Clinical trial optimization approaches to Phase III trials with multiple objectives

Gautier Paux [1] Alex Dmitrienko [2]

[1] Institut de Recherches Internationales Servier

[2] Mediana Inc.

Trend and Innovation in Clinical Trial Statistics
May 03, 2016

Context

General goal

- Clinical trials (CT) should be designed to ensure a high probability to detect an effect if the treatment is effective

Sample size calculations in traditional setting

- Traditional CT with two arms, a single endpoint and no interim looks
- Sample size calculations can be done using closed-form expression
- Example of sample size per treatment arm for normally distributed endpoints :

$$n = \frac{2(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2}$$

Context

Sample size calculations in complex setting

- CT sponsors are often interested in pursuing multiple objectives in Phase II or Phase III clinical trials such as :
  - Multiple dose-control comparisons
  - Multiple endpoints
  - Multiple patient populations
  - Interim looks and adaptations
- General analytic expressions of the power function do not exist in this case

Context

Goal

- How to evaluate power in clinical trials with complex clinical objectives ?

FDA Enrichment strategies for CT

- *Determining the required sample size that will provide reasonable power to test the different hypotheses while controlling type-I error [...] is challenging.*

Clinical Scenario Evaluation framework

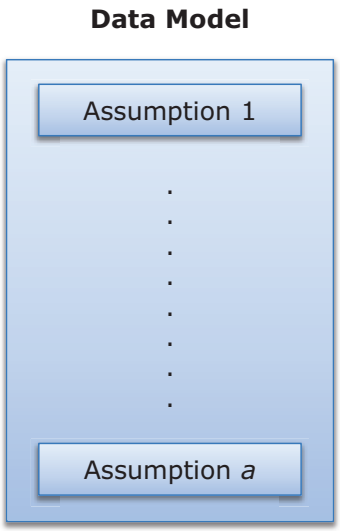
Overview

- Developed in Benda *et al.* (2010) and Friede *et al.* (2010)
- Decompose the problem of clinical trial simulations into three components

Objectives

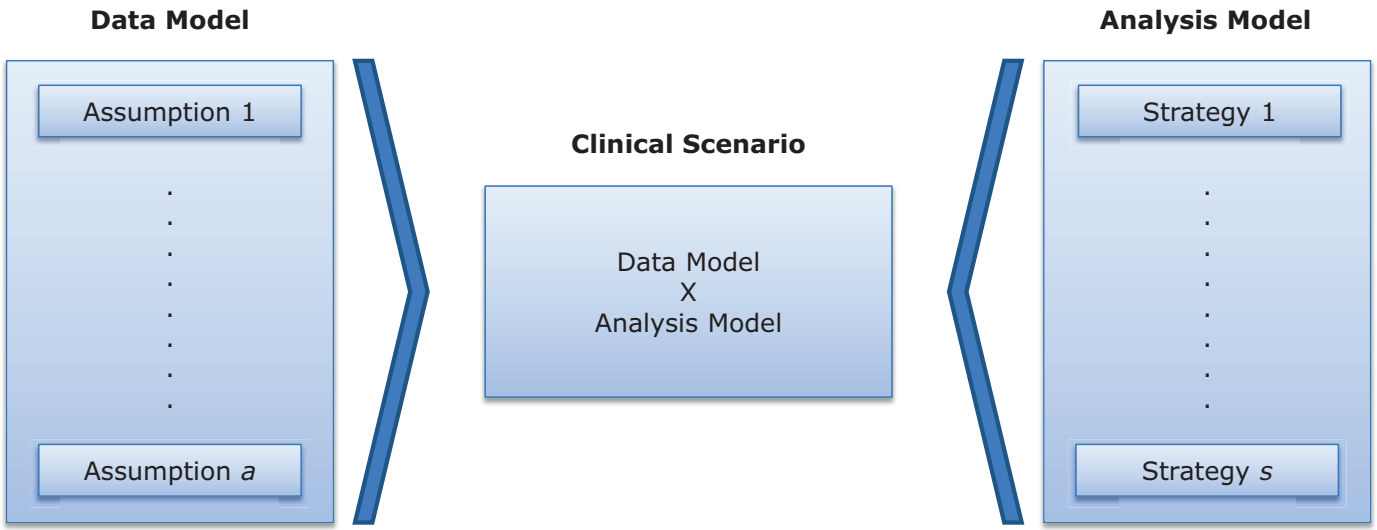
- Systematic simulation-based assessment of study designs and analysis methods
- Selection of a robust approach to clinical trial design and analysis which demonstrates optimal performance
- Sensitivity assessment of key parameters

