What usage and what hierarchical order for secondary endpoints?

Silvy Laporte, Marine Diviné, Danièle Girault

Pierre Boutouyrie, Olivier Chassany, Michel Cucherat,
Hervé de Trogoff, Sophie Dubois, Cécile Fouret,
Natalie Hoog-Labouret, Pascale Jolliet, Patrick Mismetti,
Raphaël Porcher, Cécile Rey-Coquais, Eric Vicaut

Journée Nationale de Restitution
des Rencontres de Pharmacologie et de Recherche Clinique
Pour l'Innovation et les Technologies de Santé
Paris, 23 mars 2016

Oral Rivaroxaban for the Treatment of Symptomatic Pulmonary Embolism

The EINSTEIN–PE Investigators*

RESULTS

Rivaroxaban was noninferior to standard therapy (noninferiority margin, 2.0; P=0.003) for the primary efficacy outcome, with 50 events in the rivaroxaban group (2.1%) versus 44 events in the standard-therapy group (1.8%) (hazard ratio, 1.12; 95% confidence interval [CI], 0.75 to 1.68). The principal safety outcome occurred in 10.3% of patients in the rivaroxaban group and 11.4% of those in the standard-therapy group (hazard ratio, 0.90; 95% CI, 0.76 to 1.07; P=0.23). Major bleeding was observed in 26 patients (1.1%) in the rivaroxaban group and 52 patients (2.2%) in the standard-therapy group (hazard ratio, 0.49; 95% CI, 0.31 to 0.79; P=0.003)

The EINSTEIN PE investigators. N Engl J Med 2012

Context

Two drugs in the treatment of NSCLC assessed in 2 randomised clinical trials

Primary end point: PFS

Drug 1 compared to drug X: improvement of PFS

Drug 2 compared to drug X: improvement of PFS and OS

Can this differential result be use?

Clinician's and manufacturer's standpoint

The result highlighted on the secondary endpoint seems valid since the primary endpoint is significant

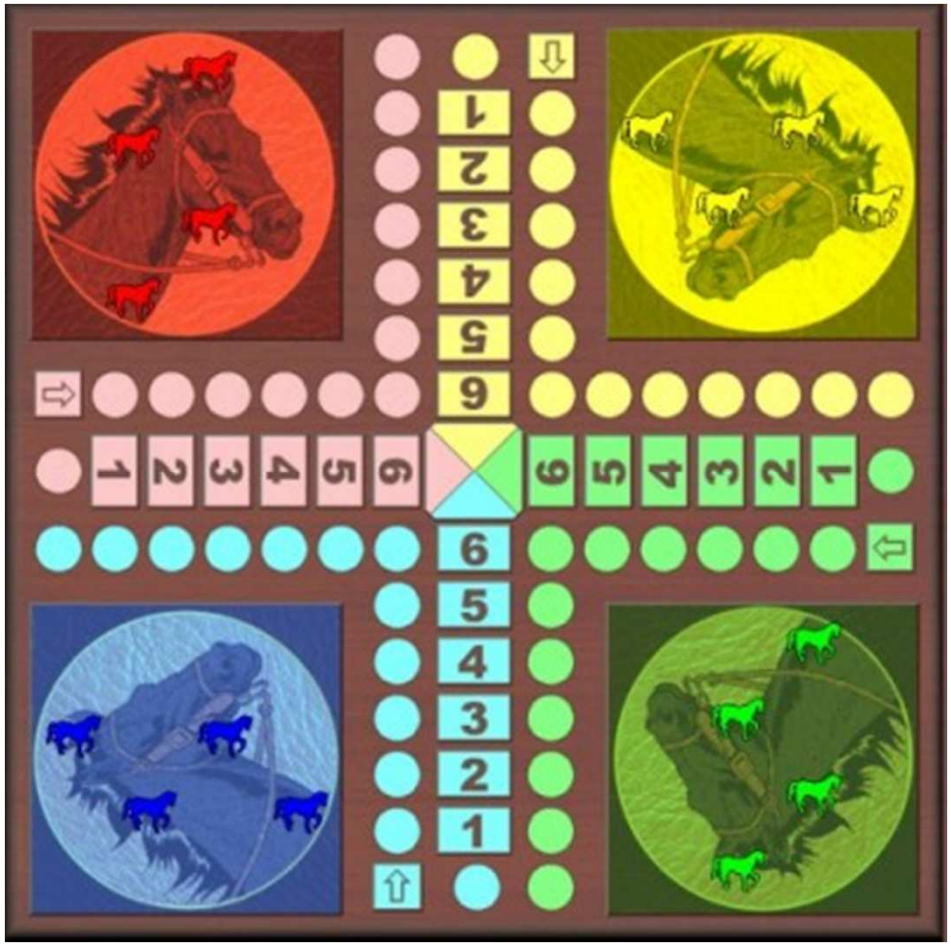
➔can be used to select the optimal treatment

