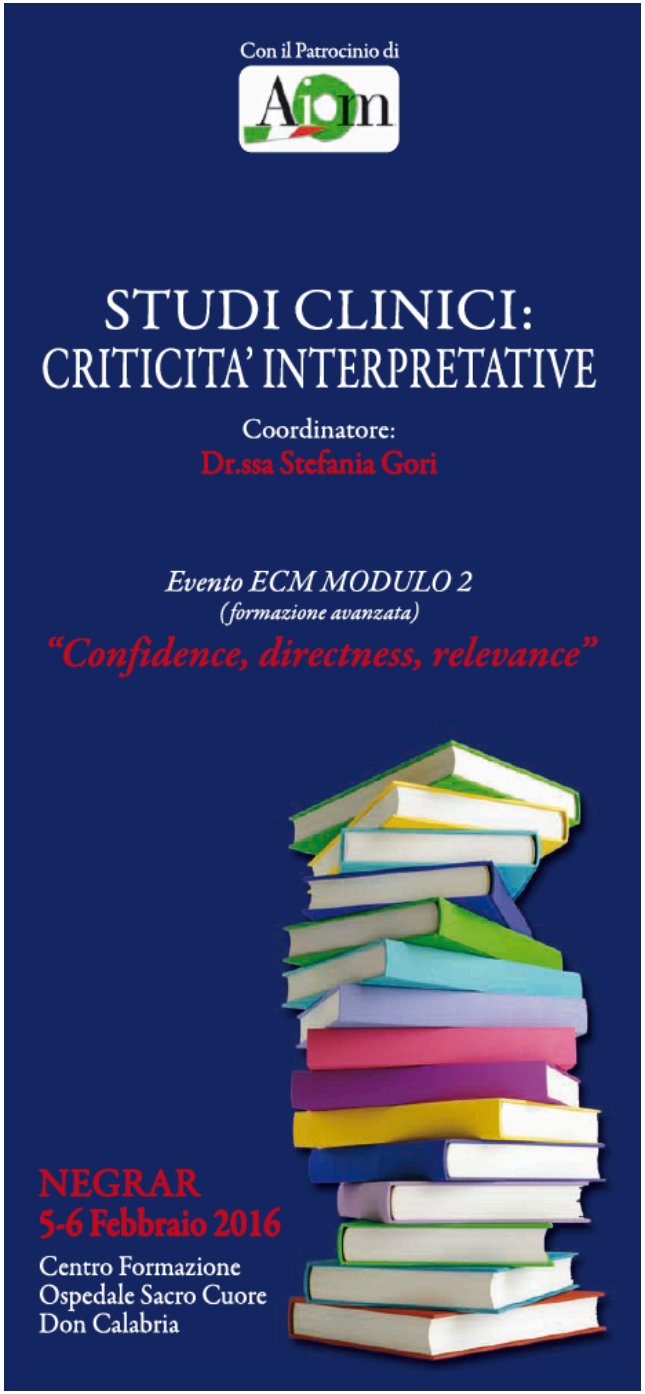


Venerdì 5 febbraio 2016

- Il quesito clinico come *primum movens* di ogni decisione terapeutica
- *Confidence*
  - ✓ rischio di bias
  - ✓ analisi per sottogruppi
  - ✓ imprecisione delle stime
  - ✓ eterogeneità delle evidenze
- *Directness*
  - ✓ adeguatezza delle evidenze al quesito P.I.C.O.
  - ✓ confronti indiretti, *network meta-analysis*
- *Relevance*
  - ✓ il *target* di rilevanza clinica