Clinical Scenario Evaluation framework

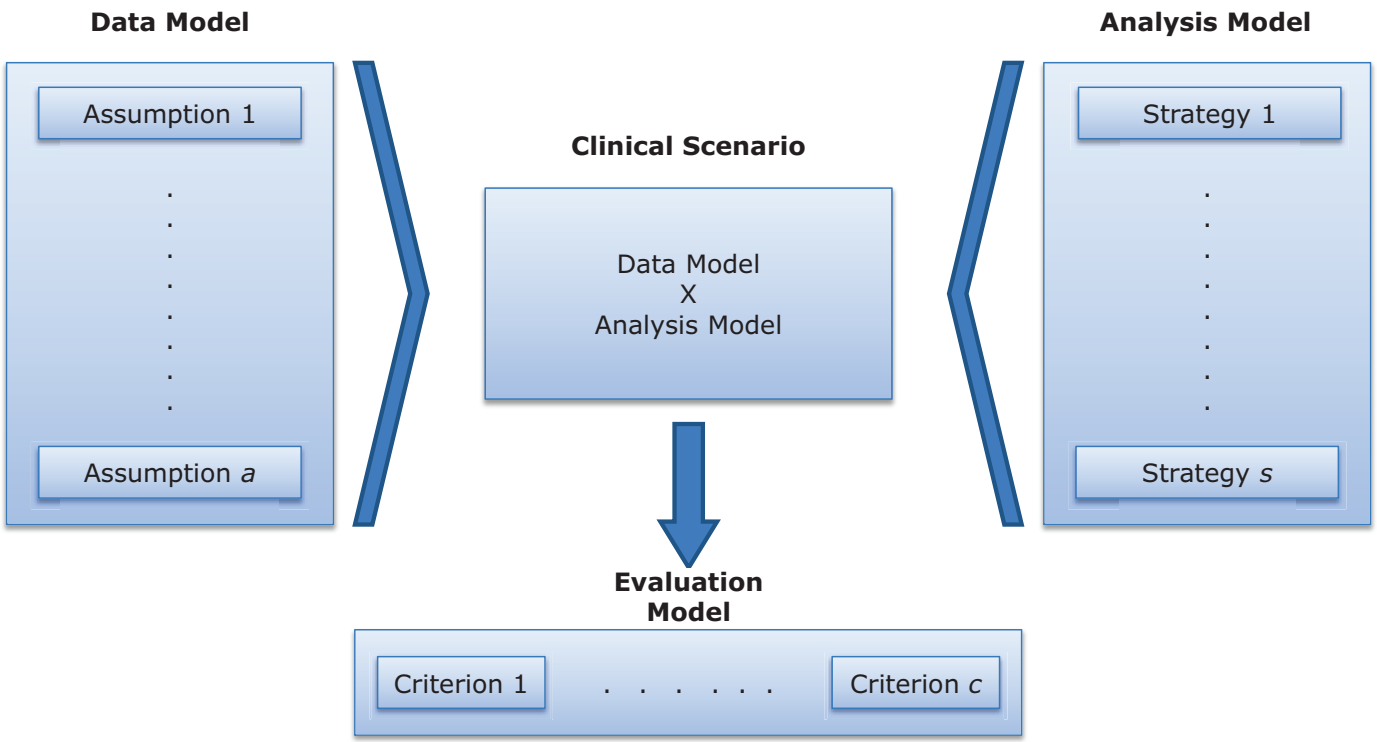


Clinical Scenario Evaluation framework

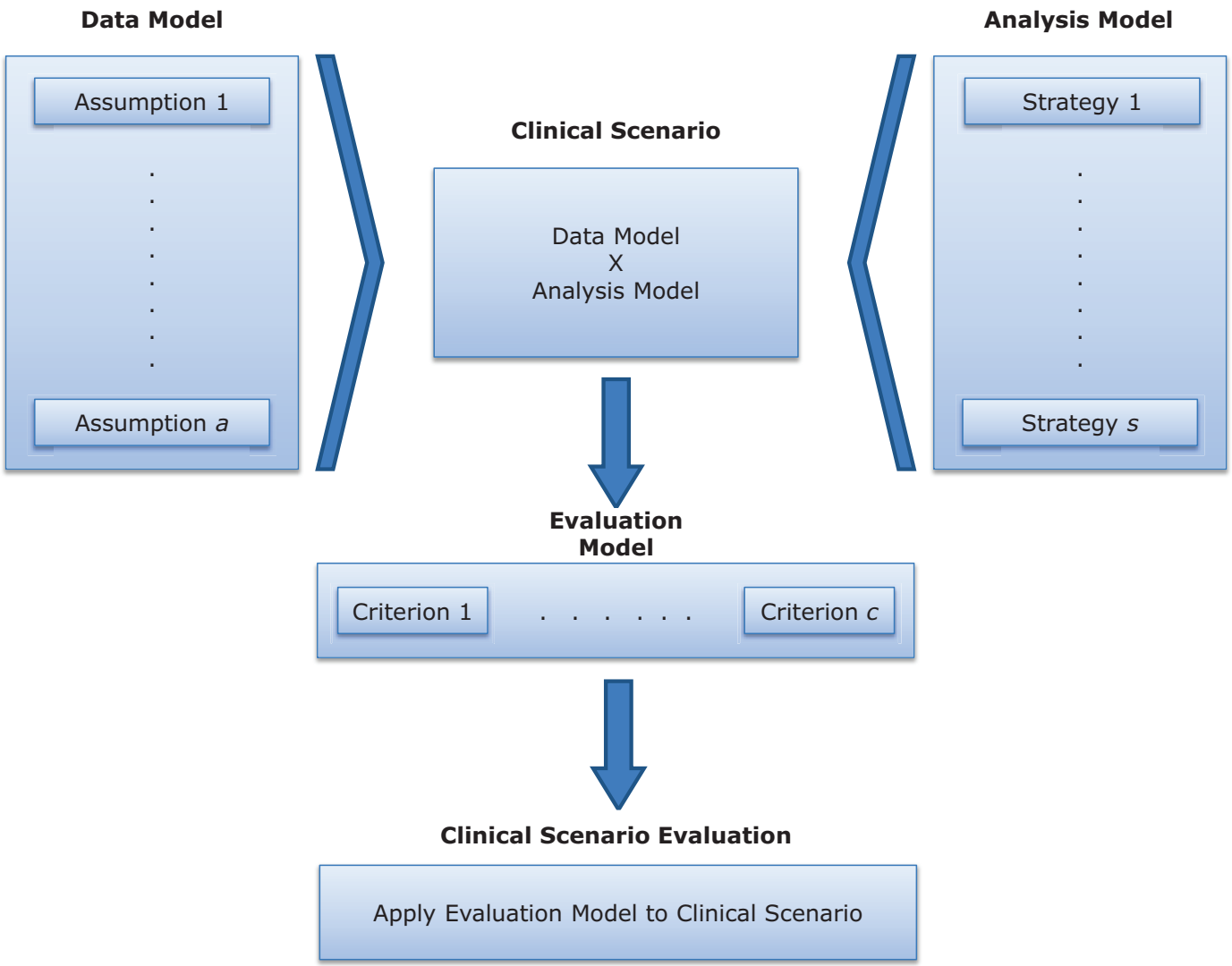
Clinical Scenario Evaluation framework



Clinical Scenario Evaluation framework



Clinical Scenario Evaluation framework



Clinical Scenario Evaluation framework

Key components

- Data model is the structure for describing the data generation parameters
- Analysis model is the structure for defining the analysis strategies applied to the data
- Evaluation model is the structure for specifying the measures for evaluating the performance of the analysis strategies

Clinical trial optimization

Optimization criterion

- Crucial to choose a relevant criterion, aligned with clinical objectives

Optimal parameters

