Multiple testing methodology in the context of subgroup analysis

Alex Dmitrienko (Quintiles)

Brian Millen (Eli Lilly and Company)

EMA Expert Workshop on Subgroup Analysis

Outline

Tailored therapeutics setting

Clinical trials with pre-specified subpopulations

Key statistical considerations

Design considerations: Multiplicity adjustment

Analysis considerations: Influence and interaction conditions

Tailored therapeutics setting

Clinical trials with multiple patient populations

Overall population

One or more subpopulations based on pre-specified genotypic, clinical or other markers

Confirmatory subgroup analysis

Overall population and subpopulations are equally important

Efficacy in at least one population provides foundation for registration

Pre-specified subpopulations

Genotypic markers

Breast cancer patients with amplified HER2 gene

Clinical markers

Patients with nonsquamous non-small cell lung cancer

Socio-demographic markers

ADHD patients who live in a stable environment

Two-population setting

Populations

Population O: Overall population

Population S_+: Marker-positive population

Population S_-: Marker-negative population

Hypothesis testing problem

H_0 and H_+, null hypotheses of no effect in Populations O and S_+

Successful outcome if at least one null hypothesis is rejected

Multiplicity adjustment

Error rate control

Control familywise error rate for $\{H_0, H_+\}$ at one-sided $\alpha = 0.025$ to enable regulatory claims in both populations

Account for logical relationships

H_0 and H_+ are interchangeable

Account for distributional information

Test statistics for H_0 and H_+ are positively correlated

Multiplicity adjustment procedures

Fixed-sequence procedure

Chain procedures

Bonferroni-based chain procedures (Bretz et al., 2009)

Parametric chain procedures (Millen and Dmitrienko, 2011)

Feedback procedures

Feedback procedures (Zhao, Dmitrienko and Tamura, 2010)

Fixed-sequence procedure

Decision rules

$\alpha = 0.025$, Familywise error rate

1. Test H_0 at 0.025
2. Test H_+ at 0.025 only if H_0 is rejected

Logical relationships are not taken into account

Bonferroni-based chain procedures

α allocation rule

αw_0 and αw_+ are assigned to H_0 and H_+

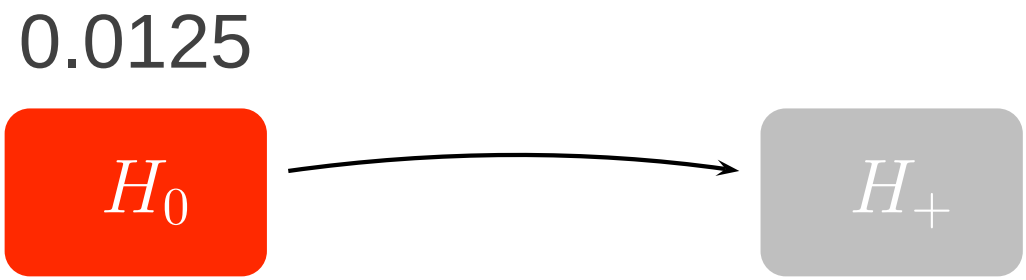
w_0 and w_+, non-negative weights with $w_0 + w_+ = 1$

α propagation rule

If H_0 is rejected, its significance level is transferred to H_+ and vice versa

Bonferroni-based chain procedure

Step 1



$w_0 = w_+ = 0.5$, Equally weighted analyses

Test H_0 at 0.0125

Carry 0.0125 forward if H_0 is rejected

Bonferroni-based chain procedure

Step 2

Test H_+ at 0.0125 if H_0 is not rejected and at 0.025 if H_0 is rejected

Carry 0.0125 backward if H_+ is rejected

Bonferroni-based chain procedure

Step 3

Retest H_0 at 0.025 if H_+ is rejected

Logical relationships are taken into account but distributional information is ignored

Distributional information

Correlation

Test statistics for H_0 and H_+ are generally strongly positively correlated

Correlation depends on the relative size of the marker-positive population

Example

Correlation=0.7 if 50% of patients are marker-positive ($n_+ = n_0/2$)

