Optimized graphical testing procedures

Dr. rer. nat. Dennis Görlich (dennis.goerlich@uni-muenster.de)
Dr. rer. nat. Robert Kwiecien,
Prof. Dr. rer. nat. et med. habil. Andreas Faldum

Westfälische Wilhelms-Universität Münster
Institut für Biometrie und Klinische Forschung
Schmeddingstrasse 56
48149 Münster
Overview

1. Motivation
2. Graphical multiple testing procedures
3. Evolutionary algorithm for optimization
4. Simulation study
5. Outlook
Motivation

• Clinical trials often are planned to answer *several equally important* hypotheses simultaneously.

• **Graphical multiple testing procedures** as propopsed by Bretz et al. provide a viable tool to set up statistical designs for clinical trials. (applicable for suitable set of hypotheses)

• The development of a *suitable graphical design* is a task that requires much experience and thought on the desired properties of the statistical tests.

What is a good graphical multiple testing procedure, given a planning alternative?
Graphical multiple test procedures

After Bretz et al. (2009):

Given a set $H$ of $m$ elementary hypotheses $H_i$, the specification of

- a local significance levels $\alpha = (\alpha_1, \ldots, \alpha_m)$ with $\sum_i \alpha_i = \alpha$
- a $m \times m$ transition matrix $G = (g_{ij})$
  with $0 \leq g_{ij} \leq 1$, $g_{ii} = 0$, $\sum_k g_{ik} \leq 1$, $\forall i, j = 1, \ldots, m$
- an update algorithm

defines a short cut for a consonant closed test procedure with weighted Bonferroni tests for the intersection hypotheses.

Bretz, Maurer, Brannath, Posch, 2009,
„A graphical approach to sequentially rejective multiple test procedures“, Statistics in Medicine, 28:586-604

Workshop „Adaptive Designs and Multiple Testing procedures“
26.06.-28.06.2015
Graphical multiple test procedures

Fixed sequence procedure

Bonferroni-Holm procedure

Procedure from Bretz et al. 2011
Graphical multiple test procedures

$\alpha = 0.05$

$p_1 = 0.0001$

$p_2 = 0.823$

$p_3 = 0.046$

$p_4 = 0.012$
Can we optimize these graphical designs?
Optimization problem

A multiple test problem among the elementary hypotheses \((H_i)_{i \in I}\) a sequential rejective testing procedure \(T\) is characterized by the transition weight matrix \(G = (g_{ij})_{i,j \in I}\) and node weight vector \(w = (w_i)_{i \in I}\).

Given

- a planning alternative \(K\)
- and a fitness function \(f(G,w,K)\)

the optimal multiple testing problem \((G^*,w^*)\) has to be identified by maximizing \(f(G,w,K)\).

Here power is used as fitness measure

- at-least-one rejection
- reject all hypotheses
- reject an arbitrary subset of hypotheses
Evolutionary algorithm

- **Get simulated local p-values**
- **1. Fitness evaluation**
  - Determine power by performing gMCP
- **2. Selection**
  - $T_{best}^{gen} = \arg\max_{T_k} (power(T_k))$
  - $T_{worst}^{gen} = \arg\min_{T_k} (power(T_k))$
- **Pop**
  - $T_1 = (G_1, w_1)$
  - $T_2 = (G_2, w_2)$
  - $T_3 = (G_3, w_3)$
- **Power()**
  - at-least-one rejection or all rejected or any combination of $H_i$
- **gen = gen + 1**
- **4. Replace worst**
  - $T_{worst}^{gen} = T_{best}^{gen}$
- **3. Mutation**
  - $T_{best}^{gen} = mutate(T_{best}^{gen})$
  - Mutates the transition weight and node weights.

Implementation: R, package gMCP
A simulation study

Clinical trial with
• 1 group of \( N \) patients
• 6 continuous test statistics \( \sim N(d_i, 1), i = 1, \ldots, 6 \)
• One-sample t-Tests with \( H_i: d_i = 0, i = 1, \ldots, 6 \)
• Planning alternatives:

\[
\begin{align*}
K_1: & \quad d_1 = 0 \\
K_2: & \quad d_2 = 0.1 \\
K_3: & \quad d_3 = 0.1 \\
K_4: & \quad d_4 = 0.2 \\
K_5: & \quad d_5 = 0.3 \\
K_6: & \quad d_6 = 0.4 \\
\end{align*}
\]

Initialisation = Bonferroni-Holm

\[
\begin{align*}
w_1 & = 0.1667 \\
w_2 & = 0.1667 \\
w_3 & = 0.1667 \\
w_4 & = 0.1667 \\
w_5 & = 0.1667 \\
w_6 & = 0.1667 \\
\end{align*}
\]
Evolutionary algorithm

1. Fitness evaluation
   Determine power by performing gMCP

2. Selection
   \[ T_{\text{best}} = \arg\max_{T_k} (\text{power}(T_k)) \]
   \[ T_{\text{worst}} = \arg\min_{T_k} (\text{power}(T_k)) \]

3. Mutation
   \[ T'_{\text{best}} = \text{mutate}(T_{\text{best}}) \]
   \text{mutate(\(T_{\text{best}}\))}
   Mutates the transition weight and node weights.

4. Replace worst
   \[ T_{\text{worst}}^{\text{gen}} = T'_{\text{best}}^{\text{gen}} \]

\[ \text{Pop} \]
\[ T_1 = (G_1, w_1) \]
\[ T_2 = (G_2, w_2) \]
\[ T_3 = (G_3, w_3) \]

\[ \text{power(\(T\))} \]
\text{at-least-one rejection or all rejected or any combination of } H_i

\[ \text{gen} = \text{gen} + 1 \]

Implementation: R, package gMCP
Exemplary optimization run
Exemplary optimization run
Expected number of rejections

Graph: Expected number of rejections for best individual (50 optimization runs, 10000 simulations per generation)

Init: Bonferroni-Holm

Patients
Power (at least one rejection)
Time to best individual

Time to best individual
(50 optimization runs, 10000 simulations per generation)

Patients

Generations

5 10 15 20 30 40 50

50 100 150 200
Does time have an effect on fitness?

Slope = \(-1.5E-6\)
p-value = 0.798

Slope = \(9.8E-6\)
p-value = 0.31

Slope = \(5.9E-6\)
p-value = 0.558

Slope = \(2.3E-5\)
p-value = 0.056

Slope = \(6.2E-6\)
p-value = 0.518

Slope = \(5.9E-6\)
p-value = 0.654

Slope = \(1.1E-5\)
p-value = 0.112

Power (at least one rej.)

Generations
Summary

• Graphical multiple test procedures can be optimized with respect to power.

• Given a planning alternative the main factor for optimization seems to be the node weights \( w \).

• Length of opt. runs: Later found optimal solutions does not seem to be much better with respect to power.
Outlook

• **EA parameters / characteristics** (population size, adaptive mutation strength).

• **Constraints** on the design need to be maintained during optimization.

• Optimization for **multiple fitness values** simultaneously, e.g., maximizing the power to reject two primary hypotheses and the expected rejections.

• Optimization and simultaneous **sample size determination** given a desired power.
Optimized graphical testing procedures

Dr. rer. nat. Dennis Görlich (dennis.goerlich@uni-muenster.de)
Dr. rer. nat. Robert Kwiecien,
Prof. Dr. rer. nat. et med. habil. Andreas Faldum

Westfälische Wilhelms-Universität Münster
Institut für Biometrie und Klinische Forschung
Schmeddingstrasse 56
48149 Münster

Workshop „Adaptive Designs and Multiple Testing procedures“
26.06.-28.06.2015