- Important to assess the performance of analysis models under optimal and nearly optimal configuration of parameters

Sensitivity assessment

- Critical to ensure that optimization is robust to reasonable deviations from original assumptions

Case study

MDD clinical trial

- Phase III clinical trial in patients with Major Depressive Disorder (MDD)
- Design : three doses of new treatment (Dose L, Dose M and Dose H) versus placebo
- Trial objective : demonstrate that at least one dose is effective
- Primary endpoint : change from baseline to end of treatment in MADRS score
- Reference : Dmitrienko, Paux and Brechenmacher (2015)

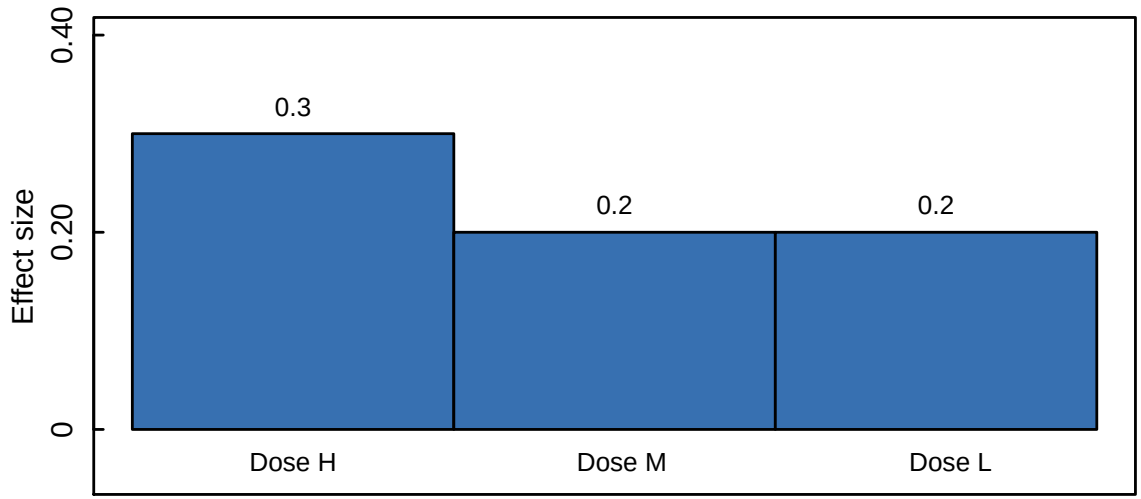
Data Model

Sample size

- Balanced groups
- *n* per group is set to 200

Outcome distribution

- Primary endpoint is normally distributed
- Expected effect sizes :



