The Unique Challenges of Cell Therapies: JMP® Scripts for Power and Sample Size Calculations

W. Heath Rushing
Andrew Karl
Jeff Hofer
Rick Burdick

JMP Discovery Americas 2023

ADSURGO
ascend with analytics

Outline

- Cell and Gene Therapies (CGTs)
- The need for comparability
- Comparability
- JMP® Sample Size Explorer
- JMP® Scripts

DSURGO
ascend with analytics

Biopharmaceutical's New Frontier

- The most novel, innovative and promising therapeutics in biopharmaceuticals are gene and cell therapies.

- The first such therapy approved in the United States was Kymriah in 2017. Since that time, there has been countless companies exploring these novel approaches to treating patients.

- One of the main barriers to regulatory approval is development of a process that consistently meets the needs of patients (efficacy and safety). One obvious challenge is modeling either patient-to-patient or donor-to-donor variability that can (and usually does) account for a large portion of the variability.

DSURGO
ascend with analytics

Gene Therapy

- Many diseases such as cancer are caused by defect genes. You can be born with these defective genes or they mutated over time.
- It is possible to modify or replace defect genes with healthy ones, usually to produce a necessary protein. This can ultimately cure (or prevent) diseases.
- Gene therapy
  - Replace one gene with another healthy one
  - Add or 'turn off' genes
- This can be do in vivo or in vitro
  - In vivo: Deliver (inject) healthy genes into the patient
  - In vitro: blood or bone marrow taken from patient, genes inserted, then delivered back into patient
- Must use a vehicle to deliver the genes to the patient, referred to as a vector (an example would be a virus).
- Vector (virus) carries information on how to fight illness to either the chromosomes or the nucleus of cells and then are broken down in the body.
- Efficacy and safety
  - Does it work?
  - Is it safe (patient could reject these cells then body starts killing itself)?
- Challenge: a patient is a lot/batch (patient-to-patient variability)

https://www.fda.gov/consumers/consumer-updates/what-gene-therapy-how-does-it-work

DSURGO
ascend with analytics

Cell Therapy

- Your body contains specialized cells which has specific functions.
- Cell therapies transfer healthy human cells into a patient to treat disease. These cells either replace diseased ones or missing cells. These cells either come from the patient themselves or a healthy (cell) donor (think bone marrow transplants) .
- Cell therapies are done in vitro to cure or prevent diseases (both vaccines and therapies):
  - Inject or transplant healthy cells into a patient.
  - Replace missing cells or diseased ones.
- Efficacy and safety.
- Challenge: donor-to-donor variability

https://www.bsgct.org/education/what-is-cell-therapy.aspx

DSURGO
ascend with analytics

Biopharmaceutical's New Frontier

One of the main barriers to regulatory approval is development of a process that consistently meets the needs of patients (efficacy and safety).

- Type I (manufacturers risk) error rate and Type II error (patient risk) rate (and power).
- Manufacturers do not want to inflate the Type I error rate
- Regulatory agencies want manufacturers to maximize power

One obvious challenge is modeling either patient-to-patient or donor-to-donor variability that can (and usually does) account for a large portion of the variability.

This donor-to-donor variability can (and usually does) account for a large portion of the variability in the response.

DSURGO
ascend with analytics

The Need for Comparability

In CGTs, like pharma and biopharma, there are changes to analytical methods and processes.

- Analytical
  - Qualification of a new reference standard
  - Assay transfer
  - Bridging studies (change in analytical procedures)

- Process:
  - Viral vector
  - Technology transfer (change in scales, manufacturing site changes)
  - Process change

The regulatory agencies recognize the need for changes, but they also recognize the need to ensure the therapeutics that are measured or manufactured are “similar” in terms of product quality.

Comparability

Comparability says they do not have to be the same, the have to be comparable or similar. What does “similar” mean? Similarity condition.