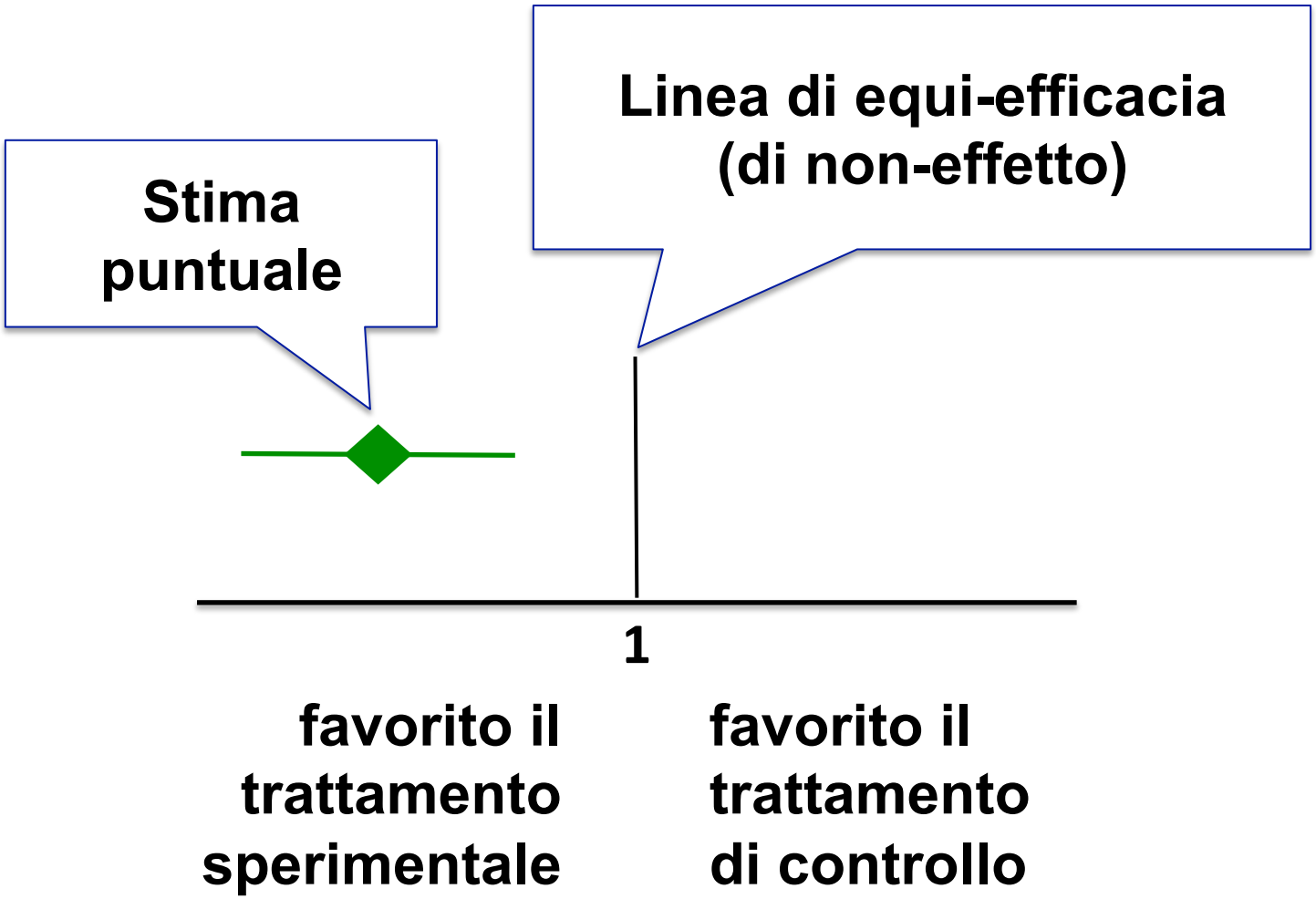
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot

Why a forest plot?