Scenario 1

Analysis model

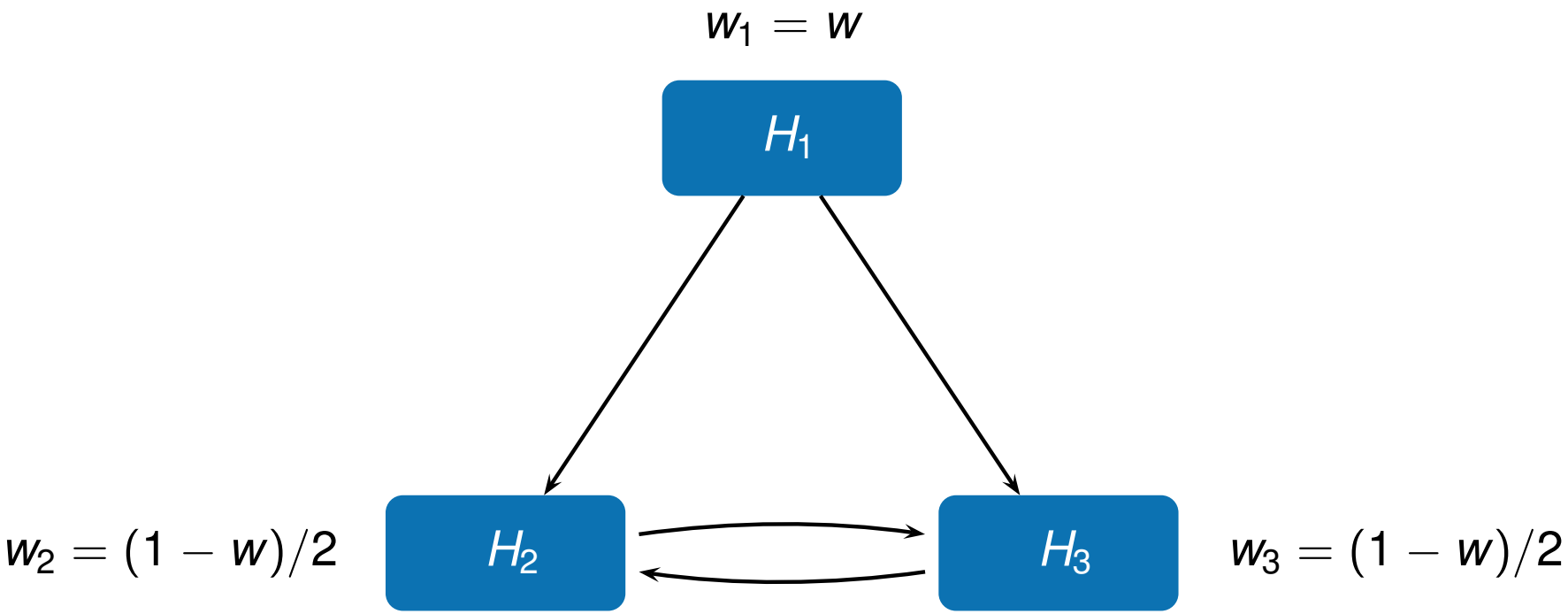
Tests

- H_1, null hypothesis of no effect for Dose H versus Placebo
- H_2, null hypothesis of no effect for Dose M versus Placebo
- H_3, null hypothesis of no effect for Dose L versus Placebo
- Null hypotheses are tested with Student t-tests

Multiple testing procedure

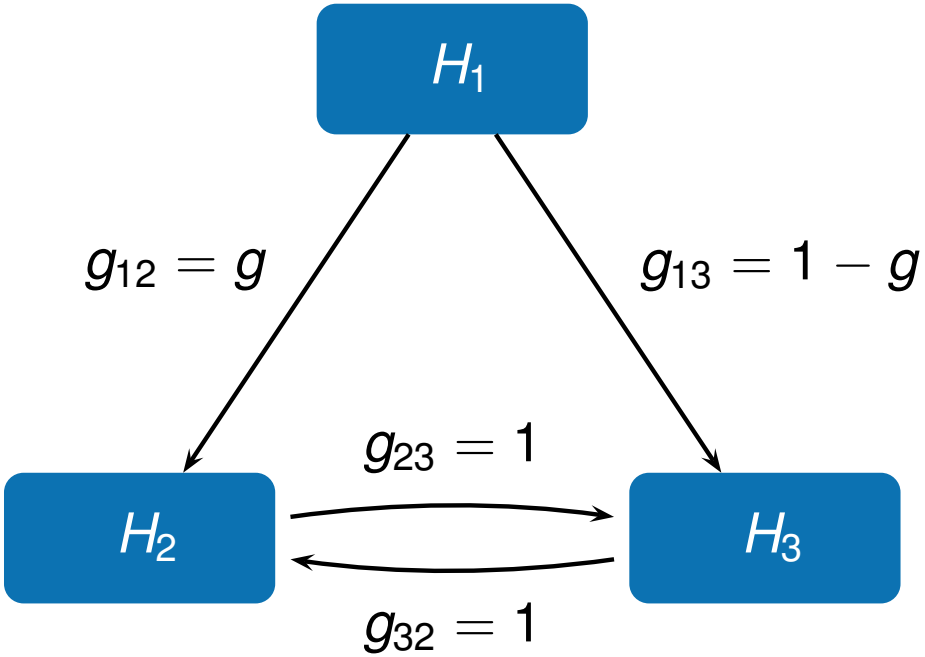
- Bonferroni-based chain procedure to control the Type I error rate
  - α allocation rule
  - α propagation rule

α allocation rule



Weight for H_1 is set to w and the remaining weight $(1 - w)$ is split between H_2 and H_3

α propagation rule



Error rate is split between H_2 and H_3 after H_1 is rejected

Error rate is transferred between H_2 and H_3

Evaluation model

Disjunctive power

- Probability to reject at least one hypothesis

$$\psi_D(w, g) = P(\text{Reject } H_1 \text{ or } H_2 \text{ or } H_3)$$

Weighted power

- Weighted sum of the probability to reject each hypothesis

$$\psi_W(w, g) = \sum_{i=1}^{3} v_i P(\text{Reject } H_i)$$

- v_i ($\sum_{i=1}^{3} v_i = 1$), relative importance of a significant treatment effect at dose i

Optimal selection of procedure parameters

Exhaustive search

- Explore multiple sets of parameters of chain procedure (hypotheses weight w and transition parameter g) to maximize an appropriate optimization criterion

