

# **Subgroup Analysis of mCRPC Trials**

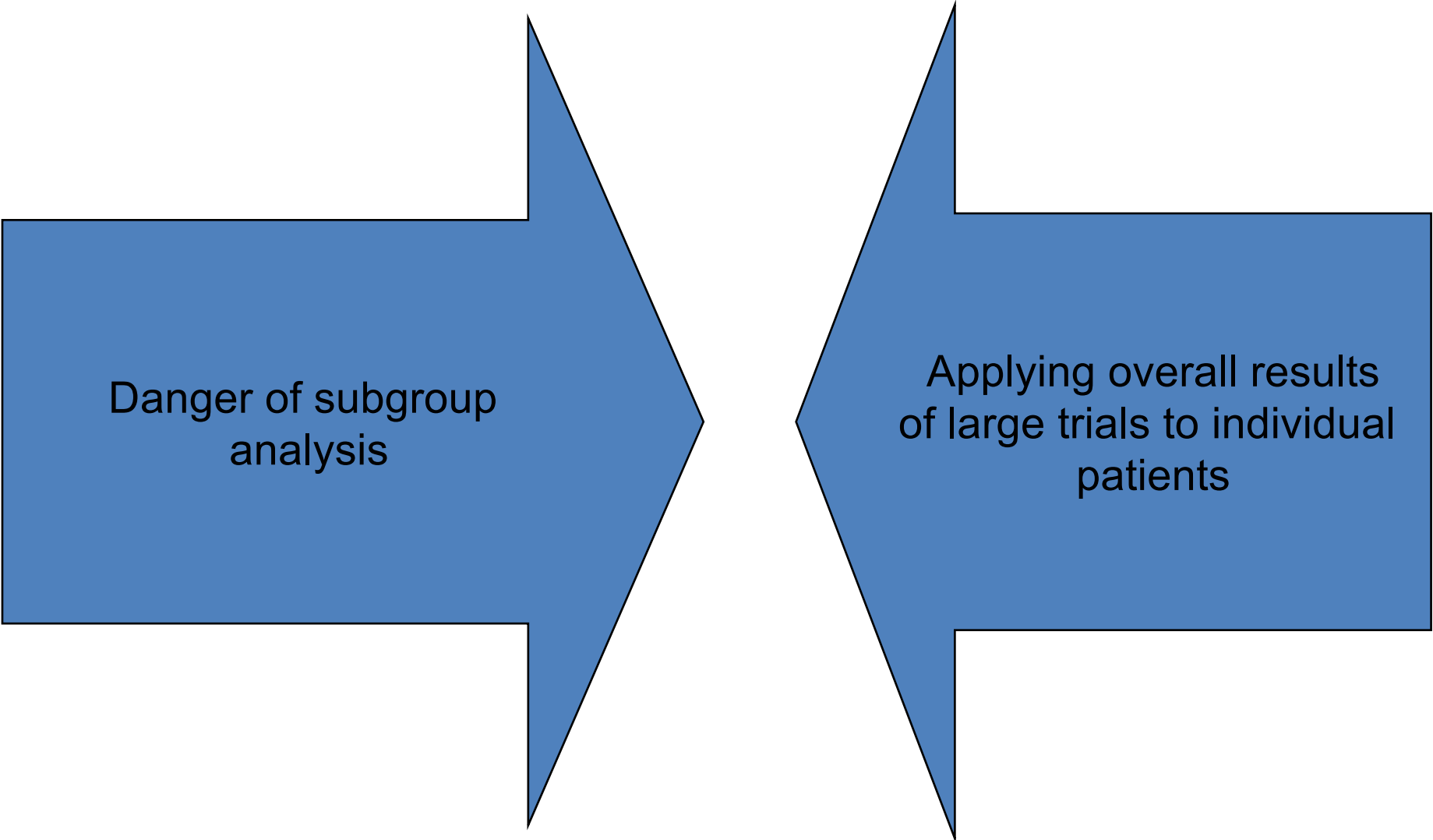
# Conflict of Interest

None

# General Assumption

- Hypothesis tested usually address an overall or 'average' treatment effect in the study population
- No assumption of homogeneity of effect across subgroups