Distributional information

Parametric chain procedures

Powerful procedures that extend Bonferroni-based chain procedures

Feedback procedures

Powerful “adaptive” procedures that extend parametric fallback procedures (Huque and Alosh, 2008; Alosh and Huque, 2009)

Parametric chain procedures

α allocation rule

αw_0 and αw_+ are assigned to H_0 and H_+

w_0 and w_+, non-negative weights with $w_0 + w_+ = 1$

α propagation rule

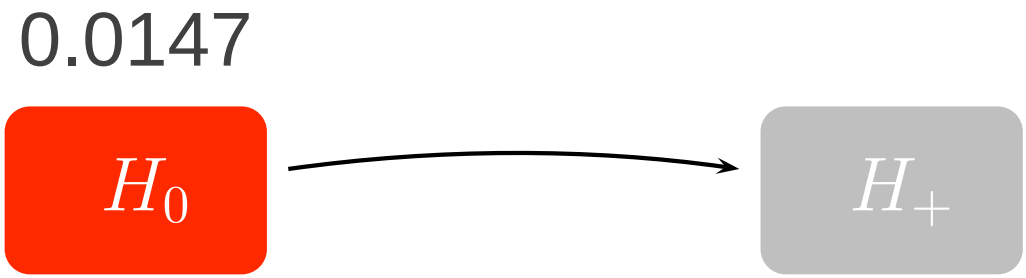
If H_0 is rejected, its significance level is transferred to H_+ and vice versa

Distributional information

Hypothesis test statistics follows a bivariate normal distribution

Parametric chain procedure

Step 1



$w_0 = w_+ = 0.5$, Equally weighted analyses

50% of patients are marker-positive ($\rho = 0.7$)

Test H_0 at 0.0147

Carry 0.0103 forward if H_0 is rejected

Parametric chain procedure

Step 2

Test H_+ at 0.0147 if H_0 is not rejected and at 0.025 if H_0 is rejected

Carry 0.0103 backward if H_+ is rejected

Parametric chain procedure

Step 3

Retest H_0 at 0.025 if H_+ is rejected

Logical relationships and distributional information are taken into account

Analysis considerations

Outcomes

Case 1: Significant effect in the overall population

Case 2: Significant effect in the marker-positive population

Case 3: Significant effects in both populations

Case 3

Regulatory claims for both populations?

Influence and interaction conditions (Millen et al., 2011) play a key role in decision making process

Influence condition

Case 3

Significant treatment effects in both populations

Influence condition

Beneficial effect in the overall population is not restricted to the marker-positive population

Influence condition is not met

Overall population

Effect size 0.1

Marker-positive population [50%]

Effect size 0.3

Marker-negative population [50%]

Effect size -0.1

Influence condition

Labeling implications

If the influence condition is not met, beneficial effect in the overall population is driven by highly beneficial effect in the marker-positive population

Regulatory claim may be restricted to the marker-positive population

Interaction condition

Case 3

Significant treatment effects in both populations

Interaction condition

Beneficial effect in the marker-positive population is appreciably greater than beneficial effect in the overall population

Interaction condition is not met

Overall population

Effect size 0.3

Marker-positive population [50%]

Effect size 0.3

Marker-negative population [50%]

