A new Bayesian adaptive decision-theoretic design for multi-arm multi-stage clinical trials illustrated by an application in exercise oncology

Peter van de Ven
Department of Data Science and Biostatistics
University Medical Center Utrecht
Disclosures

• I am seconded for 0.4 FTE to the Dutch Medicines Evaluation Board (CBG-MEB) as a methodology assessor
• I do not have any disclosures that are relevant for the presented work
• I here present my personal views
Outline

• Motivation
• Outline of the methodology
• Simulation studies and results
• Retrospective application to a trial in exercise oncology
• Concluding remarks
Motivation

- Rapid increase in available treatments (in development/approved/used in practice)
- A large proportion of phase III trials is negative (in oncology approximately 65%\(^1\))
- Long timespan until approval of new drugs (10-15 years)
- Most phase II trials are single-arm trials and most phase III trials are two-arm trials
- How to compare effectiveness of different treatments that are all standard of care?

There is a clear need for more efficient trials that compare multiple treatment options and use decision-criteria that fit a trial’s objective

Outline of methodology: Trial design

Multi-arm multi-stage design: equal randomization to active treatment arms in each stage

Setting 1: Only experimental arms

Randomize

Experimental arm 1

Experimental arm 1

Experimental arm 2

Experimental arm 1

Interim analysis

Continue

Experimental arm 1

Experimental arm 2

Stop

Final decision

Stop

Final decision

Setting 2: Control arm included

Randomize

Control arm

Experimental arm 1

Experimental arm 2

Control arm

Experimental arm 1

Experimental arm 2

Stop

Final decision

Stop

Final decision
Outline of methodology: Decision-theoretic framework

Loss functions
Bassi et al. (SMMR, 2020): 0-1 loss function with loss of 1 for incorrect decisions and loss of 0 for correct decision: expected loss is probability of making an incorrect final decision

Setting 1: Pick-the-winner
Select the experimental treatment with highest response rate

<table>
<thead>
<tr>
<th>Possible final trial decisions</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment 1 the best</td>
<td>1</td>
</tr>
<tr>
<td>Treatment 2 the best</td>
<td>1</td>
</tr>
<tr>
<td>Treatment 3 the best</td>
<td>1</td>
</tr>
<tr>
<td>Treatment 4 the best</td>
<td>0</td>
</tr>
</tbody>
</table>

Setting 2: Pick-all-treatments-superior-to-control
Select all experimental treatments that outperform the control treatment by an absolute margin of $\delta$

<table>
<thead>
<tr>
<th>Possible final trial decisions</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither treatment 1 nor 2 superior to control</td>
<td>1</td>
</tr>
<tr>
<td>Only treatment 1 superior to control</td>
<td>1</td>
</tr>
<tr>
<td>Only treatment 2 superior to control</td>
<td>0</td>
</tr>
<tr>
<td>Treatment 1 and 2 superior to control</td>
<td>1</td>
</tr>
</tbody>
</table>
Outline of methodology: Interim analyses

Model
• Response of subject in treatment arm $j$: Bernoulli distributed with probability $p_j$
• Independent, non-informative uniform priors for $p_j$ ($j = 1, .., T$)

Interim analyses
• Equal number of subjects per stage (except for first)
• One-stage ahead approach, comparing
  • Expected loss in case of stopping now and making a final decision
  • Expected loss in case of continuing for a single stage and making a final decision:
    • Keeping all active arms in the trial
    • After dropping a single treatment arm from the trial
• Based on economic principle of diminishing returns: Continue trial when the reduction in expected loss exceeds a predefined threshold
Simulation studies and results: Experimental arms only

- **Pick-the-winner**: 3 treatment arms
- Threshold for continuation fixed at 1/2500 for design B1
- Designs B2 and B3 matched in terms of (expected) sample size
- 12 subjects per stage

**Design B1**: Bayesian adaptive decision-theoretic design with dropping of arms

**Design B2**: Bayesian adaptive decision-theoretic design without dropping of arms

**Design B3**: Single stage (non-adaptive) Bayesian decision-theoretic design
Simulation studies and results: Comparison to control