# The Challenge



Danger of subgroup  
analysis

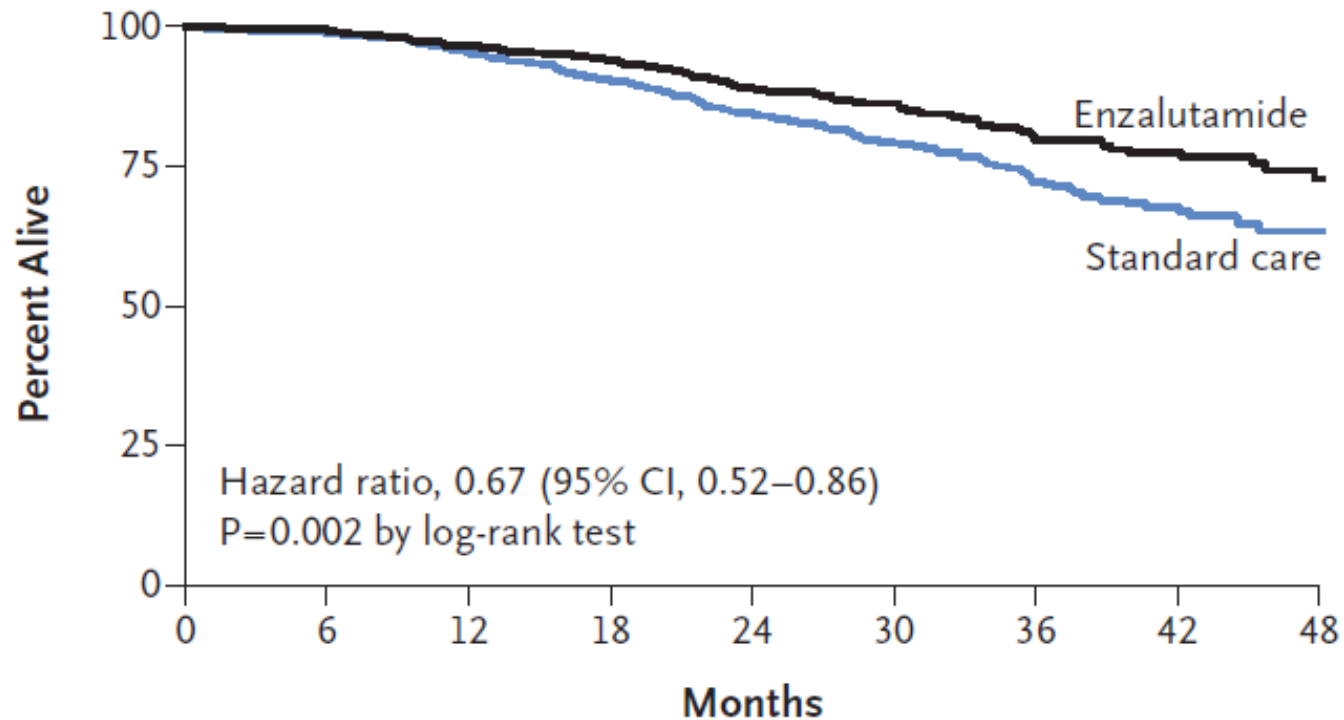
Applying overall results  
of large trials to individual  
patients

# Subgroup Analyses - Pervasive in Clinical Trials

- **Positive trial**
  - To characterize patients who benefit from the therapy vs. those who may not
- **Negative trial**
  - To identify at least some patients with treatment benefit

