

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine — reporting of subgroup analyses in clinical trials. *N Engl J Med* 2007;357:2189-94.

Supplementary Appendix I:

METHODS

We reviewed all original articles published between July 1, 2005 and June 30, 2006 that reported primary results from randomized clinical trials. For each paper, we recorded whether the trial was single- or multi-center, a superiority or non-inferiority design, the number of randomized subjects, the disease area (classified as Cardiovascular, Infectious Disease, Oncology, Respiratory Disease, Pediatric, Psychiatry/Neurology, Metabolic, Endocrine and Gastrointestinal Disease, Gynecology, Other), whether a “significant” result was reported for at least one primary endpoint (that is, finding a nominally significant difference among treatment groups in a superiority trial or declaring non-inferiority/equivalence in a non-inferiority/equivalence trial), and whether any subgroup analyses were reported.

For trials reporting subgroup analyses, we recorded where within the article the subgroup analyses were reported, i.e., whether results were presented only in the text versus also in tables or figures, either in the main body of text or in supplementary material; whether treatment comparison information was presented for the different levels of a subgrouping variable, if so, whether it was reported consistently. We further attempted to determine the number of subgroup analyses that were performed and the number of subgroup analyses that were reported, whether it was clear that each subgroup analysis was pre-specified or post hoc, whether interaction tests were used to assess heterogeneity of treatment effects, whether any heterogeneity in treatment effects was claimed and, if so, whether statistical inferences were corrected for multiple comparisons that may have been made. When counting the number of subgroup analyses, each statistical analysis evaluating treatment differences for a specific endpoint among different levels of one baseline factor is considered as one subgroup analysis. Univariate and multivariate logistic regression models were used to assess the association between trial characteristics and reporting of subgroup analyses. All p-values are 2-sided.

Supplementary Appendix II:

Table 1. Details of Trials Reporting Heterogeneous Treatment Effects in Subgroups

Trial Number	PE ¹	# Conducted ²	# Reported ³	(# Pre-specified ⁴ , # Post Hoc ⁵)	# Heterogeneous ⁶	Basis ⁷	Where ⁸	MC ⁹
1	NS	U	13+	(0, 13+)	2	1,2	A,R,D	N
2	S	U	10	(0, 10)	1	1	R, D	N
3	S	U	2	(2, 0)	2	1,2	R, D	N
4	NS	U	16	(10, 6)	4	1,2	R, D	N
5	S	U	3+	(U, U)	2	2	R, D	N
6	NS	U	19+	(U, U)	6	1,2	R, D	N
7	NS	U	6	(U, U)	1	2	R	N
8	NS	U	62+	(U, U)	3	1,3	R, D	Y
9	S	U	5	(U, U)	2	1,2	A, R, D	N
10	NS	U	11	(11, 0)	1	1	R,D	N
11	NS	U	7	(U, U)	2	2	R	N
12	S	U	10	(U, U)	4	1,2	R,D	N
13	NS	U	18	(7, 11)	2	1	A, R, D	N
14	NS	U	27	(23, 4)	4	1,2	A, R, D	Y
15	S	U	2	(U, 1+)	1	1	R, D	N

¹ PE (primary endpoint): S=Significant; NS=Not Significant.

² # Conducted: Number of subgroup analyses conducted. U=Unclear.

³ # Reported: Number of subgroup analyses reported.

⁴ # Pre-specified: Number of pre-specified subgroup analyses among reported. U=Unclear.

⁵ # Post Hoc: Number of post hoc subgroup analyses among reported. U=Unclear.

⁶ # Heterogeneous: Number of subgroup analyses claiming heterogeneity in treatment group differences.

⁷ Basis for claim: 1=Significant interaction; 2=Within subgroup comparisons; 3=Not Specified.

⁸ Where: Where was the claimed heterogeneity reported?

A=Abstract; R=Results; D=Discussion.

⁹ MC: Were multiple comparison issues addressed? N=No; Y=Yes.