- Side-by-side plots

- Quality Range Approach
  - *k* number standard deviations around the reference group
  - Ensure the measured quality attributes are within this range

- Equivalence Test (high risk attributes)
  - Two One Sided Tests (TOSTs)
  - Determination of eqivalence acceptance criterion (EAC)
  - EAC is the acceptable difference (**d**) to ensure the two processes are comparable

DSURGO
ascend with analytics

Equivalence Testing

- A t-test for a difference in means (DOE used in PD):

$H_0: \mu_L - \mu_H = 0$	The means are not different.
$H_A: \mu_L - \mu_H \neq 0$	The means are different.
$\alpha = 0.05$	95% confidence

- An equivalence test:

$H_0: \lvert\mu_H - \mu_N\rvert > d$	The means differ by more than d.
$H_A: \lvert\mu_H - \mu_N\rvert \leq d$	The means differ by at most d.
$\alpha = 0.05$	95% confidence

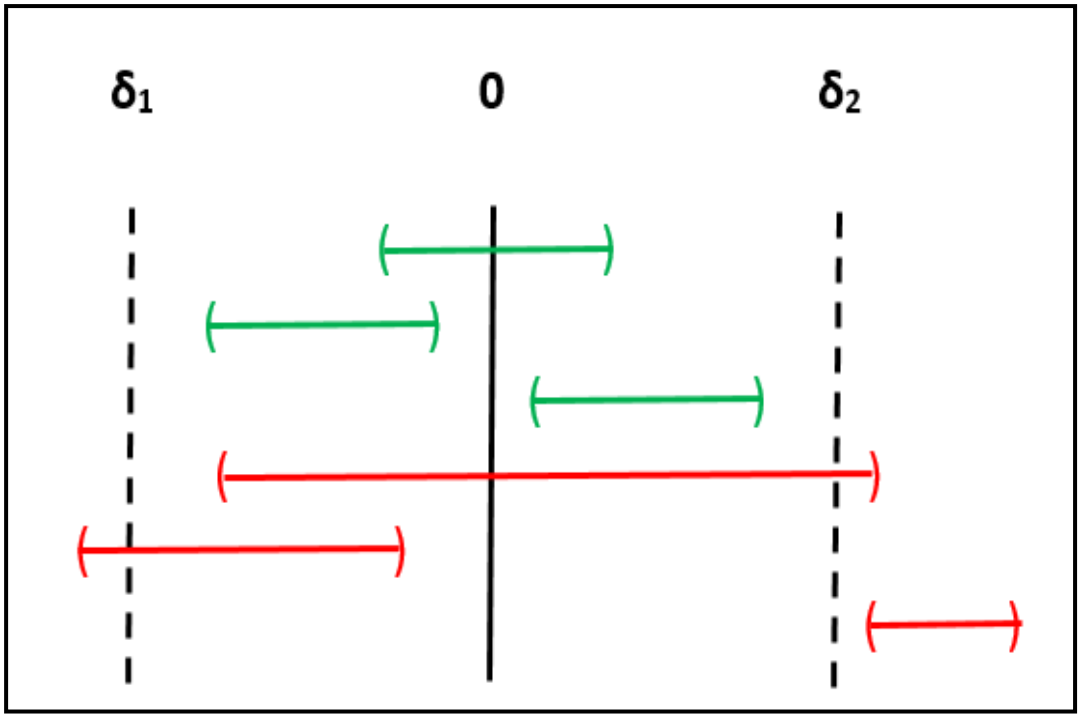
$H_0: \mu_H - \mu_N < +d$
$H_A: \mu_H - \mu_N \geq -d$
$\alpha = 0.05$

$H_0: \mu_H - \mu_N > -d$
$H_A: \mu_H - \mu_N \leq +d$
$\alpha = 0.05$

DSURGO
ascend with analytics

Equivalence Margin

– Selection of the equivalence acceptance criteria (d) is the key to the outcome of similarity.



Reference: Tsong, Yi, and OB CMC Analytical Biosimilar Method Development Team (Meiyu Shen, Cassie Xiaoyu Dong). 2015. *Development of Statistical Approaches for Analytical Biosimilarity Evaluation* [PowerPoint]. DIA/FDA Statistics Forum.

DSURGO
ascend with analytics

Equivalence Testing – Sample Size and Power

- In order to pass the TOST, manufacturers must reject both null hypothesis. This is the same as having a 90% confidence interval within the ± EAC bounds.
  - Therefore, manufacturers want to minimize their Type II error rate (manufacturers risk), maximize their power. How? Minimize the width of this confidence interval.
- If manufacturers increase the Type I error, this would increase the probability of claiming equivalence when the processes are not equivalent.
  - Therefore, regulatory agencies want to ensure the Type I error rate (patient risk) remains at 0.05 (that it is not inflated!)