# Positive Trial: ENZAMET

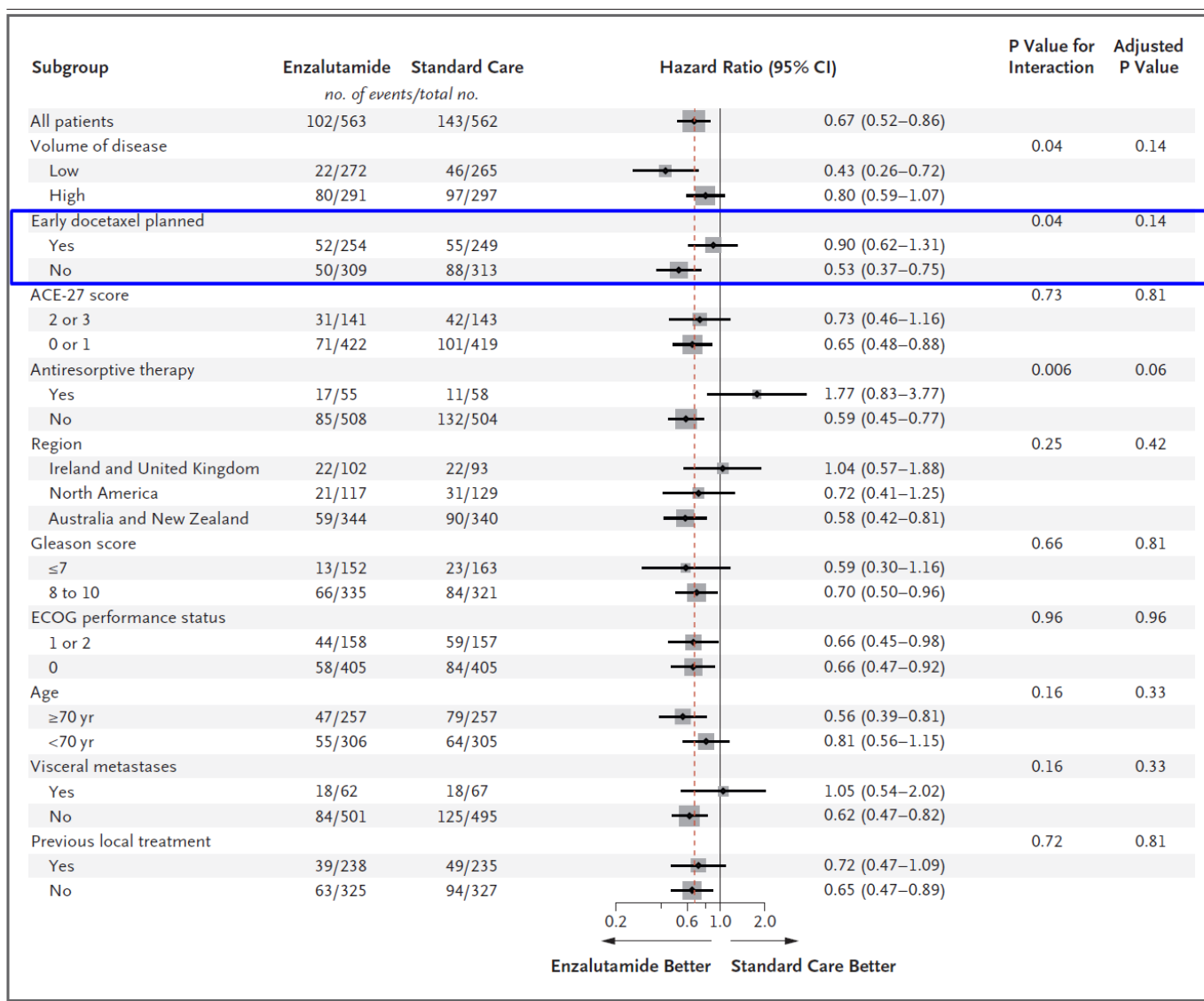
## A Overall Survival



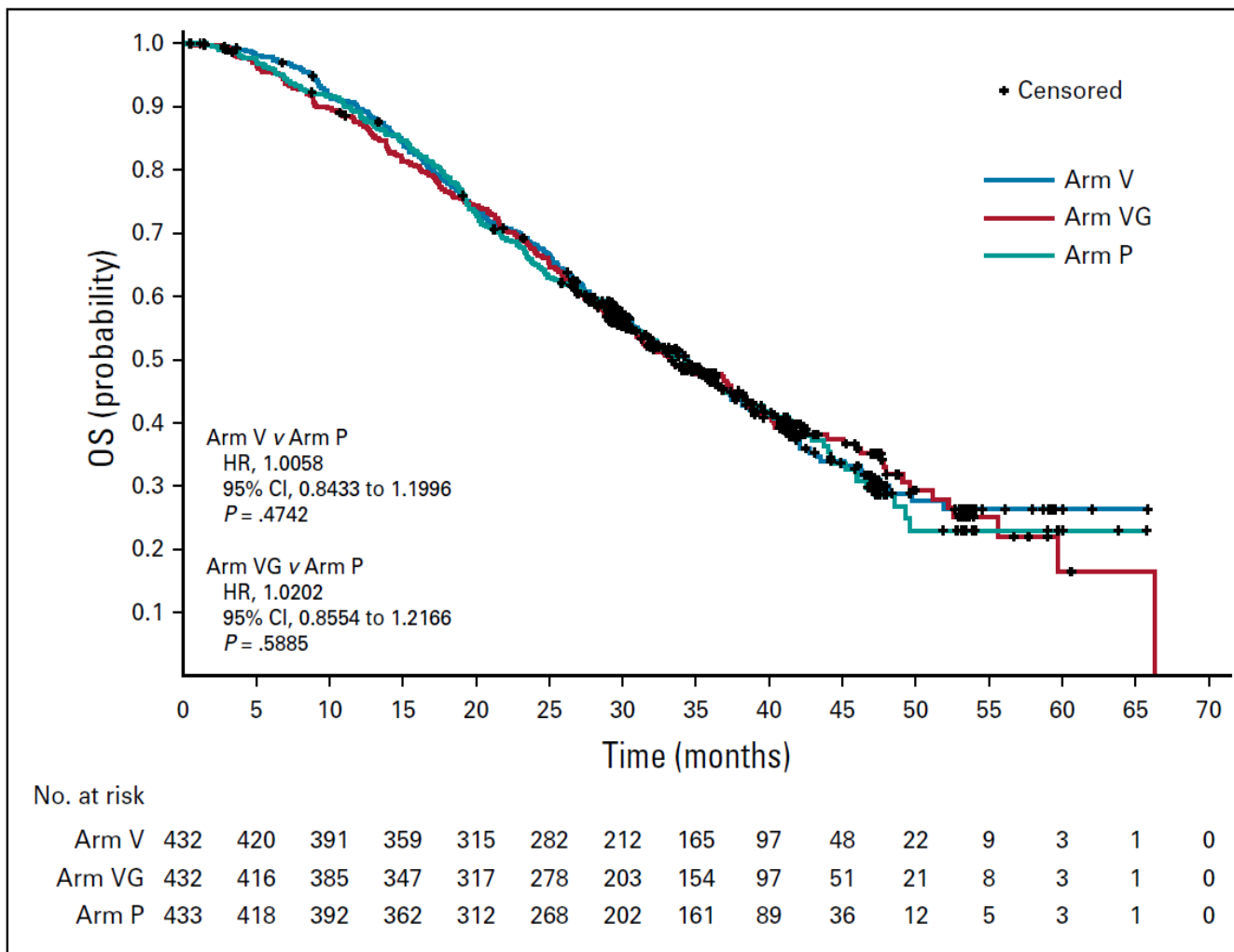
### No. at Risk

Enzalutamide	563	558	541	527	480	340	189	106	45
Standard care	562	551	531	501	452	311	174	86	32