Methodological standpoint

Numerous secondary endpoints then multiplicity of tests

➔ may induce a risk of erroneously concluding that there is an additional treatment benefit

The more you play, the more you win



$$\alpha_{overall} = 1 - (1 - \alpha)^k$$

1 dieroll proba = 1/6

5 dierolls proba > 1/2

1 test $\alpha = 0.05$

5 tests $\alpha = 0.23$

One single endpoint for one benefit?

A unique primary endpoint

- Only one, defined a priori, one statistical test (one dieroll)

Once the existence of a treatment benefit concluded

- Other effects that may increase its clinical interest?
- No official recognition of an additional treatment benefit based on secondary endpoints
- Even in the primary endpoint is significant

☹ **A single variable to summarise the potential benefit of a treatment**

Very simplistic

A set of primary or co-primary endpoints?

At least 2 endpoints are needed to declare a treatment benefit

- Test of each of the co-primary endpoints must be significant
- Type I error 0.05 for all endpoints
- No threshold reduction required

Example : 2 doses of AINS on pain

- numeric pain scale over 5 days
- total intake of concomitant analgesics over 5 days

Riou B et al. Comparison of two doses of ketoprofen to treat pain: a double-blind, randomized, noninferiority trial. Fundam Clin Pharmacol 2014.

Risk of type II error in the trial…

A set of primary or co-primary endpoints?

When one of these co-primary endpoints is sufficient to declare a treatment benefit

- Distribution of the risk of type I error between the co-endpoints
- One significant endpoint is sufficient
- Conventional approaches for multiple tests

