# BASS November 2015:

# OPTIMALITY WHEN ANALYZING CROSSOVER DESIGNS WITHOUT WASHOUT PERIODS

Ben Trzaskoma
Sr Statistical Scientist
Genentech Inc

# Background and Rationale: The IMPART Study

- Patients with cystic fibrosis have a high treatment burden and daily adherence to lifelong therapies can be challenging1

- The Pari eRapid device is a general purpose electronic nebulizer utilizing vibrating membrane technology2

  – Smaller, lighter, quieter, and more portable than conventional jet nebulizer/air compressor systems

  – In vitro experiments have demonstrated that eRapid delivers a comparable dose of Pulmozyme® (dornase alfa) with similar aerosol characteristics compared to the Pari LC® Plus jet nebulizer3

- Study Rationale: Provide CF community and FDA with robust clinical data to evaluate whether Pulmozyme can be used effectively with eRapid

1. Sawicki GS, et al. High treatment burden in adults with cystic fibrosis: challenges to disease self-management. *J Cyst Fibros*. 2009;8:91-96; Zemanick ET, et al. Measuring and improving respiratory outcomes in cystic fibrosis lung disease: opportunities and challenges to therapy. *J Cyst Fibros*. 2010;9:1-16
2. Pari received FDA 510(k) clearance for eRapid in 2012
3. Scherer T, et al. A technical feasibility study of dornase alfa delivery with eFlow® vibrating membrane nebulizers: aerosol characteristics and physicochemical stability. *J Pharm Sci*. 2011;100(1):98-109

# IMPART Key Objective

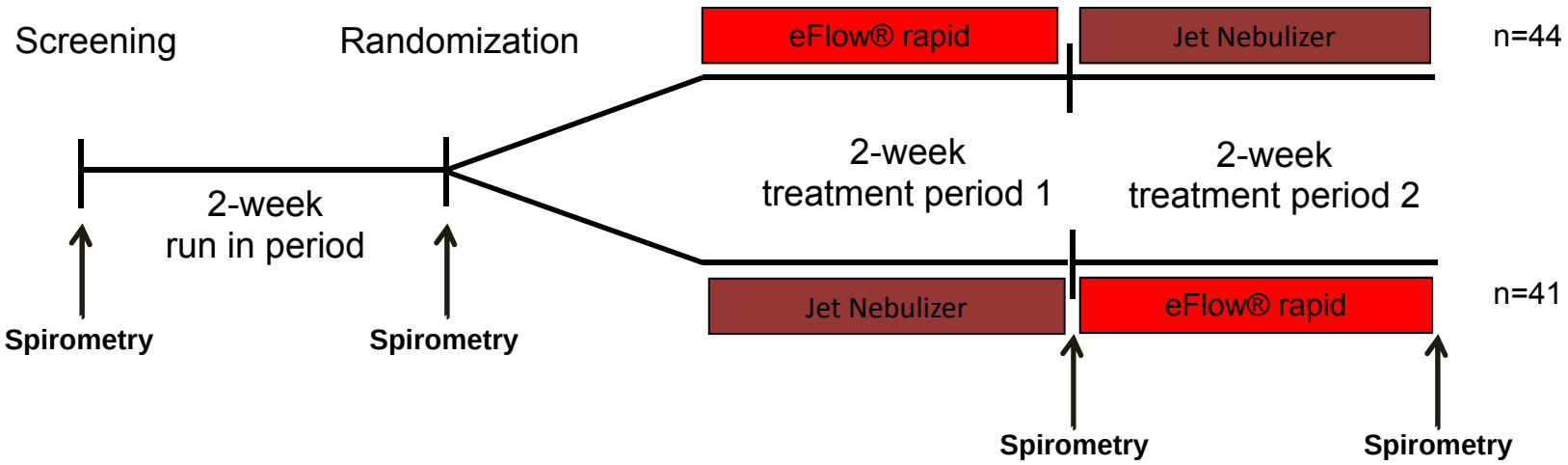
- Primary objective was to compare:

  – Stability of lung function as assessed by FEV1 when Pulmozyme is delivered by eRapid vs Pari LC Plus jet nebulizer system

# IMPART: A Phase 4, Multicenter, Randomized, Open-Label, Two-Period Crossover Study

- Patients: ≥ 6 years of age with confirmed diagnosis of CF and receiving Pulmozyme chronically for at least 6 months prior to screening

- Randomized in 1:1 fashion to either treatment sequence 1 or 2




Screened, N=103

Randomized, n=87

Completed treatment, n=85
Completed study, n=85

# Traditional 2x2 Crossover Model

- Each subject is randomly assigned to either sequence 1 (RT) or sequence 2 (TR):

| | | |
|---|---|---|
| 1 (RT) | Reference Formulation Data" Yi11 | Test Formulation Data" Yi21 |
| 2 (TR) | Test Formulation Data" Yi12 | Reference Formulation Data" Yi22 |

