

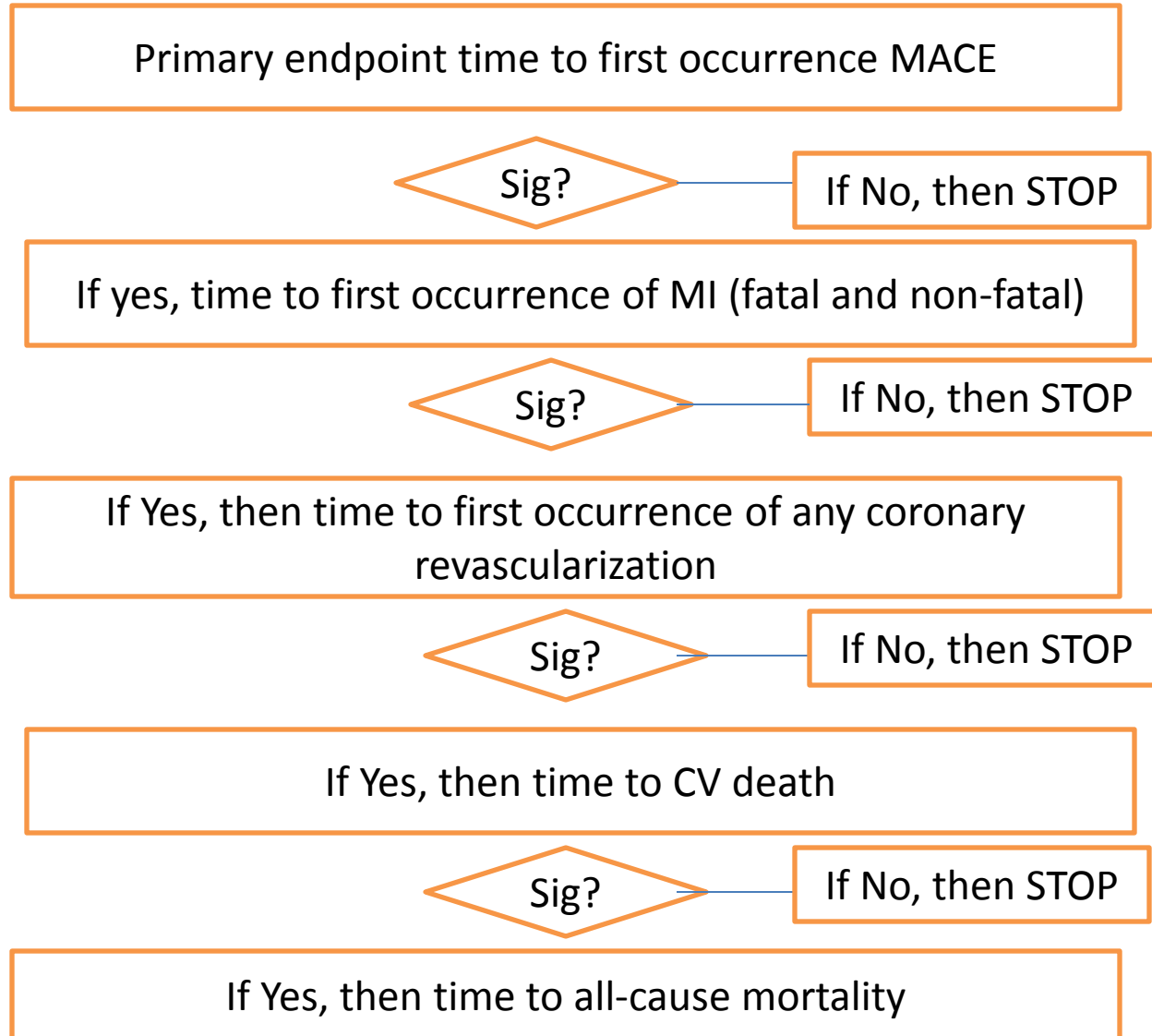
# Multiplicity Issues in Defining the Testing Strategy for Two Large Outcome Studies

By: Jennifer Shannon, Rebekkah Brown,  
Greg Cicconetti, and Rich Davies

# STABILITY and SOLID-TIMI 52

- Primary Endpoint is first occurrence MACE
  - CV death
  - Non-fatal MI
  - Non-fatal stroke
- Plan for 1500 adjudicated MACE events in each study
- 15,828 subjects randomized in STABILITY
- 13,027 subjects randomized in SOLID-TIMI 52