Sensitivity assessment

- Evaluate the impact of random deviations from the initial dose-response assumptions on the performance of optimal chain procedure

Optimal selection of procedure parameters

Exhaustive search - Disjunctive power

- When w is close to 1, H_1 serves as a gatekeeper for H_2 and H_3
- $\psi_D(w, g)$ is equivalent to $P(\text{Reject } H_1)$
- $\psi_D(w, g)$ is mostly driven by w and unaffected by g

Transition parameter \ Hypothesis weight	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.9	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.8	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.7	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.6	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.5	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.4	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.3	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.2	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0.1	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**
0	71.50	75.60	78.10	80.10	81.30	82.40	83.10	83.90	84.60	**84.90**

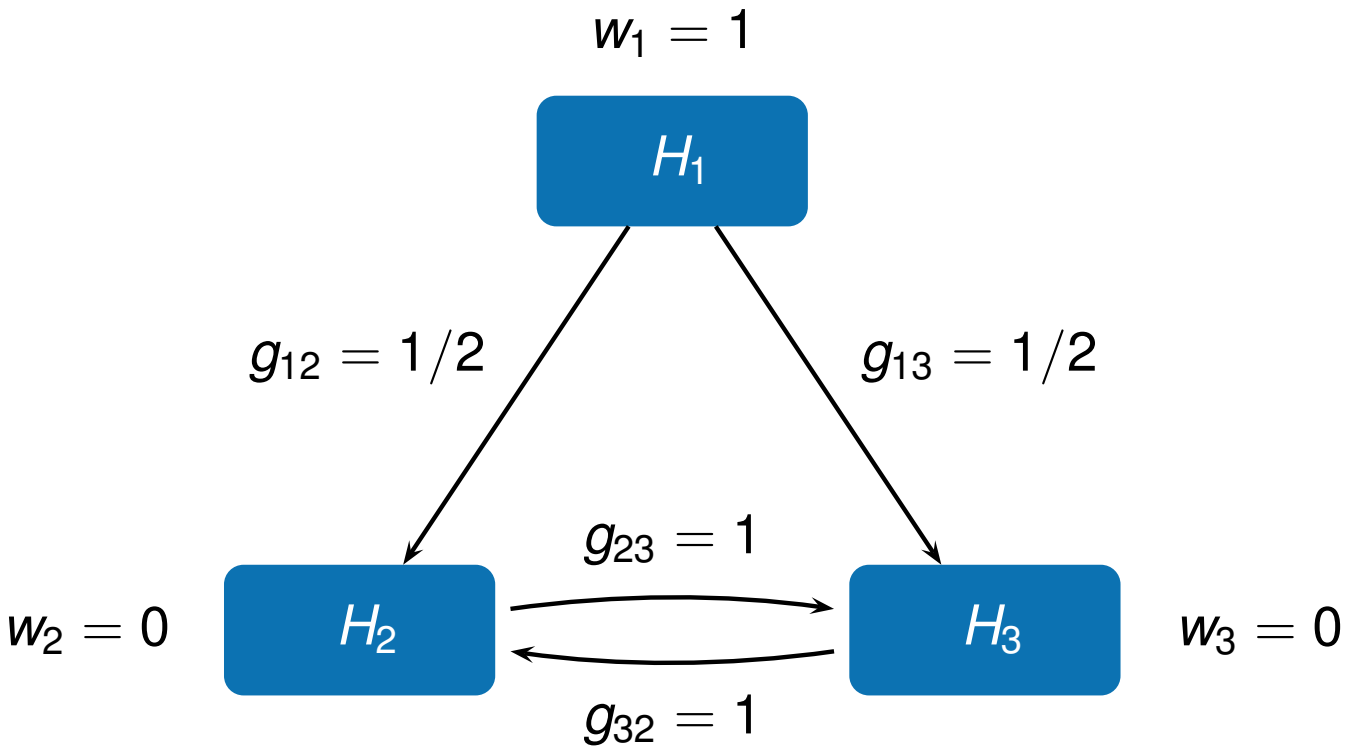
Optimal selection of procedure parameters

Exhaustive search - Weighted power

- $\psi_W(w, g)$ is maximized for several combinations of the w and g parameters
- Example : $w = 1$ and $g = 0.5$

Transition parameter \ Hypothesis weight	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	48.30	51.20	52.70	53.80	54.40	55.00	55.30	55.60	55.80	55.50
0.9	48.30	51.20	52.70	53.90	54.60	55.10	55.40	55.80	56.10	56.20
0.8	48.40	51.10	52.70	53.90	54.60	55.20	55.60	55.90	56.30	56.40
0.7	48.40	51.20	52.70	53.80	54.60	55.20	55.50	56.00	56.30	**56.60**
0.6	48.40	51.20	52.70	53.90	54.60	55.20	55.60	56.00	56.30	56.50
0.5	48.40	51.20	52.70	53.90	54.60	55.20	55.60	56.00	56.40	**56.60**
0.4	48.40	51.30	52.80	54.00	54.70	55.30	55.70	56.10	56.40	**56.60**
0.3	48.40	51.30	52.80	53.90	54.60	55.20	55.60	55.90	56.30	56.50
0.2	48.40	51.20	52.70	53.90	54.50	55.10	55.50	55.80	56.20	56.30
0.1	48.40	51.20	52.70	53.80	54.50	55.00	55.30	55.70	56.00	56.10
0	48.40	51.20	52.60	53.80	54.40	54.90	55.20	55.50	55.70	55.30

