Correlating Survival Times with Gene Expression in Interim Analyses

Andreas Leha, Tim Beißbarth, Klaus Jung
Department of Medical Statistics, University Medical Center Göttingen, 37099 Göttingen, Germany
Contact: andreas.leha@med.uni-goettingen.de

INTRODUCTION

Survival studies combined with DNA microarray experiments are performed to detect correlations between expression levels of genes and patient survival. Survival studies are typically long-lasting and interim analyses that obtain early results are essential to provide the patients with the best possible treatment. When testing a single hypothesis interim analyses for early results are essential to provide the patients with the best possible treatment. Adjusted significance levels. In the case of high-dimensional microarray data, where many hypotheses are tested simultaneously, adjustment for interim analyses has been shown to be unnecessary under certain conditions [1].

We provide a simulation framework to explore error and power rates when performing interim analyses of high-dimensional microarray data [2]. A special focus of this framework are microarray experiments within survival studies. Additionally, the possibility of early stopping in interim analyses is studied.

STUDY DESIGN

• Arrival of patients within the phase [0, l1]: surgery + measurement of gene expression in tissue sample
• Length of bars: survival of patients
• Follow-up phase [l1, l2]: no new patients, update of survival data
• Patients marked with an ‘x’ have censored survival times at the final analysis
• Several interim analyses during arrival period and follow-up period

METHODS

Correlating Survival Times with Gene Expression

Detecting Survival Related Genes
Gene-wise Cox-regression:
\[ h(t) = h_0(t)e^{\beta x_i}, \quad i = 1, \ldots, \#Genes \]  

Multiple Hypothesis Testing
Control of the false discovery rate:
\[ FDR = E\left(\frac{\# \text{false positives}}{\max(\# \text{positives}, 1)}\right) \]  

Assessment of average power rate:
\[ APR = E\left(\frac{\# \text{true positives}}{\# \text{non-true null hypotheses}}\right) \]

Interim Analyses
Following [1] no individual adjustment of FDR level in interim analyses
Overall FDR level: 5%

RESULTS

Simulation Study
Parameters
• 50 patients
• Arrival times \( \sim \text{unif}([0, l_1]) \)
• Survival times \( \sim \text{exp}(\lambda) \)
• 5 years recruitment phase
• 5 years follow-up phase
• 10,000 genes
• Autoregressive correlation structure

Obtained error and power rates:

Example: Breast Cancer Data [3]
• 295 patients
• 11 years recruitment
• More than 7 years follow-up
• Simulated arrival times \( \sim \text{unif}([0, 11]) \)

FUNDING

Deutsche Forschungsgemeinschaft (KFO 179), German Federal Ministry of Education and Research (BMBF)
[NGFN network IG Prostate-Cancer (01GS0890), Medical Systems Biology network BreastSys]

REFERENCES


CONCLUSIONS

• Adjustment of FDR level in interim analyses not necessary
• Early stopping of survival studies possible
• In case of small effects: Alternative stop criterion (e.g. prediction accuracy of survival times)