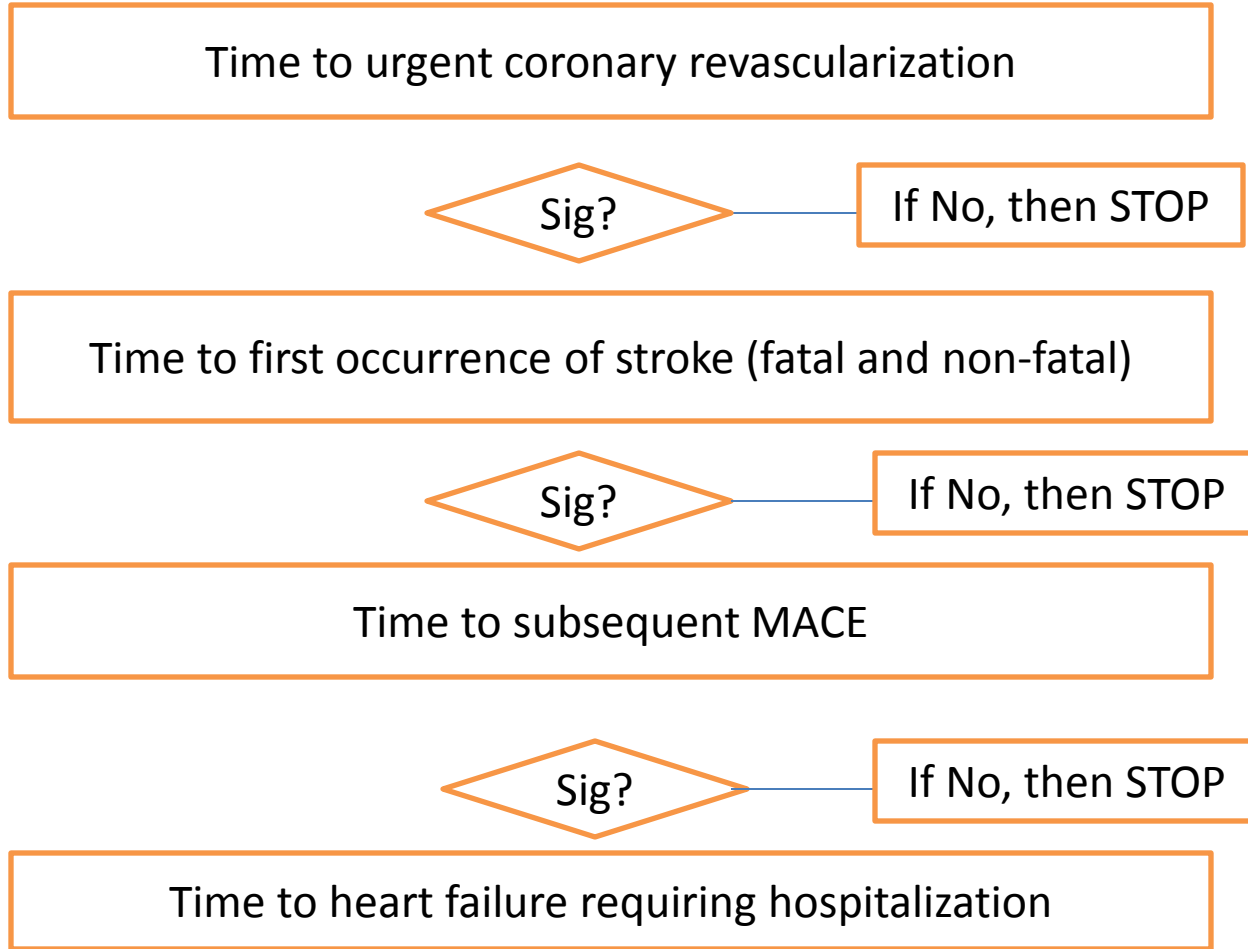
# Multiplicity Procedure – Individual Studies



# Gatekeeper Strategy

- P-value for primary endpoint MACE in both STABILITY and SOLID-TIMI 52 is  $<0.20$  with at least one study being statistically significant (per alpha spending function)
- Integrated p-value for MACE  $<0.01$ .
- Test of homogeneity for the integrated data based on treatment by study interaction is not qualitatively meaningful and statistically significant at 0.05.

# Multiplicity Procedure –Integrated Analysis



# Simulation Study for Gatekeeper Strategy

- Assumption: log of hazard ratios are approximately normal

Endpoint	MACE	MACE	UCR	UCR
Study	STABILITY	SOLID	STABILITY	SOLID
Underlying HR	0.845	0.845	1.00	1.00
Events	1500	1500	400	400

# Simulation Results

	Scenario (alpha levels one-sided)	Power
A	Integrated MACE <0.005	0.979
B	AND STABILITY MACE <0.025	0.896
C	AND SOLID-TIMI 52 MACE <0.10	0.879
D	AND INTEGRATED UCR <0.025	0.045
E	Integrated MACE <0.005 AND SOLID-TIMI 52 MACE <0.025	0.896
F	AND STABILITY MACE <0.10	0.879
G	AND Integrated UCR < 0.025	0.045
H	Scenario D OR Scenario G	0.048

# Question for Discussion

- Since there are no common endpoints in the individual studies and the integrated analysis, do the discussants agree that the type 1 error is adequately addressed in the integrated analysis through the gatekeeper strategy and the hierarchical testing approach?
  - Is this strategy overly conservative, especially with respect to the alpha levels tested for the secondary endpoints in the situation where the study is stopped for positive efficacy (and the stopping boundaries require very robust efficacy to stop)?