance limits are random variables depending on the observed variability
  - Strictly speaking, $\Delta$ is not defined beforehand
  - The model is based on the true – but unknown – population parameter $\sigma_{wR}$, whereas the study is assessed based on the sample $s_{wR}$
  - This may lead to a misclassification, i.e.,
    - the limits are scaled (because $CV_{wR} > 30\%$), although the drug is not highly variable and hence,
    - the chance to pass increases, compromising the patient’s risk\textsuperscript{1,2}

Inflation of the Type I Error

- Example TRTR | RTRT

ABEL

Inflated \(TIE\) with \(CV_{WR} \sim 24 - 42\%\)
- low dependency on sample size
  \((n = 20: 0.0800, n = 120: 0.0838)\)
- Maximum empiric \(TIE\) at true \(CV_{WR} = 30\%\)

RSABE

Inflated \(TIE\) with \(CV_{WR} < 30\%\)
- high dependency on sample size
  \((n = 20: 0.1251, n = 120: 0.2421)\)
Inflation of the Type I Error

- Example TRTR | RTRT
  - \(10^6\) simulated\(^1\) studies
  - \(n = 24, 36, 48\)
  - True \(CV_{WR} = 20 - 65\%\)
  - True T/R-ratio = \(\theta_{s2}\)
    - a Conventional ABE
    - b ABEL (EMA, EEU, and others)
    - c ABEL (Health Canada)
    - d ABEL (GCC)
    - e RSABE (implied limits)\(^2\)
    - f RSABE (desired consumer risk model)\(^2\)

1. No exact method exists for power and hence, the TIE in the implemented regulatory frameworks. Therefore, extensive simulations under the Null are required.
Realized $\Delta_r$

- **Example TRTR | RTRT**
  - 500 studies simulated for ABEL
  - $n = 34$ (81.2% power)
  - True $CV_{wR} = 35\%$ ($\Delta = 22.77\%$)
  - True T/R-ratio = 0.90
    - 417 studies passed (83.4%)
    - Realized $CV_{wR} 22.30 – 51.25\%$ ($\Delta_r 20.00 – 30.16\%$)

- **Every study sets its own rules, awarding ones with high $CV_{wR}$**
  - Without access to the study report, $\Delta_r$ is unknown to physicians, pharmacists, and patients alike
  - This is an unsatisfactory situation – we put the cart before the horse
Critical Remarks on Reference-Scaled Average Bioequivalence

Thank You!
Благодарю вас!

Helmut Schütz
BEbac
Consultancy Services for Bioequivalence and Bioavailability Studies
1070 Vienna, Austria
helmut.schuetz@bebac.at