DSURGO
ascend with analytics

JMP Sample Size Explorer

- One-sample Equivalence Test of Means
  - Population standard deviation known or not?
  - Equivalence, Superiority, or Non-inferiority?

- Two-sample Equivalence Test of Means
  - Population standard deviation known or not? Equal or unequal variance?
  - Equivalence, Superiority, or Non-inferiority?

- Currently does not address the challenges associated with with a split apheresis design (paired t-tests)
  - Sample size and power calculations
  - Overlay different scenarios for comparison
  - Dependendent of the proportion of donor-to-donor variability (ρ)

DSURGO
ascend with analytics

JMP Sample Size Explorer

- One-sample Equivalence Test of Means
  - Population standard deviation known or not?
  - Equivalence, Superiority, or Non-inferiority?

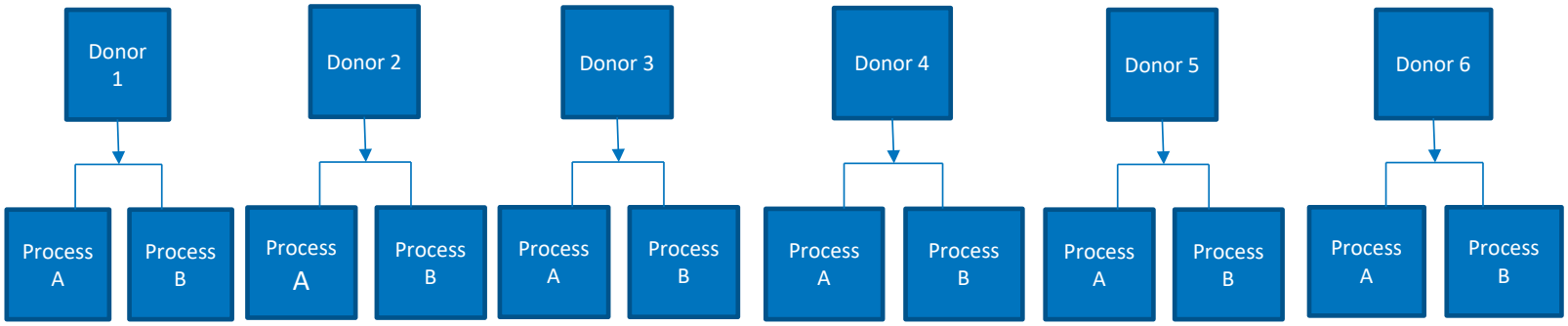
- Two-sample Equivalence Test of Means
  - Population standard deviation known or not?
  - Equivalence, Superiority, or Non-inferiority?

- Currently does not address the challenges associated with with a split apheresis design (paired t-tests)
  - Sample size and power calculations
  - Overlay different scenarios for comparison
  - Dependendent of the proportion of donor-to-donor variability (ρ)

DSURGO
ascend with analytics

Equivalence Tests for CGTs - Paired

- Split apheresis design
  - Used in CGTs to partition out variation among donors
  - Same donor material across process A & B