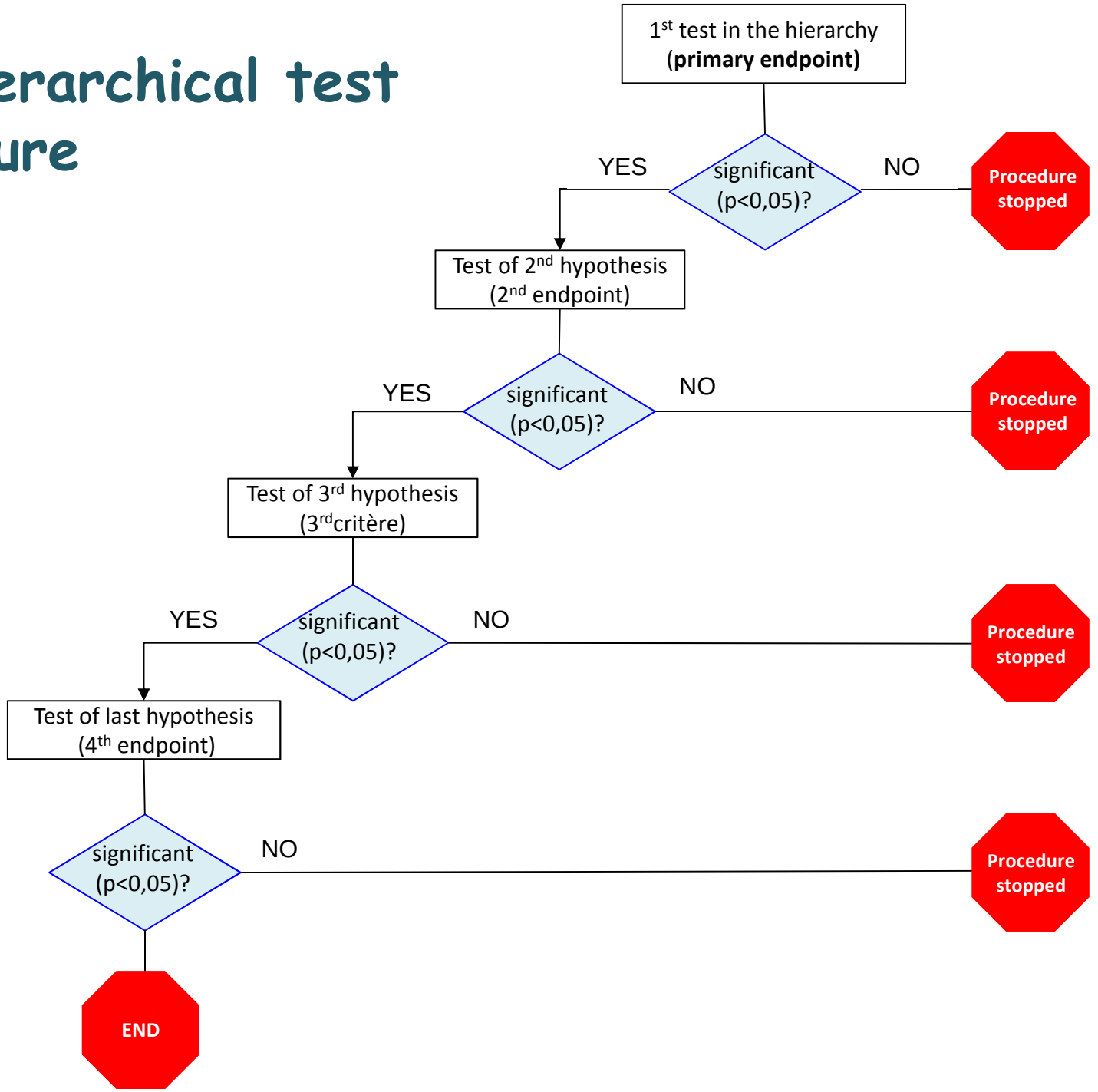
Example :	OS	α 0.025	α 0.046
	PFS	α 0.025	α 0.005

Gilbert MR et al. A Randomized Trial of Bevacizumab for Newly Diagnosed Glioblastoma. N Engl J Med 2014

☹ **Increase in the sample size (α 0.01 for 5 endpoints)**

Maximum 2-3 co-endpoints…

The hierarchical test procedure



The hierarchical test procedure: example

Table 3. Major Efficacy End Points at 12 Months.*

End Point	Ticagrelor Group	Clopidogrel Group	Hazard Ratio for Ticagrelor Group (95% CI)	P Value†	
Primary end point: death from vascular causes, MI, or stroke — no./total no. (%)	864/9333 (9.8)	1014/9291 (11.7)	0.84 (0.77–0.92)	<0.001‡	1
Secondary end points — no./total no. (%)					
Death from any cause, MI, or stroke	901/9333 (10.2)	1065/9291 (12.3)	0.84 (0.77–0.92)	<0.001‡	2
Death from vascular causes, MI, stroke, severe recurrent ischemia, recurrent ischemia, TIA, or other arterial thrombotic event	1290/9333 (14.6)	1456/9291 (16.7)	0.88 (0.81–0.95)	<0.001‡	3
MI	504/9333 (5.8)	593/9291 (6.9)	0.84 (0.75–0.95)	0.005‡	4
Death from vascular causes	353/9333 (4.0)	442/9291 (5.1)	0.79 (0.69–0.91)	0.001‡	5
Stroke	125/9333 (1.5)	106/9291 (1.3)	1.17 (0.91–1.52)	0.22	6

