Patient-selected and patient-ranked outcomes in clinical trials

PSI 2026

David Robertson

david.robertson@mrc-bsu.cam.ac.uk

Introduction

- Growing emphasis on patient-focused drug development, e.g. FDA (2018–2023)
  - Need to measure outcomes of importance to patients in clinical trials
- **Challenge:** Medical conditions affect multiple symptom domains $\rightarrow$ multiple outcomes
- Traditionally, stakeholders (including patients) come to a consensus around the relative importance of these different outcomes $\rightarrow$ selection of (co-)primary and secondary endpoints
- This ignores individual patient preference heterogeneity.

Case Study: FDA 2018 Migraine Guidance

A regulatory precedent for using individual patient preferences in practice

- **Historical standard:** demonstration of effect on four co-primary endpoints: pain, nausea, photophobia, *and* phonophobia

Case Study: FDA 2018 Migraine Guidance

A regulatory precedent for using individual patient preferences in practice

- **Historical standard:** demonstration of effect on four co-primary endpoints: pain, nausea, photophobia, *and* phonophobia
- **New approach**: each patient is able to *select* the symptom that matters most to them $\rightarrow$ two co-primary endpoints
  1. Freedom from headache pain at 2 hours.
  2. Effect on the *Most Bothersome Symptom (MBS)* out of nausea, photophobia and phonophobia at 2 hours

Incorporating patient preferences

1. **Patient-Selected Outcomes** (Top-ranked): Analysis focuses on the single outcome selected by the patient as most important

2. **Patient-Ranked Outcomes**: Analysis considers the full hierarchy of outcome importance for each patient

Methods

- Two-arm RCT comparing an experimental treatment ($k = 1$) against a control ($k = 0$) with a total of $n = n_0 + n_1$ patients
- $\boldsymbol{Y_i} = (Y_{i,1}, \cdots, Y_{i,m})$ denotes the vector of m outcomes for patient i
- $a_i \in \{0, 1\}$ denotes treatment allocated to patient i
- Assume $Y_{i,j} \sim N(\mu_{a_i,j}, \sigma^2_{a_i})$ independently

Baseline Method: Standard Univariate Analysis

The standard approach that ignores individual preferences: single outcome j^* is chosen *a priori* as the primary outcome for *all* patients

- **Test Statistic:** Welch t-test comparing means of the chosen outcome j^* between groups
- **Hypothesis:** $H_0 : \mu_{0,j^*} = \mu_{1,j^*}$ vs $H_A : \mu_{1,j^*} > \mu_{0,j^*}$

Method 1: Mean Patient-Selected Outcome

Let s_i be the outcome selected by patient i. We analyse Y_{i,s_i}

- **Assumption:** Selection s_i is determined prior to randomization
- **Test Statistic:** Welch t-test comparing means of selected outcomes
  - To ensure valid null distribution, sufficient condition is to use stratified randomisation by selected outcome so that $n_{1,j}/n_1 = n_{0,j}/n_0$
- **Hypothesis:** $H_0 : \mu_{0,j} = \mu_{1,j}$ vs $H_A : \mu_{1,j} \geq \mu_{0,j}$ for all j

Method 2: Patient ranked outcomes

Composite Desirability of Outcome Ranking (DOOR) approach (Lu et al., 2022)

- Uses full ranking vector $\boldsymbol{R_i} = (R_{i,1}, \cdots, R_{i,m})$
- Pairwise comparison of every treated patient vs every control patient
- **Algorithm:**
  1. Compare set of outcomes that share the same set of rankings, starting from the smallest possible set to the largest
  2. Patient A 'wins' if all outcomes in set are tied or better than for Patient B

Method 2: Patient ranked outcomes

Composite Desirability of Outcome Ranking (DOOR) approach (Lu et al., 2022)

- Uses full ranking vector $\boldsymbol{R_i} = (R_{i,1}, \cdots, R_{i,m})$
- Pairwise comparison of every treated patient vs every control patient
- **Algorithm:**
  1. Compare set of outcomes that share the same set of rankings, starting from the smallest possible set to the largest
  2. Patient A 'wins' if all outcomes in set are tied or better than for Patient B
- **Estimand** (θ): Probability that, for a randomly selected pair of participants, the patient on the treatment arm has a more favourable composite DOOR than the patient on the control arm ('winning' probability')
- **Test statistic:** $\hat{\theta}$ uses an asymptotic normal approximation
- **Hypothesis:** $H_0 : \theta = 0.5$ versus $H_A : \theta > 0.5$

