What decision(s) are we making?

- The decision is whether two products (groups) can be considered equivalent.

What are the consequences of making an incorrect decision(s)?

- Incorrect decision could occur in 2 ways –
  - First: products are equivalent but deemed non-equivalent & could result in unnecessary additional testing or product changes and possible delays in market entry
  - OR
  - Products are non-equivalent but deemed equivalent & could result in products being brought to market that have different in-vitro properties compared to current product, meaning initial development work may not be appropriate

What approach is being proposed in this talk to improve decision making?

- This talk reviews properties of a decision making tool already in place, the population bioequivalence statistical criterion
Meeting the Quality Challenge for Orally Inhaled Drug Products

Review of the Performance Properties of the Population Bioequivalence Statistical Criterion

Beth Morgan, PhD
GlaxoSmithKline
Acknowledgements

• Presented on behalf of the IPAC-RS Population Bioequivalence (PBE) Working Group
  - Co-chairs:
    • Dave Christopher, Merck
    • Beth Morgan, GlaxoSmithKline

• Thank you to the Working Group for engagement and to the IPAC-RS Board of Directors for support of this research

• Special thank you to Stephanie Chen, North Carolina State University, for statistical and programming support
Outline

• Introduction of Population Bioequivalence Criterion

• Update PBE WG Activities
  - Initial key findings
  - Database description
  - High level overview of process followed

• Conclusion and Next steps
Decision Making Based on PBE

• Statistical tool used to determine whether 2 populations are equivalent
  - PBE is current standard referenced in FDA guidance
  - Typically applied to delivered dose, single actuation content, aerodynamic particle size distribution, droplet size distribution, ovality ratio

• Types of equivalency decisions:
  - Comparison between innovator & generic products
  - Comparison between pre-approval or post-approval changes to the same product
Regulatory Background: US

• First appeared in general 2001 guidance on BE:

• Appeared in 2003 for in-vitro data
  - 2003 Guidance on BA/BE Studies for Nasal Aerosols and Nasal Sprays for Local Action:

• Most recently documented in Product Specific Guidances
    • Provided defined equivalency criterion and incorporated multiple measurements within a unit (called life stages)
The equation is an expression of differences in means and variances in the metric in which normality is assumed, *log normal* scale.

Complex in terms of understanding changes in 4 parameters mapped to 1 parameter.

- Equivalency defined as theta <= 2.089
- No clear understanding within *in vitro* arena of how changes in these 4 parameters relate to equivalency
• IPAC-RS established a PBE Working Group (WG) to study **performance properties** of the PBE applied to in vitro data
  • Limited research work conducted in this area based on realistic *in vitro* data scenarios to understand impact to decision making

• **Activities to date**
  • Designed **database** to collect realistic industry data
  • **Studied** database to understand specific patterns and potential impact on the PBE performance
  • Developed statistical **model** for simulations: focus on dose
  • Conducted initial **simulations** to understand performance properties of the statistical criterion
Key Findings to Date

• Re-parameterized the criterion in terms of parameters in the original scale
  - *Constant scaling* implemented when RSD of reference product $\leq 10$
  - When constant scaling used and variation comparable, equivalence defined as test mean, approximately, *-13% to +16% of reference mean*

• Probability of declaring equivalence within this region will increase:
  - If *test mean higher than reference mean* for the same difference
  - If *between batch variation lower than within batch* for the same total variability
  - If analysis based on *multiple life stages*
Overview of Database

• Currently based on MDI data
• Contains several tests: delivered dose, fine particle mass, impactor sized mass, mass balance
• Total of 59,277 measurements
  • Across 25 products and 1043 batches
  • Contains release and stability data
  • Contains data from comparability protocols - as well as commercial and development data
• Data “double blinded”
  • Historical database built across multiple in-vitro tests
  • Challenges to maintain variance structure of original data but in log scale
Patterns Different for Delivered Dose Compared to APSD data

Presence of outliers can impact performance properties of a statistical test
Statistical Modeling and Simulation Considerations

• Focused first on delivered dose because patterns differ across tests

• Developed a statistical model that allows flexibility to incorporate:
  • Batch to batch variation
  • Systematic change across life stages
  • Life stage effect that can vary across batches and across units within a batch
  • Outlier distribution
  • Mixture of normal distribution with outliers
Model Assessment Output

- Used a variety of visual and statistical tools to assess how well the simulated data fit to the actual data:
  - Histograms, Box plots
  - Q-Q plots for normally distributed data
  - Variability plots to understand changes across life stage
Original Data

Simulated Data

Simnum=1

Simnum=50

Simnum=80
Actual Data
Each line represents a can

BOU               MOU           EOU

Sim=1

BOU               MOU           EOU

Sim=50

BOU               MOU           EOU

Sim=80
Simulated Data

Simnum=1

Mean: 99.42, Std: 5.81

Simnum=50

Mean: 100.12, Std: 5.70

Simnum=80

Mean: 100.10, Std: 6.29

*QQ-plot horizontal reference line at 115
Types of questions being considered

• How does theta < 2.089 equate to **parameter changes** in **original** scale?

• What is the impact of **constant scaling**?

• What is the impact of the **log transformation**?

• What is the impact of the presence of **outliers**?

• What is the impact of changing these factors:
  • Different amounts of between **batch and within batch variation**
  • Analysis based on **1 life stage or multiple life stages**
  • **Differing amounts of variability** between 2 populations
Impact of Log Transformation and Constant Scaling

• Theta criterion reduces to a function of logs of ratios between means and differences in RSD in original scale

• This creates asymmetry for differences between means; equivalence more likely if Test Mean > Reference Mean

• Constant scaling allows a greater difference in means than would be allowed based on reference scaling
Effect of Constant Scaling: RSD=2%, Variance Equal

Mean In Log Scale

Reference Mean at 45, 100, 200

Red Area=Range for Test Mean for Equivalency with Reference Scaling
-5 % to + 6 %

Green Area=Range for Test Mean for Equivalency with Constant Scaling
-13% to +16 %
Effect of Varying Between Batch and Within Batch Variation

• PBE assumes no between batch variation and pools all data into a “super batch”

• If there is between batch variation then equivalence less likely to be declared
  - If the total variation is the same but more of the variation is due to between batch, then, the probability of equivalence decreased within the acceptance region
Effect of varying between batch and within batch variation when variation comparable

Between Batch Var < Within Batch

RSD=4%

RSD=10%
Effect of including 2 life stages

- Equivalence more likely to be declared with 2 life stages

- This is because the sample size is increasing and the variation in the test statistic is decreasing
Effect of increasing number of life stages

- **RSD=4%**
- **RSD=10%**
Work Ahead

• Further analysis within dose to better understand:
  - Variation changing between test and reference products
  - Extreme values present
  - Publish work to date

• Apply overall modeling and simulation approach to fine particle mass and impactor sized mass
Concluding Thoughts

• Blinded industry data can serve a critical scientific role

• This type of work can be a powerful approach for better understanding realistic impact of decision making tools

• Work to date has lead to better understanding of what is defined as equivalent and the types of factors that impact the decision outcome and can serve the industry as a whole
THANK YOU