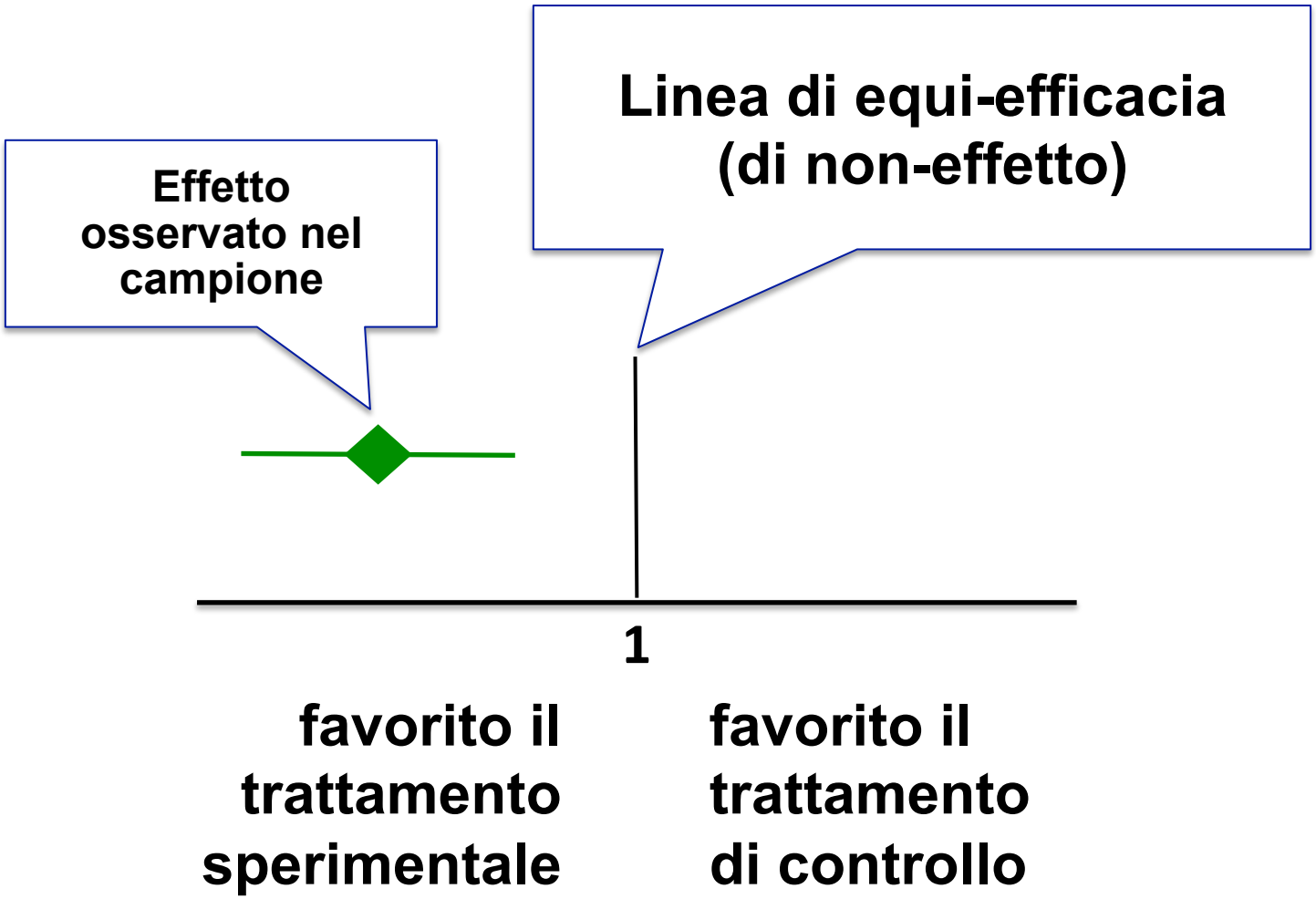
The plot was not called a "forest plot" in print for some time, and the origins of this title are obscured by history and myth. At the September 1990 meeting of the breast cancer overview, Richard Peto jokingly mentioned that the plot was named after the breast cancer researcher Pat Forrest, and, at times, the name has been spelt "forrest plot." However, the phrase actually originates from the idea that the typical plot appears as a forest of lines.