Optimal selection of procedure parameters



Test H_1 at full α-level and split the error rate equally between H_2 and H_3 after H_1 is rejected

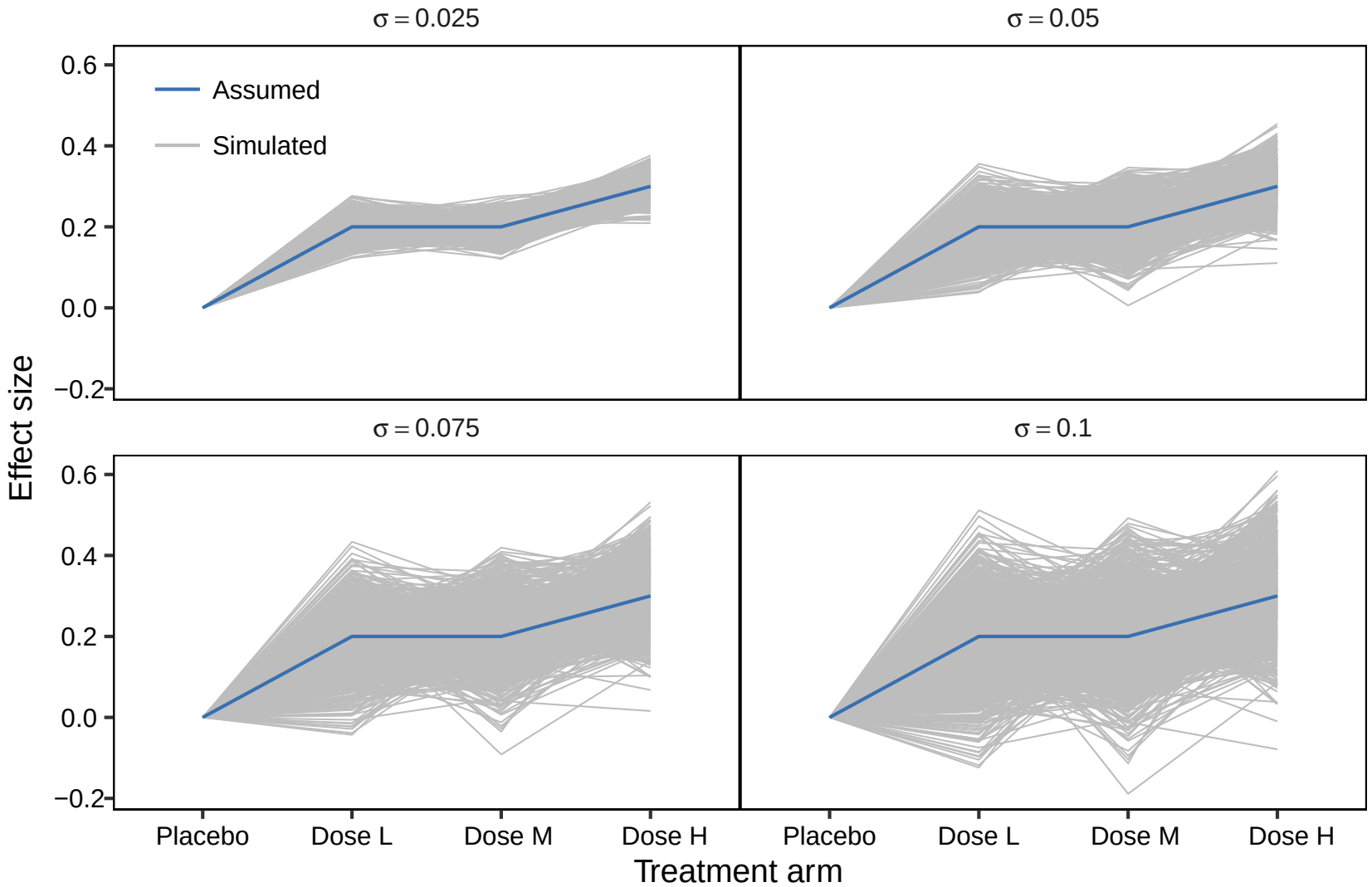
Sensitivity assessment

Algorithm

- $\theta = (\theta_1, \theta_2, \theta_3)$, assumed dose-response function
- Step 1 : Generate k "true" dose-response functions, i.e., $\theta_i^* = (\theta_{i1}^*, \theta_{i2}^*, \theta_{i3}^*)$, $i = 1, \ldots, k$, where $\theta_{ij}^* \sim N(\theta_{ij}, \sigma)$ and σ quantifies the amount of random perturbation
- Step 2 : Compute weighted power for optimal chain procedure using each true dose-response function (10,000 simulation runs)
- Step 3 : Summarize weighted power over k dose-response functions
  - Distribution of weighted power
  - Performance loss