Note: Dosing periods are "usually" separated by a washout period

# General 2x2 Crossover Model

$$Yijk = \mu + Sik + Pj + F(j,k) + C(j-1,k) + eijk\ ,$$

where

i(subject) = 1,2,…,nk

j(Period) = 1,2

μ is the overall mean

Sik ~ iid N(0, σ2s) is the random between-subject effect

Pj is the fixed effect of the jth period

F(j,k) is the fixed effect for treatment in the jth period, kth sequence

C(j-1,k) is the fixed carryover effect of treatment in the kth sequence which

# Summary of Fixed Effects

- Each subject is randomly assigned to either sequence 1 (RT) or sequence 2 (TR):

| | | |
|---|---|---|
| 1 (RT) | $\mu 11 = \mu + P1 + FR$ | $\mu 21 = \mu + P2 + FT + CR$ |
| 2 (TR) | $\mu 12 = \mu + P1 + FT$ | $\mu 22 = \mu + P2 + FR + CT$ |

Where:
$\mu jk = E(Yijk)$,
$P1 + P2 = 0$, $FR + FT = 0$,
$CR + CT = 0$

# Background Data from Registry

- From our large CF registry, we found that for patients treated with Pulmozyme for at least 6 months have:

  -An FEV1 % predicted with a SD of about 20.

  -A 2-week change in FEV1 % predicted with a SD of about 8.

Therefore:

Var(Yijk) = V**ar** (Sik) + Var(eijk) = **σ2s + σ2e = 202** ,

Var(Yi,j+1,k - Yi,j,k) = Var(eijk – ei,j+1,k) = **2 σ2e = 82**.

So that, **σe = 5.7 and σs = 19.2.**

# The Million Dollar Question

*So if a 2 week change has a std deviation of 8 and an individual timepoint has a std of 20,  shouldn't we lean towards using a 2-week change over an individual timepoint for the analysis? (Hint: We don't have a washout, so it isn't straightforward)

# Which Endpoint to Use?

- "Standard" four point design:

Diff1 = (Yi,2,k - Yi,1,k ) - (Yi,4,k - Yi,3,k )

Day 1 Day 14 washout Day 28 Day 35

- Our design analyzing changes between periods:

Diff2= (Yi,2,k - Yi,1,k ) - (Yi,3,k - Yi,2,k )



Day 1 Day 14 Day 28

# Which Endpoint to Use?, contd…

- Our design, comparing value at each timepoint:

Diff3= (Yi,3,k - Yi,2,k )

Day 1 Day 14 Day 28

- Our design a[illegible]anges f[illegible]inal baseline:

Diff4= (Yi,2,k - Yi,1,k ) - (Yi,3,k - Yi,1,k )



Day 1 Day 14 Day 28

# Variances of 4 approaches

- Four-point design:

Var(Diff1)=Var{(Yi,2,k - Yi,1,k ) - (Yi,4,k - Yi,3,k )} =

Var{ c + (ei,2,k-ei,1,k)-(ei,4,k-ei,3,k) } = **4 σ2e**

- Our design, changes btwn periods:

Var(Diff2)=Var{ (Yi,2,k - Yi,1,k ) - (Yi,3,k - Yi,2,k ) } =

Var{ c + ei,3,k - 2ei,2,k + ei,1,k } = **6 σ2e**

- Our design, comparing timepoints:

Var(Diff3)= Var{(Yi,3,k - Yi,2,k )} = Var{ c + (ei,3,k - ei,2,k)}=**2 σ2e**



- Our design, orig change from BL:

# Original Proposal: Non-Inferiority Test

Ho: $\mu T - \mu R \leq -2.5$

HA: $\mu T - \mu R > -2.5$ (Not Inferior)

at $\alpha = 0.025$

| Endpoint | Variance of statistic | Power of Test |
|---|---|---|
| Four point Design | **4 σ2e** | 73% |
| Our design: Change between periods | **6 σ2e** | 56% |
| Our design: Comparing each timepoint | **2 σ2e** | 95% |
| Our design: Change from original baseline | **2 σ2e** | 95% |

# Statistical Methods after FDA Type C Meeting

- The FDA felt that 140-160 patients with non-inferiority testing was “overkill” and suggested that we just gather data on a smaller set to evaluate how the nebulizers compare to each other

- After negotiating a bit, it was decided that 60-90 patients would be adequate and that the principles of “bioequivalence” could be employed.