# Positive Trial: ENZAMET

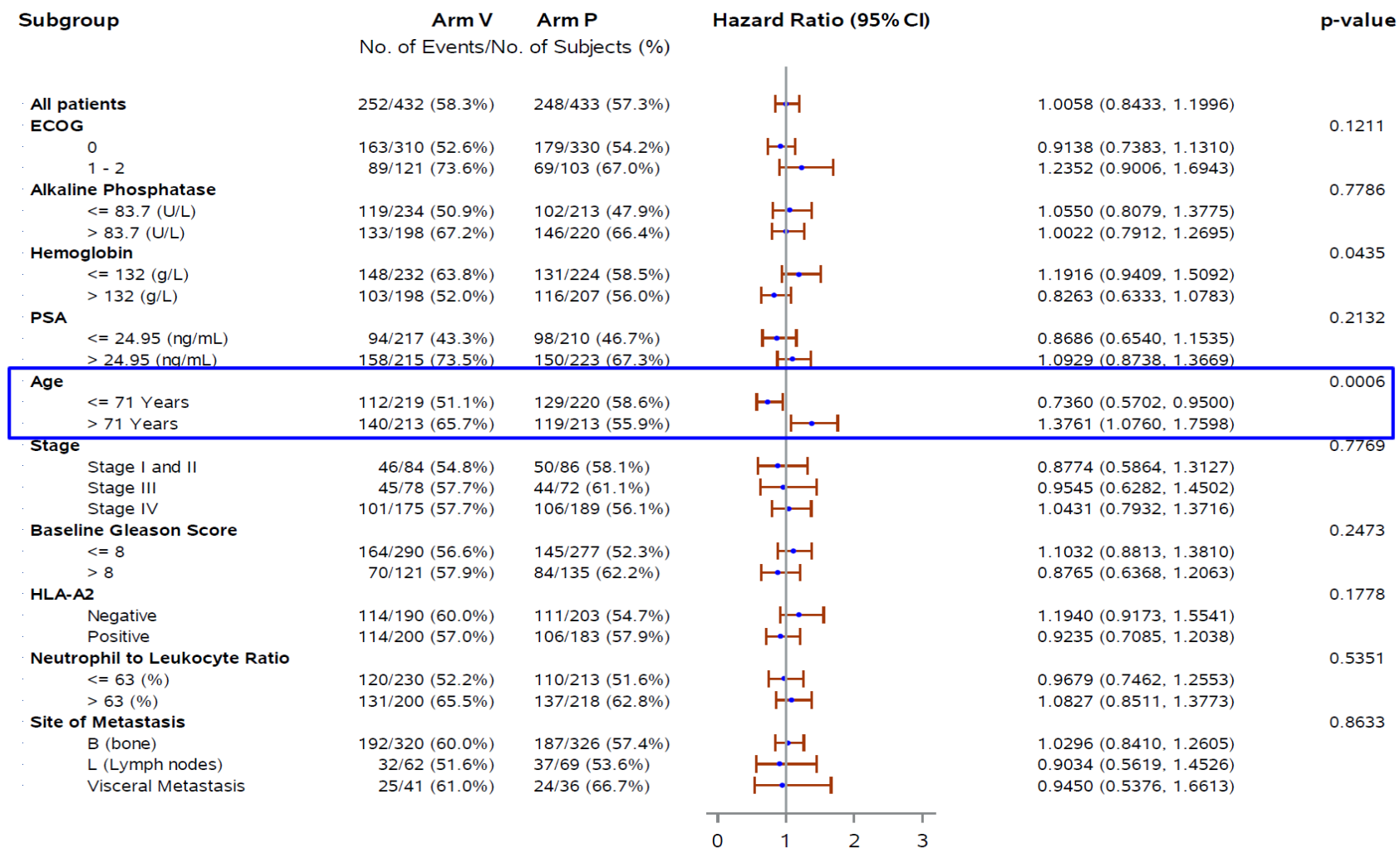


# Negative Trial: PROSTVAC





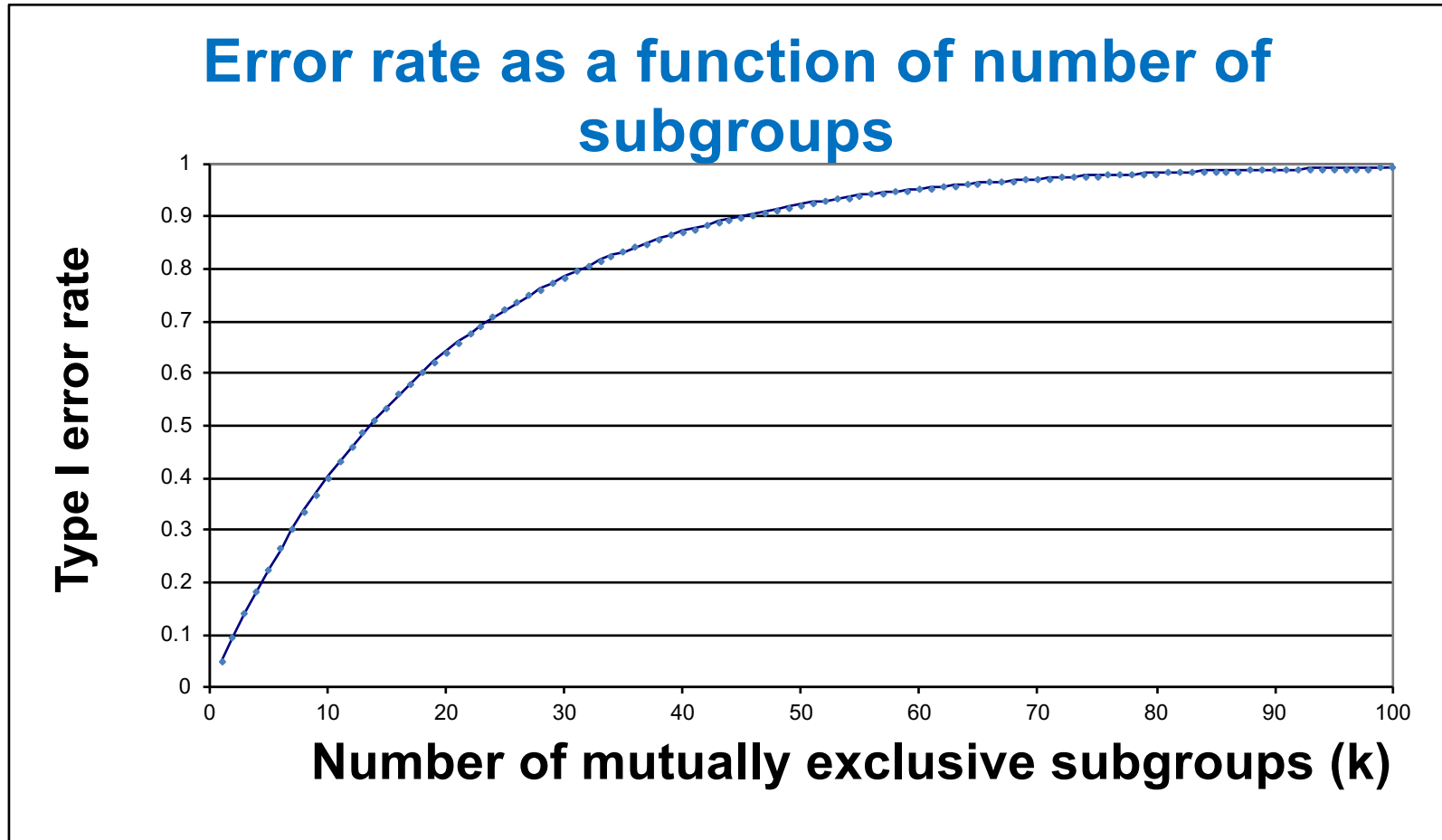
# Negative Trial: PROSTVAC



# Warning: Subgroup Analysis

- A machine for producing false negative and false positive results.