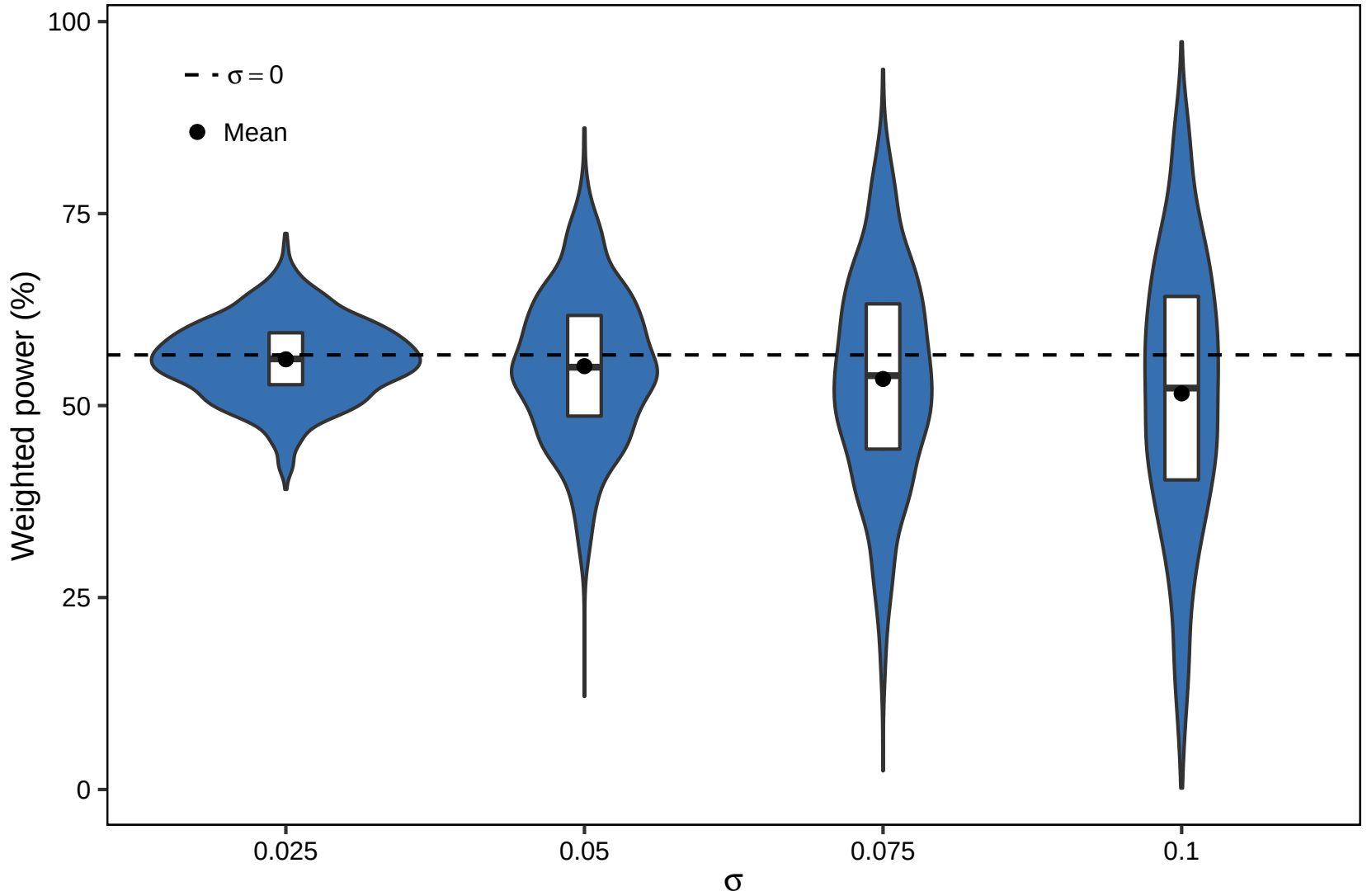
Sensitivity assessment

Simulated dose-response functions



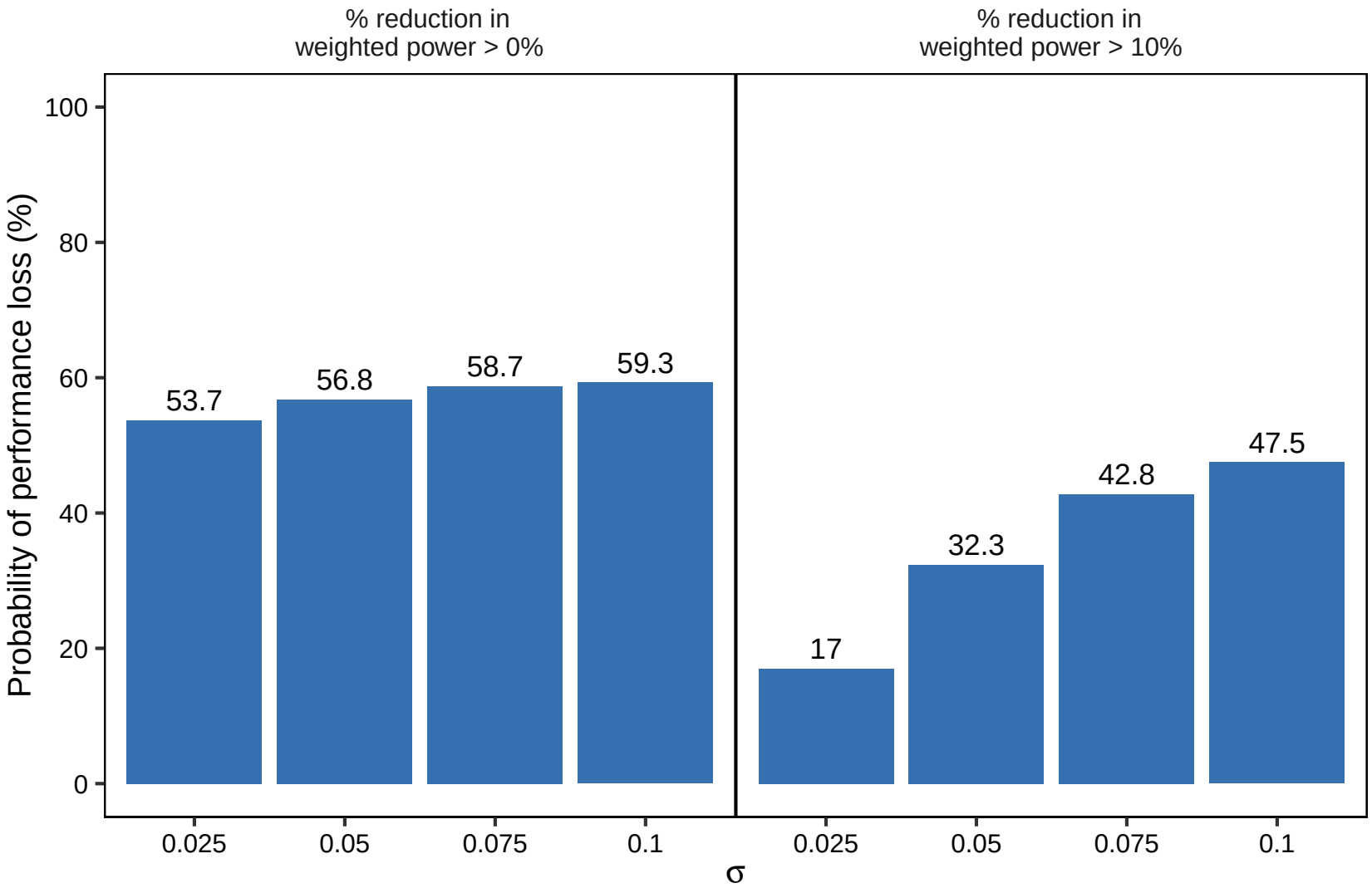
Sensitivity assessment

Weighted power distribution



Sensitivity assessment

Probability of performance loss



Case study

Key messages

- A range of sample sizes should be evaluated
- Several expected dose-response functions should be considered
- Irrelevant optimization criterion could lead to incorrect choice of optimal multiple testing procedure
- Important to assess the robustness of optimal procedure's performance by randomly perturbing the assumed dose-response functions

Mediana R package

Objective

- To provide a standard framework for clinical trial simulations typically performed in Phase II or Phase III trials
- To create best practices for clinical trial simulations
- To create reproducible simulation-based calculations

Overview

- Based on the Clinical Scenario Evaluation approach
- Supports a broad class of data, analysis and evaluation models
- Flexible framework easily extensible to define custom options in data, analysis and evaluation models
- High-performance computing

Mediana R package

Release

- First version was released on CRAN in Q3 2015 (currently v1.0.3)
- Dmitrienko, A., Paux, G., Brechenmacher, T. (2016). Power calculations in clinical trials with complex clinical objectives. *Journal of the Japanese Society of Computational Statistics*. 28, 15–50.
- Dmitrienko, A., Paux, G., Pulkstenis, E., Zhang, J. (2016). Tradeoff-based optimization criteria in clinical trials with multiple objectives and adaptive designs. *Journal of Biopharmaceutical Statistics*. 26, 120–140.

New features for next version

- Support to adaptive designs
- Interim analysis decision rules for futility or overwhelming efficacy