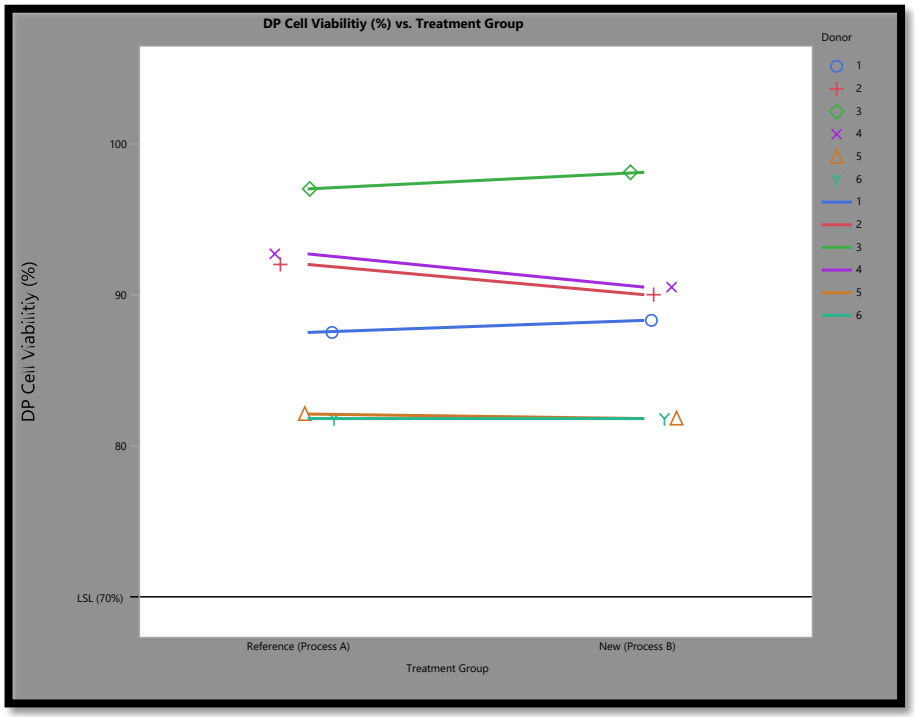


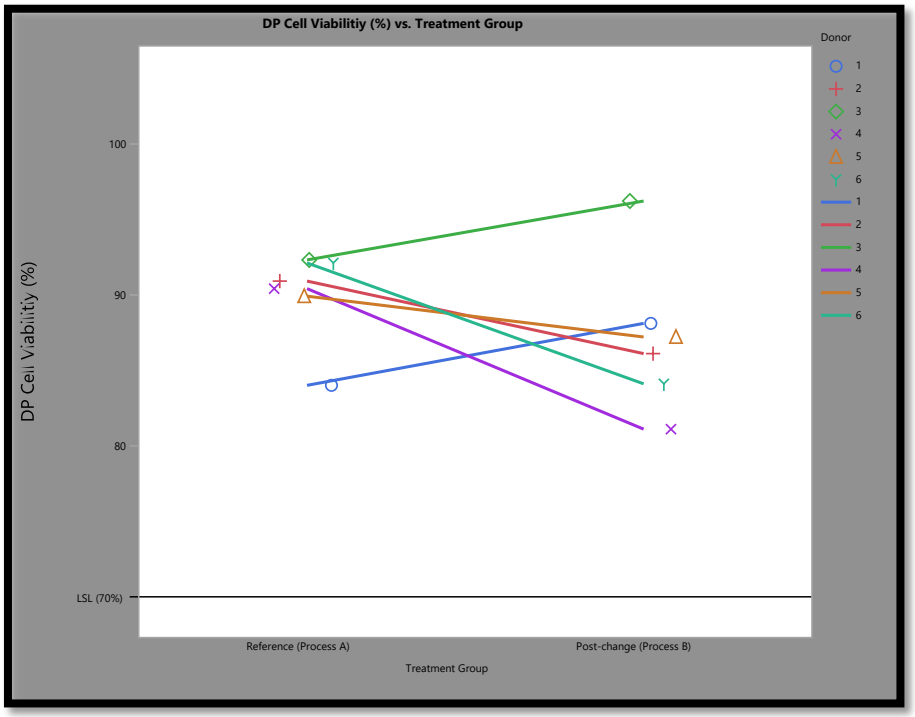
"For studies that compare two cellular manufacturing processes using the split-donor starting material, the product data from each donor are paired. In such cases, you should select a statistical test suitable for analysis of the difference between paired data, rather than a test that assumes an independent data structure."

- "Draft Guidance for Industry Manufacturing Changes and Comparability for Human Cellular and Gene Therapy Products" July 2023.

DSURGO
ascend with analytics

Donor-to-donor Variability





$$\rho = \frac{\sigma^2_{Donor}}{\sigma^2_{\text{Analytical+Process}} + \sigma^2_{\text{Donor}}}$$

DSURGO
ascend with analytics

JMP Scripts – Power and Sample Size

- Paired approach - Known versus Unknown Standard Deviation
- Is the Type I and Type II error rates (Power) affected by the proportion of donor-to-donor variability (ρ)?
- Different values of paired lots (n2) and *k* values (used to calculate EAC).
- How the standard deviation is calculated for the EAC (*d*)?
- How does the Paired approach compare to the Independent approach for varying levels of ρ?

DSURGO
ascend with analytics

Operating Characteristic (OC) or Power Curves

- EMA 2021: The evidence needed to declare comparability must be based on expected operating characteristics (OC curves).

