Statistics•Collaborative

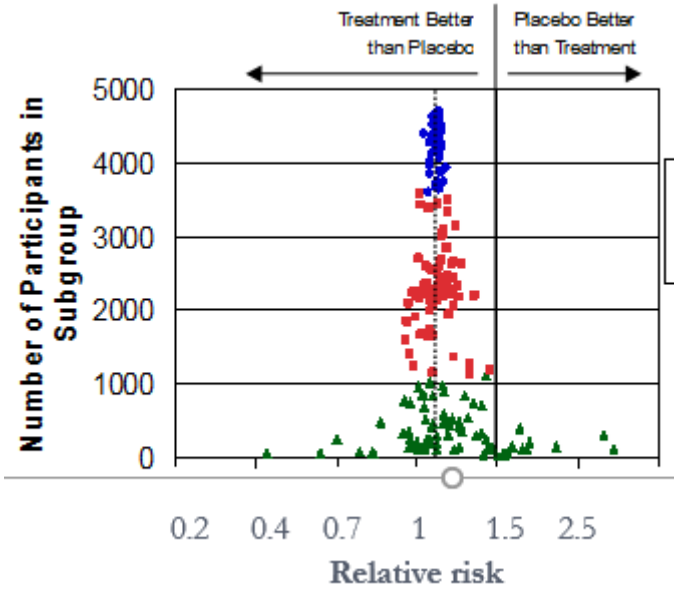
Evolving thoughts on subgroups

Janet Wittes, PhD

14th Annual University of Pennsylvania Conference
on Statistical Issues in Clinical Trials

April 12, 2022

SHEP – Relative Risk of CV Events



Interim Results of OvaRex Clinical Trial Demonstrate Clinical Benefit in Ovarian Cancer Patients

"The proportion of high-risk patients who achieved a disease-free survival of 6 months is significantly higher ($P = .0397$) among those treated with OvaRex (79%) than among those receiving placebo (39%)."

- ONCOLOGY Vol 15 No 8, Volume 15, Issue 8 (2001)

Interim Results of OvaRex Clinical Trial Demonstrate Clinical Benefit in Ovarian Cancer Patients