Method 3: Weighted Winning Probability (WWP)

A Composite DOOR approach only using top-ranked data (Lu et al., 2022)

- **Strategy:**
  1. Stratify patients based on their top-ranked outcome j
  2. Calculate winning probability $\hat{\theta}_j$ within each stratum
  3. Compute weighted average based on stratum size $\hat{p}_j$:

$$\hat{\theta}_{WT} = \sum_{j=1}^{m} \hat{p}_j \hat{\theta}_j$$

Simulation study

- Multiple Sclerosis (MS) trial comparing cognitive and behaviour therapy (CBT) against standard of care
- Three clinical outcomes measured using PROMIS scoring:
  1. Fatigue
  2. Pain
  3. Depression
- Primary outcome = normalised reduction in PROMIS scores from baseline after one year

Simulation study

- Multiple Sclerosis (MS) trial comparing cognitive and behaviour therapy (CBT) against standard of care
- Three clinical outcomes measured using PROMIS scoring:
  1. Fatigue
  2. Pain
  3. Depression
- Primary outcome = normalised reduction in PROMIS scores from baseline after one year
- Prior to randomisation a patient survey elicited patient preferences of relative importance of outcome

$$(p_{123}, p_{132}, p_{213}, p_{231}, p_{312}, p_{321}) = (0.42, 0.17, 0.24, 0.05, 0.08, 0.04)$$

Simulation Scenarios

$n = 60$ with $n_0 = n_1 = 30$, moderate correlation structure assumed between outcomes

- **S1: Null Scenario**
  No treatment effect on any outcome
- **S2: Uniform Improvement**
  CBT uniformly improves all 3 outcomes
- **S3: Fatigue Only**
  CBT improves Fatigue only (ranked top by 59%)
- **S4: Depression Only**
  CBT improves Depression only (ranked top by 12%)
- **S5:** CBT improves fatigue by 1, depression by 0.5, no effect for pain
- **S6:** CBT improves top-ranked (= patient-selected) outcome by 1, second-ranked outcome by 0.5 and no effect for bottom-ranked outcome
- **S7:** CBT improves bottom-ranked outcome by 1, second-ranked outcome by 0.5 and no effect for top-ranked (= patient-selected) outcome no effect on depression

Results

Power (%) under real-world MS preference distribution

Method	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**
UV1 (Fatigue)	4.7	98.7	98.7	4.7	98.7	97.4	7.7
UV2 (Pain)	5.0	98.4	5.0	5.0	5.0	45.0	70.1
UV3 (Depression)	4.9	98.5	4.9	98.5	60.0	13.6	94.3
Mean Patient-Selected	4.8	98.7	68.5	8.8	74.8	81.6	28.0
Composite DOOR	6.0	99.8	59.9	26.4	76.6	80.9	59.2
WWP	6.7	98.0	65.6	7.9	71.2	78.4	25.5

Discussion

- Where majority of patients have same top-ranked (i.e., selected) outcome, mean patient-selected outcome can be competitive with composite DOOR in terms of power
- Mean patient-selected outcome has uniformly higher power than the WWP

Discussion

- Where majority of patients have same top-ranked (i.e., selected) outcome, mean patient-selected outcome can be competitive with composite DOOR in terms of power
- Mean patient-selected outcome has uniformly higher power than the WWP
- Composite DOOR relies on full preference ranking, and has implicit assumption of equal 'distance' in importance between each rank
- By ignoring outcomes that are not the top-ranked/selected one, patient-selected outcomes could miss meaningful differences in treatment effects
- Flip side is that increased power of composite DOOR can be driven by improvements on outcome that (majority of) patients rank as least important

Discussion

- Use of patient-selected outcomes has demonstrated regulatory acceptability
- Patient-selected outcomes have a much simpler interpretation than composite DOOR or WWP
- Much scope for further development of patient-selected outcomes in clinical trials
- For more discussion and results, see arXiv preprint at `https://arxiv.org/abs/2510.11578`

References

US Food and Drug Administration (2025)
FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making

US Food and Drug Administration (2018)
Migraine: Developing Drugs for Acute Treatment Guidance for Industry
https://www.fda.gov/regulatory-information/search-fda-guidance-documents/migraine-developing-drugs-acute-treatment

Y. Lu, Q. Zhao, J. Zou, S. Yan, J. S. Tamaresis, L. Nelson, X. M. Tu, J. Chen, and L. Tian (2022)
A composite endpoint for treatment benefit according to patient preference
Statistics in Biopharmaceutical Research, 14(4):408–422.