PLATO. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 2009

How to construct the hierarchy: a strategic choice

How to select endpoints to include in the hierarchy

- Endpoints for which a benefit is sought, including safety endpoints
- Endpoints related to relevant therapeutic objective for the patient

How to construct the hierarchy: a strategic choice

How to order the endpoints in the hierarchy

- 1st endpoint: same rules as for the primary endpoint, relevant clinical endpoint, sufficiently common to be conclusive
- Rest of the hierarchy: taken into account a potential lack of power for some of them, frequent events can be preferred as long as clinically interesting,
- Compromise while staying focused on the benefit to the patients
- Ultimately, once endpoints are ordered hierarchically, all endpoints selected have the same level in terms of statistical evidence

Example :

1. Composite endpoints of morbidity and mortality
2. Overall mortality

How to construct the hierarchy: a strategic choice

Discuss the hierarchy with the Healthcare Authorities

- European Medicines Agency (EMA) for Scientific Advice
- Health Technology Assessment (HTA) agencies for Scientific Advice (Shaping European Early Dialogues)
- Conjointly (parallel EMA multi HTA early dialogue)

Conclusion

Consideration of several criteria for the primary endpoint

- Description of the treatment effect on the pathology requiring several parameters
- Use of co-endpoints accompanied by the essential type I error adjustments

Secondary endpoints to differentiate treatments that demonstrate an effect on the primary endpoint

- The hierarchical test procedure is the only simple method currently available that can demonstrate a medical benefit on several endpoints
- If respected…

Table 3. Major Efficacy End Points at 12 Months.*

End Point	Ticagrelor Group	Clopidogrel Group	Hazard Ratio for Ticagrelor Group (95% CI)	P Value†
Primary end point: death from vascular causes, MI, or stroke — no./total no. (%)	864/9333 (9.8)	1014/9291 (11.7)	0.84 (0.77–0.92)	<0.001‡
Secondary end points — no./total no. (%)				
Death from any cause, MI, or stroke	901/9333 (10.2)	1065/9291 (12.3)	0.84 (0.77–0.92)	<0.001‡
Death from vascular causes, MI, stroke, severe recurrent ischemia, recurrent ischemia, TIA, or other arterial thrombotic event	1290/9333 (14.6)	1456/9291 (16.7)	0.88 (0.81–0.95)	<0.001‡
MI	504/9333 (5.8)	593/9291 (6.9)	0.84 (0.75–0.95)	0.005‡
Death from vascular causes	353/9333 (4.0)	442/9291 (5.1)	0.79 (0.69–0.91)	0.001‡
Stroke	125/9333 (1.5)	106/9291 (1.3)	1.17 (0.91–1.52)	0.22
Ischemic	96/9333 (1.1)	91/9291 (1.1)		0.74
Hemorrhagic	23/9333 (0.2)	13/9291 (0.1)		0.10
Unknown	10/9333 (0.1)	2/9291 (0.02)		0.04
Other events — no./total no. (%)				
Death from any cause	399/9333 (4.5)	506/9291 (5.9)	0.78 (0.69–0.89)	

1
2
3
4
5
6
STOP
7

(4.0% vs. 5.1%, P=0.001) but not stroke alone (1.5% vs. 1.3%, P=0.22). The rate of death from any cause was also reduced with ticagrelor (4.5%, vs. 5.9% with clopidogrel; P<0.001). No significant difference in the rates of major bleeding was found

Recommendations

1. Use the hierarchical test to be able to claim the additional benefits relating to the secondary endpoints
2. To implement a hierarchical test sequence not initially scheduled, a protocol amendment may be considered before unblinding of the data
3. Respect the order of the hierarchy predefined in the protocol : after the first non-significant test, results can only be descriptive even with a p value < 0.05
4. The hierarchical test procedure is admissible at a statistical level but does not exempt of possible biases in the study, the clinical relevance of the endpoints, or the significance of the effects
5. Not all secondary endpoints need to be included in the hierarchy. The results for these endpoints will therefore be descriptive only.
6. Training for the method is crucial in order to obtain a legitimate interpretation of the results of the approach.