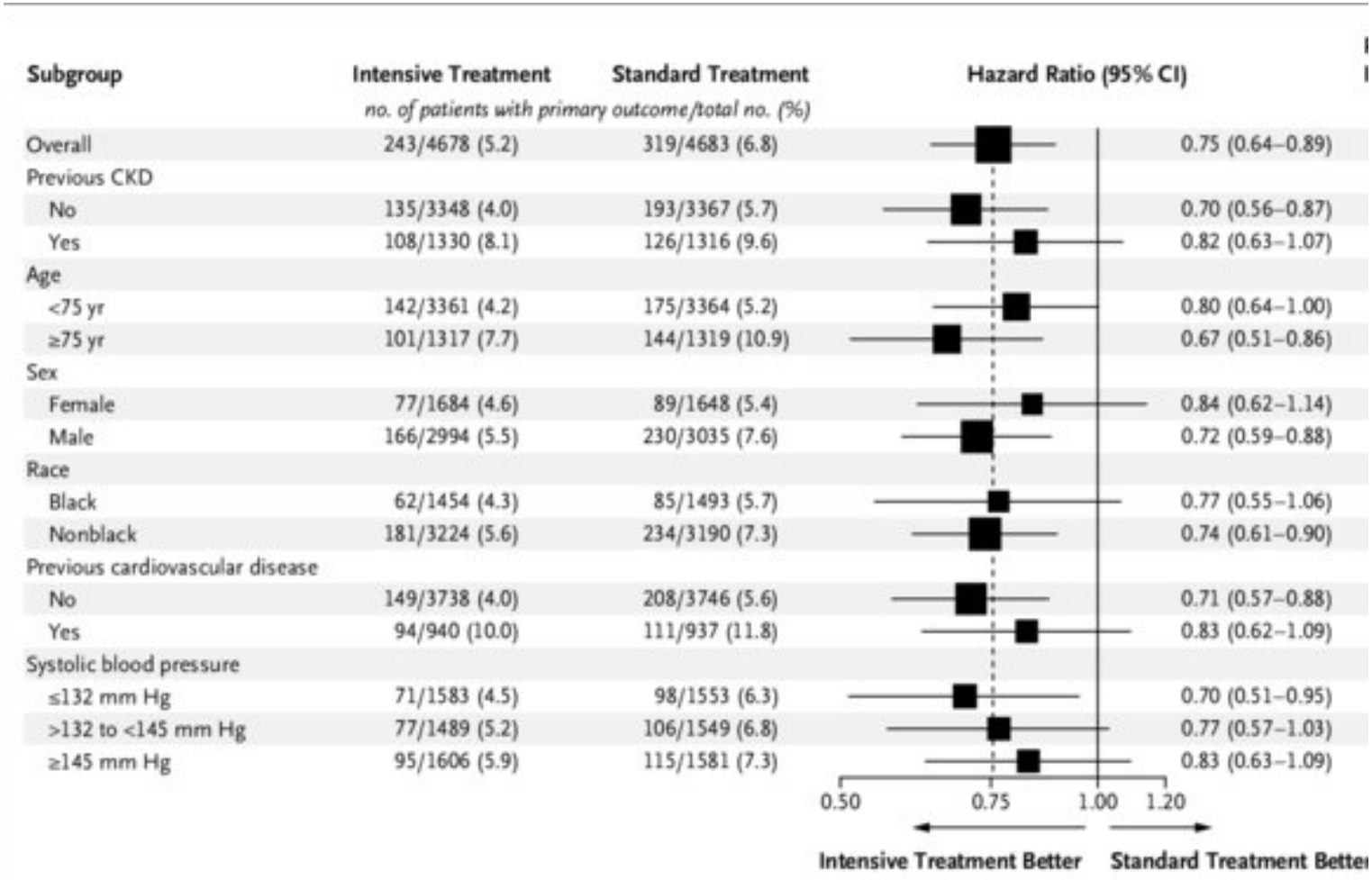
Unstated…this was ½ of the sample size; No effect overall – the lower risk half had significantly WORSE disease-free survival.

My view: before doing the next trial, investigators have to believe as strongly in the other ½ as you do in the half that shows success.

Phase 3 trials – no evidence of difference.