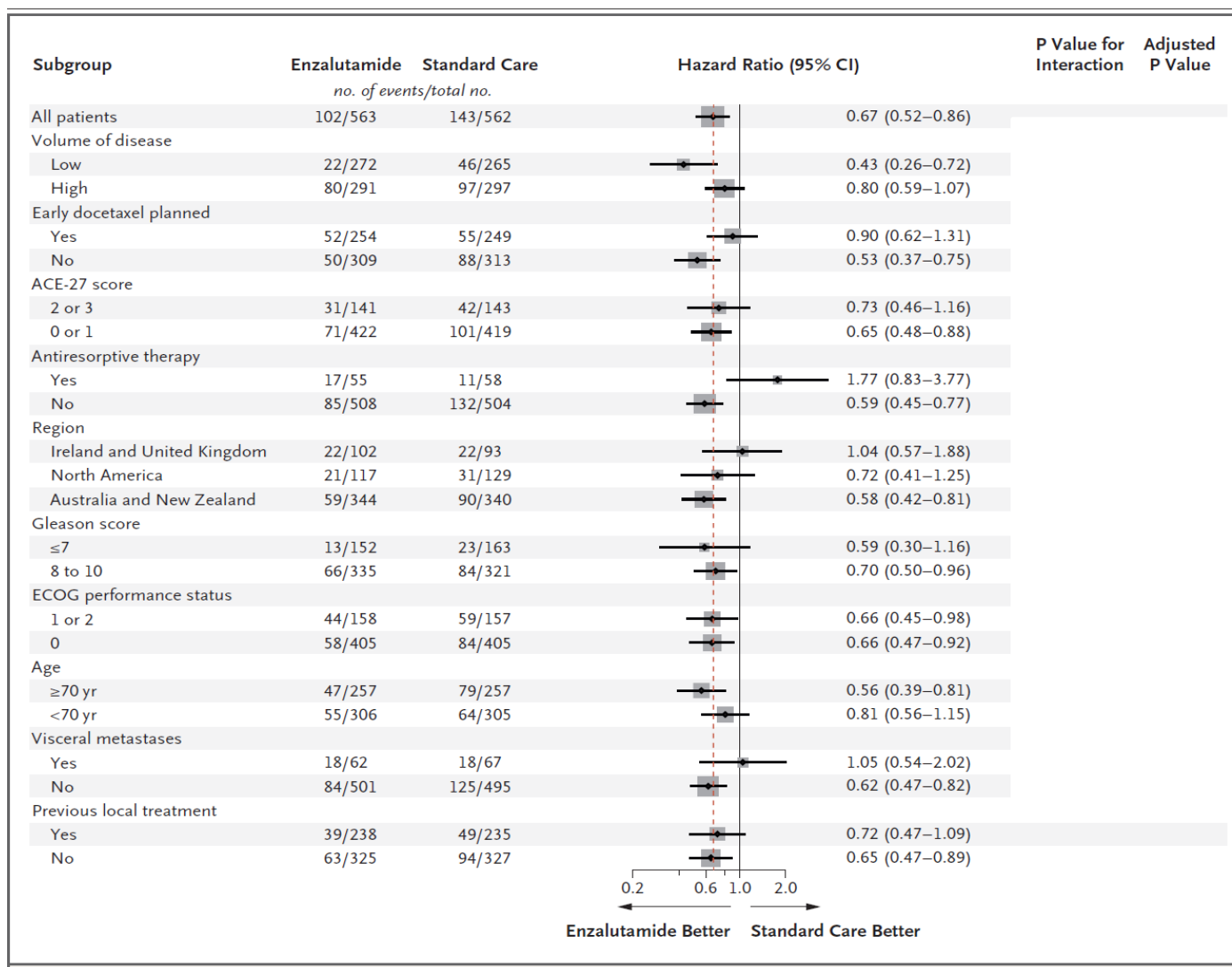
# 1. Type I Error Rate



*$k=5$ , probability is 0.23 that one comparison  $p$ -value  $< 0.05$*

*$k=10$ , probability is 0.40 at least one comparison  $p$ -value  $< 0.05$*

# Positive Trial: ENZAMET



## 2. Power Is An Issue

### Don't Be Misled

Ratio of Subgroup Events/ Total Events	Power (90%)	Power (85%)
1	0.90	0.85
0.75	0.83	0.74
0.50	<b>0.63</b>	<b>0.56</b>
0.40	<b>0.54</b>	<b>0.47</b>
0.30	<b>0.43</b>	<b>0.37</b>

