Adaptive clinical trials with subgroup selection

Tim Friede\(^1\), Nick Parsons\(^2\), Nigel Stallard\(^2\)
\(^1\)University Medical Center Göttingen, Göttingen, Germany
\(^2\)Warwick Medical School, University of Warwick, Coventry, UK

n.stallard@warwick.ac.uk

European Medicines Agency Workshop on subgroup analysis
18 November 2011
Setting

Definitive comparison of experimental treatment with control

Single predefined subgroup of interest

Research questions: is treatment effective in whole population
is treatment effective in subgroup

For confirmatory study wish to control risk of any false positive
Control familywise error rate in strong sense at specified level

Adaptive design – idea

We are testing hypotheses in full population and subgroup

Controlling false positives means it is harder to show efficacy

If we knew which hypothesis to test, could focus on that

Adaptive design idea:
use interim data to guide hypothesis to test in remainder of trial
Adaptive design – details

Stage 1: Recruit from full population
  On basis of interim data decide to continue to
  (i) test full and sub-populations at final analysis
  (ii) test full population only final analysis
  (iii) test sub-population only final analysis
      (possibly include interim data on short-term endpoint)
Stage 2: Cases (i) or (ii): continue to recruit from full population
Case (iii): recruit from subgroup only

Final analysis: Conduct selected tests
    Control error rate allowing for test selection

(Song & Chi 2007, Brannath et al 2009, Jenkins et al 2011)
Test statistics

Test $H_0^{\{F\}}$ (full population) and $H_0^{\{S\}}$ (sub population)

At stage $i$, let
$p_{i}^{\{F\}}$ and $p_{i}^{\{S\}}$ be p-values for full and sub-population
$Z_{i}^{\{F\}}$ and $Z_{i}^{\{S\}}$ be test statistics for full and sub-population

$$ p_{i}^{\{F\}} = 1 - \Phi(Z_{i}^{\{F\}}), \quad p_{i}^{\{S\}} = 1 - \Phi(Z_{i}^{\{S\}}) $$

Assume $Z_{i}^{\{F\}}$ and $Z_{i}^{\{S\}}$ normal with
unit variance

correlation $\sqrt{\tau}$ where $\tau$ is proportion in subgroup

(Spiessens and Debois 2010)
Controlling the overall error rate

**Bonferroni test**

\[ p_i^{\{F, S\}} = 2 \min\{p_i^{\{F\}}, p_i^{\{S\}}\} \]

**Simes test**

\[ p_i^{\{F, S\}} = \min\{2 \min\{p_i^{\{F\}}, p_i^{\{S\}}\}, \max\{p_i^{\{F\}}, p_i^{\{S\}}\}\} \]

(Simes 1986)

These tests ignore correlation between \( p^{\{F\}} \) and \( p^{\{S\}} \)

**Spiessens and Debois test**

Obtain p-value based on distribution of \( \max\{Z_i^{\{S\}}, Z_i^{\{F\}}\} \)

(Spiessens and Debois 2010)
Combining stages 1 and 2

Combination test
Obtain overall p-value for each hypothesis using pre-specified combination function (Bretz et al 2006)

Conditional error function approach
Obtain design assuming testing both hypotheses
At interim analysis obtain error rate for second stage conditional on stage 1 data
Can redesign stage 2 so long as conditional error rate is controlled
  e.g. restrict testing to full population or subgroup alone (Müller and Schäfer 2001, Friede et al 2011)
Simulation study

Based on Brannath et al. (2009) and Jenkins et al. (2011)

Model: Normal approximation for logrank test statistics
Standardised effects based on
haz. ratio: 0.77 in subgroup, 1 outside subgroup

\[ n_1 = 170, \quad n_2 = 748 \] (if recruit from full population)
\[ 470 \] (if recruit from subgroup)

Subgroup prevalence: 20% to 50%

Fixed designs: single study testing both hypotheses
two separate studies

Adaptive designs: CEF test with Spiessens and Debois test combination test with Spiessens test combination test with Simes test
Simulation study results

Adaptive: CEF
Adaptive: Comb test Simes
Adaptive: Comb test Spiessens
Fixed: separate studies
Fixed: single study
Conclusions

We may want to test treatment in a subgroup in addition to full population.
In confirmatory test need to control overall error rate.

Adaptive design allows us to use interim analysis data to select hypotheses for final analysis.

Appropriate methodology exists to control overall error rate.
Adaptive design appears more powerful than either:
fixed single stage design adjusting for multiple testing
selection of hypotheses in separate study.

We have assumed a pre-defined subgroup. Allowing for use of interim data to select subgroup is harder!
References