Pace Fleming

A typical forest plot



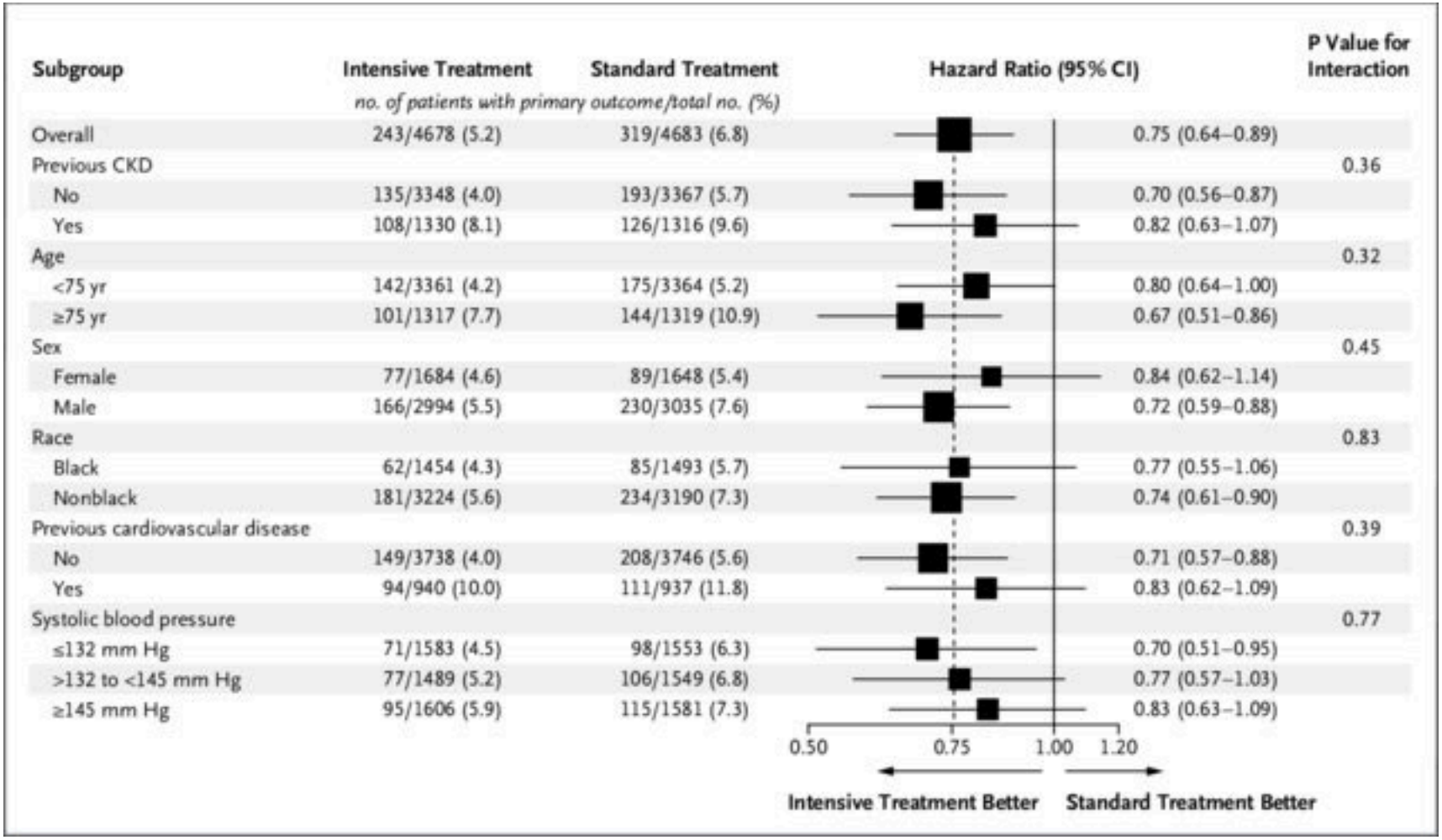
Subgroup	Intensive Treatment	Standard Treatment	Hazard Ratio (95% CI)
	no. of patients with primary outcome/total no. (%)		
Overall	243/4678 (5.2)	319/4683 (6.8)	0.75 (0.64–0.89)
Previous CKD			
No	135/3348 (4.0)	193/3367 (5.7)	0.70 (0.56–0.87)
Yes	108/1330 (8.1)	126/1316 (9.6)	0.82 (0.63–1.07)
Age			
<75 yr	142/3361 (4.2)	175/3364 (5.2)	0.80 (0.64–1.00)
≥75 yr	101/1317 (7.7)	144/1319 (10.9)	0.67 (0.51–0.86)
Sex			
Female	77/1684 (4.6)	89/1648 (5.4)	0.84 (0.62–1.14)
Male	166/2994 (5.5)	230/3035 (7.6)	0.72 (0.59–0.88)
Race			
Black	62/1454 (4.3)	85/1493 (5.7)	0.77 (0.55–1.06)
Nonblack	181/3224 (5.6)	234/3190 (7.3)	0.74 (0.61–0.90)
Previous cardiovascular disease			
No	149/3738 (4.0)	208/3746 (5.6)	0.71 (0.57–0.88)
Yes	94/940 (10.0)	111/937 (11.8)	0.83 (0.62–1.09)
Systolic blood pressure			
≤132 mm Hg	71/1583 (4.5)	98/1553 (6.3)	0.70 (0.51–0.95)
>132 to <145 mm Hg	77/1489 (5.2)	106/1549 (6.8)	0.77 (0.57–1.03)
≥145 mm Hg	95/1606 (5.9)	115/1581 (7.3)	0.83 (0.63–1.09)

Sprint Investigators (2015). A Randomized Trial of Intensive versus Standard Blood-Pressure Control. NEJM 373:2103-16.