- **Pick-all-treatments-superior-to-control**: 2 experimental arms and 0.15 margin
- Threshold tuned for design B2 to have overall type I error of 5% (one-sided testing)
- Frequentist designs F1, F2 and F3 matched to B2 on (expected) sample size
- 12 subjects per stage, 24 in first stage
- Arm 1 is the control arm

**Design B2**: Bayesian adaptive decision-theoretic design without dropping of arms

**Design F1**: Single-stage, Dunnett’s procedure

**Design F2**: Two-stage procedure of Urach and Posch, O’Brien Fleming spending function

**Design F3**: Two-stage procedure of Urach and Posch, Pocock spending function
Simulation studies and results

Simulations showed increased efficiency compared to single- and two-stage designs through:

• Adaptive stopping when probability of an incorrect decision is not expected to sufficiently reduce in the next stage
• Dropping of arms, provided that differences between the arms are large

Simulations (not presented here) further showed that:

• Frequentist two-stage procedures required average trial sizes that were 14%-67% higher (matching proportion of correct decisions)
• In pick-the-winner setting with up to five arms proportions of correct decisions of >80% could be obtained with average trial sizes of 100-150
Retrospective application: PACES trial

- Trial included 230 patients with breast cancer receiving adjuvant therapy
- Compared 2 exercise programmes to usual care

Randomize

- Usual care
- OncoMove: Home based aerobic exercise
- OnTrack: supervised aerobic and resistance exercise

Surgery → Chemotherapy

Endpoint: dose-modifications during chemotherapy

Endpoint: dose-modifications for chemotherapy (yes/no)
Retrospective application: PACES trial

Reanalysis using Bayesian adaptive decision-theoretic method

- 36 patients per stage
- Pick-all-treatments-superior-to-control setting with absolute margin of 0.10
- Total trial size between 108 and 180 depending on threshold for continuation and dropping of arms

<table>
<thead>
<tr>
<th>Threshold for continuation</th>
<th>Number of patients included in the trial</th>
<th>Posterior probability that only OnTrack was superior to UC</th>
<th>Average trial size</th>
<th>Type 1 error probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without arm dropping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>108</td>
<td>0.777</td>
<td>139</td>
<td>9.7%</td>
</tr>
<tr>
<td>0.001</td>
<td>144</td>
<td>0.810</td>
<td>195</td>
<td>5.3%</td>
</tr>
<tr>
<td>0.0001</td>
<td>180</td>
<td>0.881</td>
<td>250</td>
<td>2.5%</td>
</tr>
<tr>
<td>With arm dropping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>72</td>
<td>0.713</td>
<td>126</td>
<td>18.4%</td>
</tr>
<tr>
<td>0.001</td>
<td>144</td>
<td>0.814</td>
<td>174</td>
<td>16.3%</td>
</tr>
<tr>
<td>0.0001</td>
<td>144</td>
<td>0.814</td>
<td>216</td>
<td>14.3%</td>
</tr>
</tbody>
</table>
Concluding remarks

• We introduced a general Bayesian-adaptive decision-theoretic framework for multi-arm multi-stage trials
• We focused on binary loss functions, focusing on the posterior probability of a correct final decision
• Control of type I error possible, but requires tuning of threshold in combination with sample size for first stage (latter in presence of dropping of arms)
• Efficiency shown in various settings and scenarios
• Currently applied in multi-arm AMICO trial: Aerobic fitness or Muscle mass training to Improve Colorectal cancer Outcome
Acknowledgements

The presented work is joint work with:

- prof. Hans Berkhof
- Andrea Bassi, MSc
- dr. Laurien Buffart
- prof. Daphne de Jong

Principal investigators of the PACES trial are:

- prof. Neil Aaronson
- dr. Martijn Stuiver
- prof. Gabe Sonke

Email: p.m.vandeven-3@umcutrecht.nl