Key messages

Take home messages

- New drug development is a time-consuming and expensive process
- Need for more efficient drug development programs with innovative designs and advanced analysis strategies
- Crucial role of quantitative assessments of the performance of these designs and analysis strategies
- Mediana R package provides a turnkey solution to facilitate systematic quantitative assessment of performance for Phase II and III trial designs and analysis methods.

Thank you

Contact information

- Gautier Paux : gautier.paux@servier.com
- Alex Dmitrienko : admitrienko@medianainc.com

Website

- Mediana package website

  `http://gpaux.github.io/Mediana/`

- Mediana package on CRAN (latest released version)

  `https://cran.r-project.org/web/packages/Mediana/index.html`

- Mediana package on GitHub (latest development version)

  `https://github.com/gpaux/Mediana`

References

- Benda, N., Branson, M., Maurer, W., Friede, T. (2010). Aspects of modernizing drug development using clinical scenario planning and evaluation. *Drug Information Journal*. 44, 299–315.
- Dmitrienko, A., Paux, G., Brechenmacher, T. (2016). Power calculations in clinical trials with complex clinical objectives. *Journal of the Japanese Society of Computational Statistics*. 28, 15–50.
- Dmitrienko, A., Paux, G., Pulkstenis, E., Zhang, J. (2016). Tradeoff-based optimization criteria in clinical trials with multiple objectives and adaptive designs. *Journal of Biopharmaceutical Statistics*. 26, 120–140.
- Friede, T., Nicholas, R., Stallard, N., Todd, S., Parsons, N.R., Valdes-Marquez, E., Chataway, J. (2010). Refinement of the clinical scenario evaluation framework for assessment of competing development strategies with an application to multiple sclerosis. *Drug Information Journal* 44 :713–718.