- Primary efficacy analysis:

  - Based on “bioequivalence” principles to compare the mean percent predicted FEV1 at the end of each treatment period between the eRapid nebulizer and the Pari LC Plus nebulizer

  - According to the FDA, the two nebulizers will be considered equivalent if the 90% confidence interval for the ratio of the mean percent predicted FEV1 values at the end of each treatment period were within 80%-125%

# Fieller's Method for CI of a Ratio

- Goal is to get a confidence interval for $\rho = \frac{\mu_T}{\mu_R}$.
- Major assumption: the N pairs (xi,yi) are iid bivariate normal (or at least $(\bar{x}, \bar{y})$ is).
  - Unbiased estimators for the expected values $(\mu_R, \mu_T)$ are the sample means $\bar{x}$ and $\bar{y}$ with variances and covariances:
    $\hat{\sigma}^2_x = \frac{1}{N}\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$, $\hat{\sigma}^2_y$ = (analogous to $\hat{\sigma}^2_x$)
    $\hat{\sigma}_{x,y} = \frac{1}{N}\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$.
- $\hat{\rho} = \frac{\bar{y}}{\bar{x}}$ is an intuitive point estimate but has a complex distribution.

# Fieller's Method contd…

- Because the difference of normal variables is normal, the term $\bar{y} - \hat{\rho}\,\bar{x}$ is normal. Dividing this term by the appropriate estimate of the standard deviation gives us:

$$To = \frac{\bar{y} - \hat{\rho}\,\bar{x}}{\sqrt{(\hat{\sigma}^2_{\bar{y}} - 2\,\hat{\rho}\,\hat{\sigma}_{\bar{x},\bar{y}} + \hat{\rho}^2\,\hat{\sigma}^2_{\bar{x}})}}$$

which has a t-distribution. So to obtain confidence limits for ρ we calculate the set of ρ values for which To lies within the (1-α) quantiles of the t distribution (denoted $t_q$). This gives us confidence limits of:

$$\frac{(\bar{x}\,\bar{y} - t^2_q\,\hat{\sigma}_{\bar{x},\bar{y}}) \pm \sqrt{(\bar{x}\,\bar{y} - t^2_q\,\hat{\sigma}_{\bar{x},\bar{y}})^2 - (\bar{x}^2 - t^2_q\hat{\sigma}^2_{\bar{x}})(\bar{y}^2 - t^2_q\hat{\sigma}^2_{\bar{y}})}}{\bar{x}^2 - t^2_q\hat{\sigma}^2_{\bar{x}}}$$

# IMPART Baseline Characteristics

| Characteristic* | mITT Patients (n=85) |
|---|---|
| Age, years | 13.6 (6.9); range 6-44 |
| Categorical age: 6-13 years old, n (%) | 50 (58.8) |
| Categorical age: 14 years and older, n (%) | 35 (41.2) |
| Female, n (%) | 43 (50.6) |
| Caucasian, n (%) | 83 (97.6) |
| Time since CF diagnosis, years | 11.8 (6.4) |
| Lung function (spirometry) | |
| FEV1, liters | 2.5 (0.9) |
| FEV1, percent predicted | 97.5 (21.8) |
| FVC, liters | 3.0 (1.1) |
| FVC, percent predicted | 101.8 (17.4) |
| FEF25-75, liters/sec | 2.7 (1.2) |
| FEF25-75, percent predicted | 92.3 (37.0) |
| Bronchodilator, n (%)** | 80 (94.1) |
| Inhaled Antibiotic, n (%)** | 33 (38.8) |

*Mean (standard deviation), unless noted

**Bronchodilator and inhaled antibiotics include incidence at baseline and concomitant meds during study

# Primary Analysis: FEV1 Was "Equivalent" Between the 2 Nebulizers

| | | | | | | |
|---|---|---|---|---|---|---|
| | eRapid | | | Jet Nebulizer | | |
| | Period 1 | Period 2 | Either Period | Period 1 | Period 2 | Either Period |
| n | 44 | 41 | 85 | 41 | 44 | 85 |
| Mean (SD), % | 99.2 (20.0) | 96.9 (24.4) | 98.1 (22.1) | 94.3 (21.6) | 100.0 (19.7) | 97.2 (20.7) |
| Range, % | 51.5, 145.5 | 52.2, 142.0 | 51.5, 145.5 | 52.7, 133.1 | 50.0, 142.1 | 50.0, 142.1 |
| Mean ratio, (eRapid/Jet either period) , % | 100.9 | | | | | |
| **90% CI for ratio of means, %** | **(99.5, 102.3)** | | | | | |

*Only raw data were used in the above table after extensive simulations were performed

# Summary

- Careful consideration is required when analyzing non-standard crossover designs (e.g. No wash-out periods)

- The general 2x2 crossover model is very powerful and can incorporate a range of possible endpoints as well as fixed and random effects

- Regardless of the testing procedure, when examining unbiased estimators for the testing of treatment effects one must compute and compare the variances among the candidates.

- In the IMPART study, comparable efficacy was observed in CF patients when Pulmozyme was delivered by eRapid vs standard jet nebulizer (LC Plus) and the FDA approved the label update to include eRapid in summer 2014.