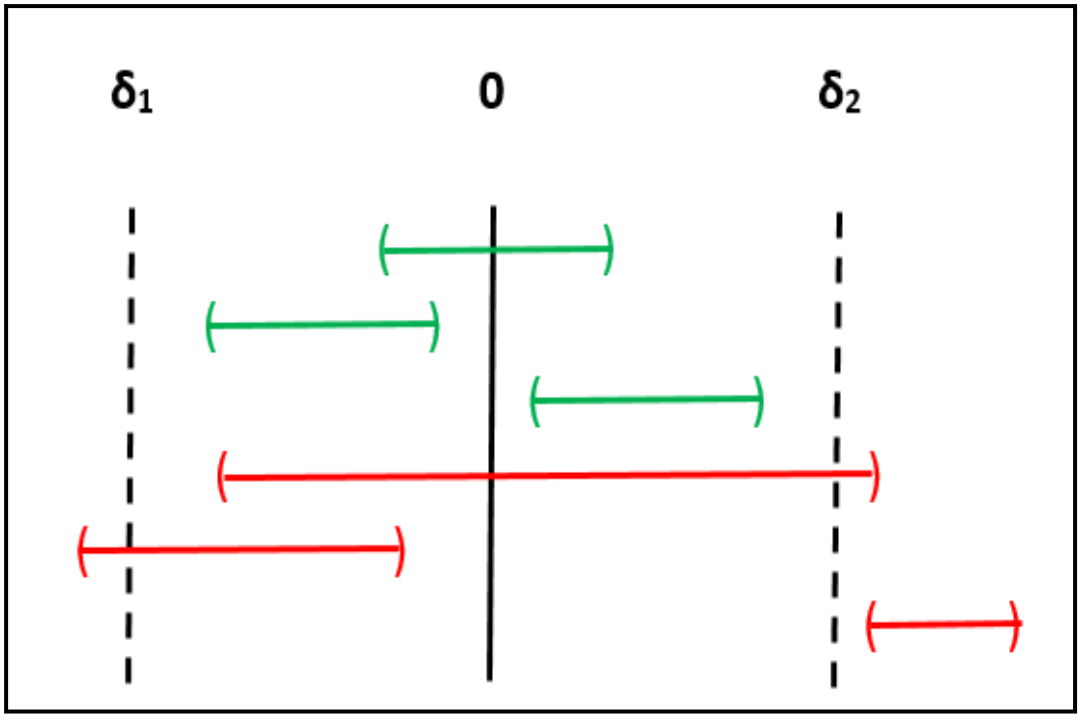
- Power - The probability of meeting the comparability criterion for a given true value of capability.

Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development EMA/CHMP/138502/2017, 26 July 2021, Committee for Medicinal Products for Human Use (CHMP).

DSURGO
ascend with analytics

Equivalence Margin

– Selection of the equivalence acceptance criteria (d) is the key to the outcome of similarity.



Reference: Tsong, Yi, and OB CMC Analytical Biosimilar Method Development Team (Meiyu Shen, Cassie Xiaoyu Dong). 2015. *Development of Statistical Approaches for Analytical Biosimilarity Evaluation* [PowerPoint]. DIA/FDA Statistics Forum.

DSURGO
ascend with analytics

JMP Scripts – Power and Sample Size

- Paired approach - Known versus Unknown Standard Deviation
- Is the Type I and Type II error rates (Power) affected by the proportion of donor-to-donor variability (ρ)?
- Different values of paired lots (n2) and *k* values (used to calculate EAC).
- How the standard deviation is calculated for the EAC (*d*)?
- How does the Paired approach compare to the Independent approach for varying levels of ρ?

DSURGO
ascend with analytics

Power of Split-apheresis design Depends on Proportion of Variance due to Donor

- Power of split-design (paired design) depends on r where

$$\rho = \frac{\sigma^2_{Donor}}{\sigma^2_{\text{Analytical+Process}} + \sigma^2_{\text{Donor}}}$$

- If a CQA has a low ρ, the paired design:
  - will be under-powered with fewer than 8 lots.
  - is no more efficient than an independent design.

DSURGO
ascend with analytics

Leveraging Historic and/or Independent Data Sets

- Practicality dictates no more than 3-6 paired lots can be produced in the best of circumstances.

- If ρ is small, this design will not provide sufficient power to perform a statistical test of equivalence with $\alpha=0.05$ and EAC = 2 or less.

- Thus, sufficient data from an independent design (including representative historic lots) should be used to supplement the split-apheresis design.

DSURGO
ascend with analytics

Any questions?

Q & A

DSURGO
ascend with analytics