Effect size 0.3

Interaction condition

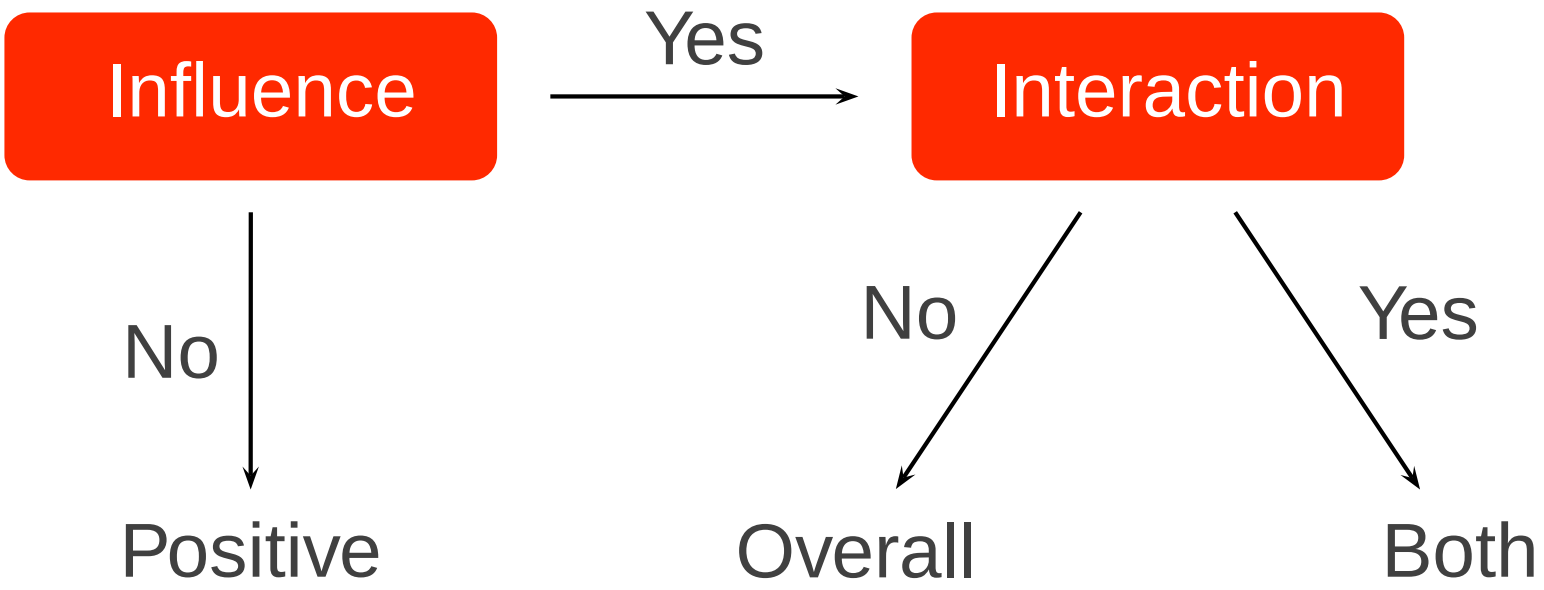
Labeling implications

If the interaction condition is not met, the marker is not informative

Regulatory claim may be limited to the overall population

Decision making process in Case 3

Regulatory claims in the overall and marker-positive populations



References

Alosh, M., Huque, M. (2009). A flexible strategy for testing subgroups and overall population. *Statistics in Medicine*. 28, 3-23.

Bretz, F., Maurer, W., Brannath, W., Posch, M. (2009). A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine*. 28, 586-604.

Dmitrienko, A., Millen, B.A., Brechenmacher, T., Paux, G. (2011). Development of gatekeeping strategies in confirmatory clinical trials. *Biometrical Journal*. To appear.

Huque, M., Alosh, M. (2008). A flexible fixed-sequence testing method for hierarchically ordered correlated multiple endpoints in clinical trials. *Journal of Statistical Planning and Inference*. 138, 321-335.

References

Millen, B.A., Dmitrienko, A. (2011). Chain procedures: A class of flexible closed testing procedures with clinical trial applications. *Statistics in Biopharmaceutical Research*. 3, 14-30.

Millen, B.A., Dmitrienko, A., Ruberg, S., Shen, L. (2011). Statistical considerations for clinical trials with tailoring objectives. In press.

Zhao, Y.D., Dmitrienko, A., Tamura, R. (2010). Design and analysis considerations in clinical trials with a sensitive subpopulation. *Statistics in Biopharmaceutical Research*. *2*, 72-83.