Hazard ratio=0.75

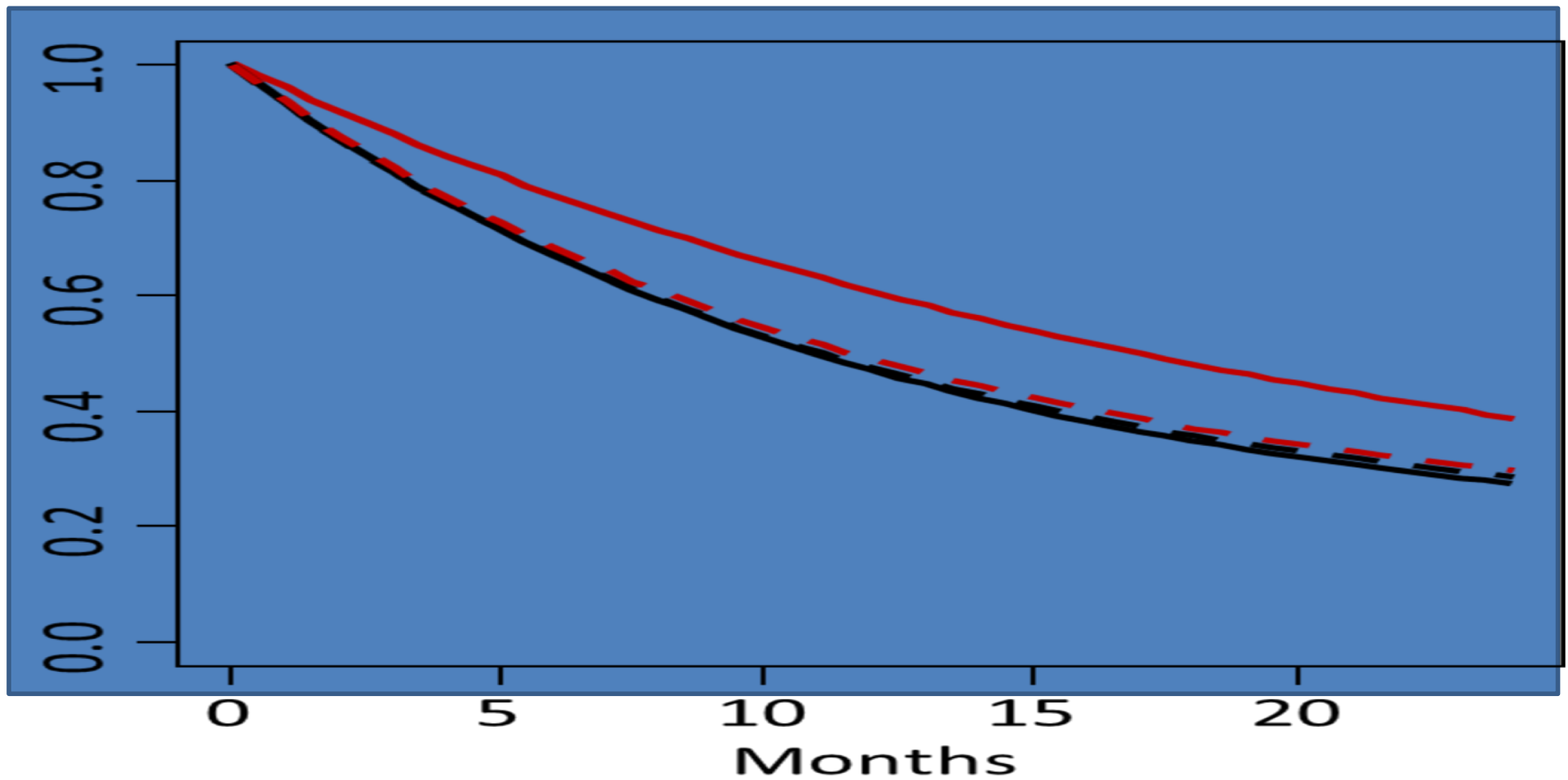
### 3. A Mistake to Avoid

- An incorrect inference that a subgroup effect is present based on separate tests of treatment effects within each level of the characteristic of interest, that is, to compare one significant and one non-significant  $p$ -value