My old view

1. Use statistical methods that capture the framework of the prior hypotheses. (In light of Kent's warnings, don't do one-subgroup at-a-time inference and consider risk.)
2. Place greater emphasis on the overall result than on what may be apparent within a particular subgroup.
3. Distinguish between prior and data-derived hypotheses. {In light of Unger's warning, don't have too many prior hypotheses!}
4. Use tests of "interactions," and/or correct for multiplicity of statistical comparisons.

A forest plot with interaction p-values



Subgroup	Intensive Treatment	Standard Treatment	Hazard Ratio (95% CI)	P Value for Interaction
	no. of patients with primary outcome/total no. (%)			
Overall	243/4678 (5.2)	319/4683 (6.8)	0.75 (0.64–0.89)	
Previous CKD				0.36
No	135/3348 (4.0)	193/3367 (5.7)	0.70 (0.56–0.87)	
Yes	108/1330 (8.1)	126/1316 (9.6)	0.82 (0.63–1.07)	
Age				0.32
<75 yr	142/3361 (4.2)	175/3364 (5.2)	0.80 (0.64–1.00)	
≥75 yr	101/1317 (7.7)	144/1319 (10.9)	0.67 (0.51–0.86)	
Sex				0.45
Female	77/1684 (4.6)	89/1648 (5.4)	0.84 (0.62–1.14)	
Male	166/2994 (5.5)	230/3035 (7.6)	0.72 (0.59–0.88)	
Race				0.83
Black	62/1454 (4.3)	85/1493 (5.7)	0.77 (0.55–1.06)	
Nonblack	181/3224 (5.6)	234/3190 (7.3)	0.74 (0.61–0.90)	
Previous cardiovascular disease				0.39
No	149/3738 (4.0)	208/3746 (5.6)	0.71 (0.57–0.88)	
Yes	94/940 (10.0)	111/937 (11.8)	0.83 (0.62–1.09)	
Systolic blood pressure				0.77
≤132 mm Hg	71/1583 (4.5)	98/1553 (6.3)	0.70 (0.51–0.95)	
>132 to <145 mm Hg	77/1489 (5.2)	106/1549 (6.8)	0.77 (0.57–1.03)	
≥145 mm Hg	95/1606 (5.9)	115/1581 (7.3)	0.83 (0.63–1.09)	

Sprint Investigators (2015). A Randomized Trial of Intensive versus Standard Blood-Pressure Control. NEJM 373:2103-16.

My old view

1. Use statistical methods that capture the framework of the prior hypotheses.
2. Place greater emphasis on the overall result than on what may be apparent within a particular subgroup.
3. Distinguish between prior and data-derived hypotheses. {In light of Unger's warning, don't have too many prior hypotheses!}
4. Use tests of "interactions," and/or correct for multiplicity of statistical comparisons.
5. …..

Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA*. 1991;266:93–98.

Evolving worry: some types of subgroups

- REGION
  - Standard of care, diet, risk factors (pace xx) differ markedly
  - Entry of China into trials is frequently late, attenuating effect of drug in trials with delayed effect
  - Large differential reporting of adverse events, even SAEs
  - TOPCAT (Russia & Georgia)

Wittes (2013). Why is this subgroup different from all other subgroups? Thoughts on regional differences in randomized clinical trials. *Proceedings of the Fourth Seattle Symposium in Biostatistics: Clinical Trials.*

More evolution: personalized medicine

- McShane – likely future subgroups…

5. Interpret the results in the context of similar data from other trials, from the architecture of the entire set of data on all patients, and from principles of biological coherence.

We have much to learn about capitalizing on these new methods