# Subgroup Analyses

P-value for interaction



# Criteria to Assess Credibility of Subgroup Analyses

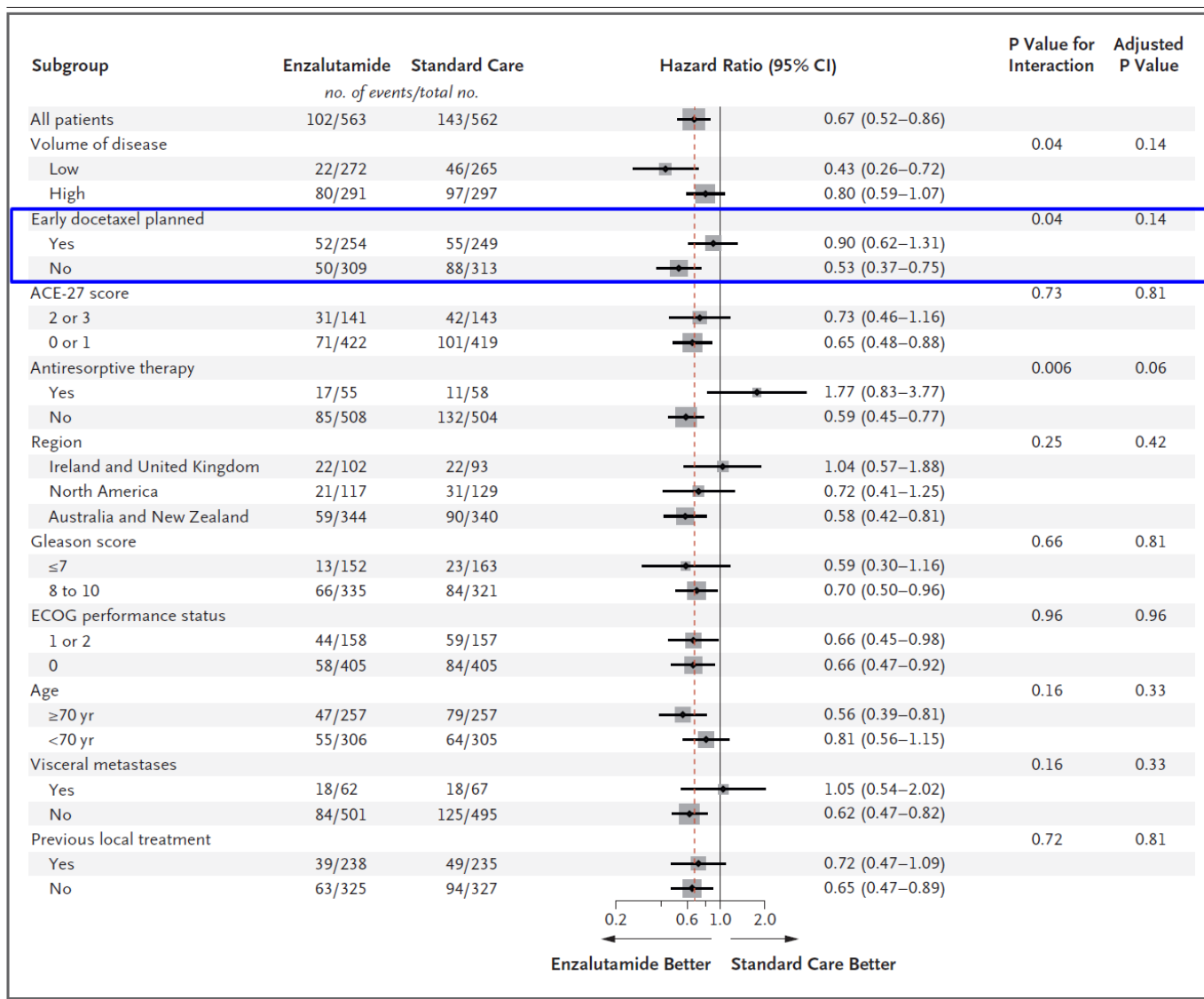
- Can chance explain the apparent subgroup effect?
- Is treatment effect consistent?
- Was the subgroup hypothesis one of a small number of hypotheses developed *a-priori* with direction specified?



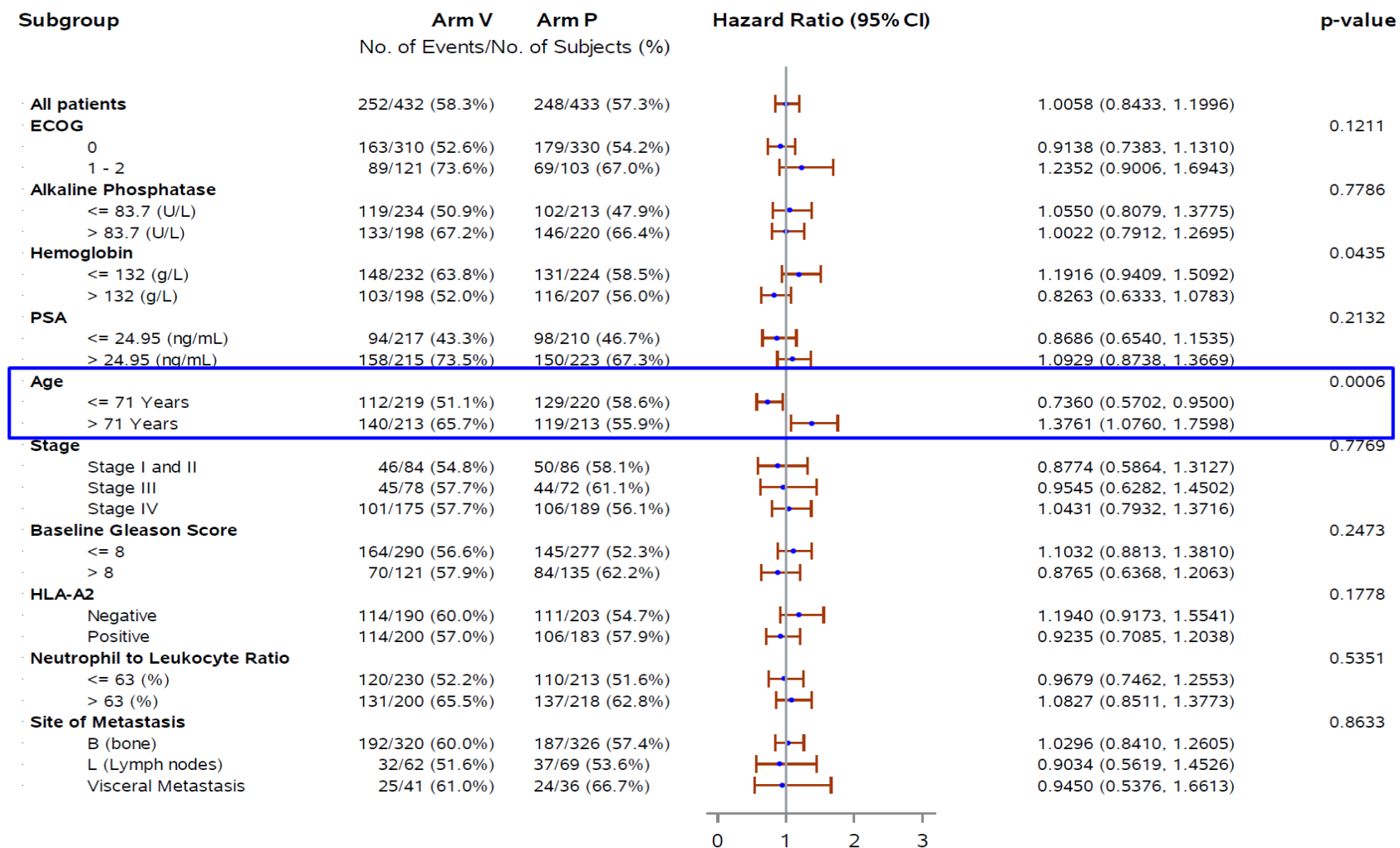
# Criteria to Assess Credibility of Subgroup Analyses

- Is there strong preexisting biological support?
- Is the evidence supporting the effect based on within- or between-study comparisons?

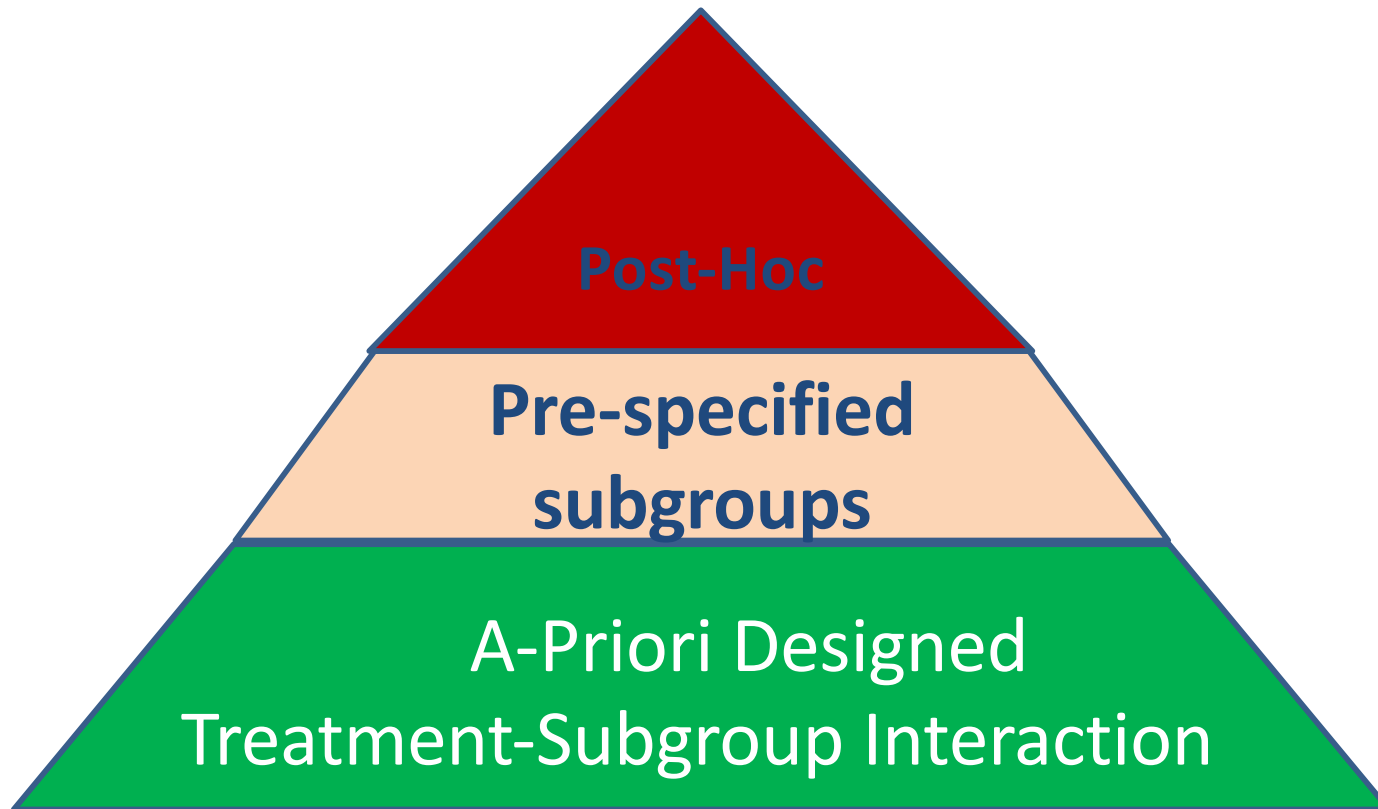
# Positive Trial: ENZAMET



# Negative Trial: PROSTVAC



# Level Of Evidence



# Safeguards: Design and Analysis Phase

- Clear description of hypothesis: direction
- Limit number of subgroup testing
- Statistical test of treatment-subgroup interaction
- Subgroup a stratification variable

# Safeguards: Interpretation

- Greater emphasis on the overall result than a subgroup
- test of treatment-subgroup interaction rather than treatment effect within subgroups
- Interpret the results in the context of other trials principles of biological rationale and coherence

# Conclusion

- Best statistical design
  - Answer primary question
  - Feasible
- Planning is key
  - Avoid “**statistical sins**”
- Pre-specified subgroup is better than post-hoc

# Conclusion

- Larger studies are needed for treatment-subgroup interaction
- Meta-analysis plays critical role



# A Final Note

“Rather than reporting isolated P values, articles should include effect sizes and uncertainty metrics.”

Waaserstein R, American Statistician 2016

The NEW ENGLAND JOURNAL of MEDICINE

## EDITORIALS



### **New Guidelines for Statistical Reporting in the *Journal***

David Harrington, Ph.D., Ralph B. D'Agostino, Sr., Ph.D., Constantine Gatsonis, Ph.D.,  
Joseph W. Hogan, Sc.D., David J. Hunter, M.B., B.S., M.P.H., Sc.D.,  
Sharon-Lise T. Normand, Ph.D., Jeffrey M. Drazen, M.D., and Mary Beth Hamel, M.D., M.P.H