BMJ VOLUME 322 16 JUNE 2001

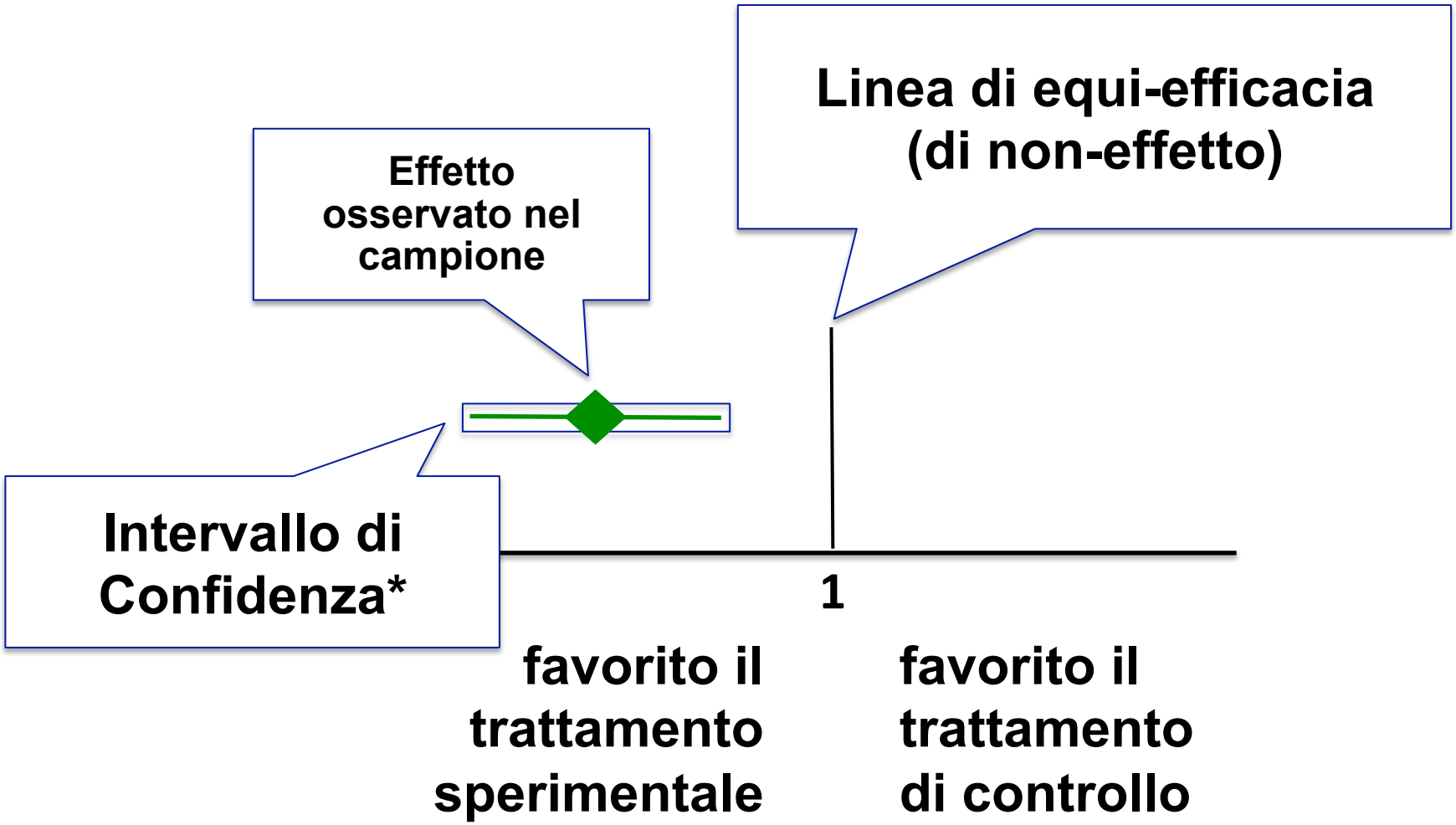
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione degli studi clinici mediante Forest (Forrest?) Plot

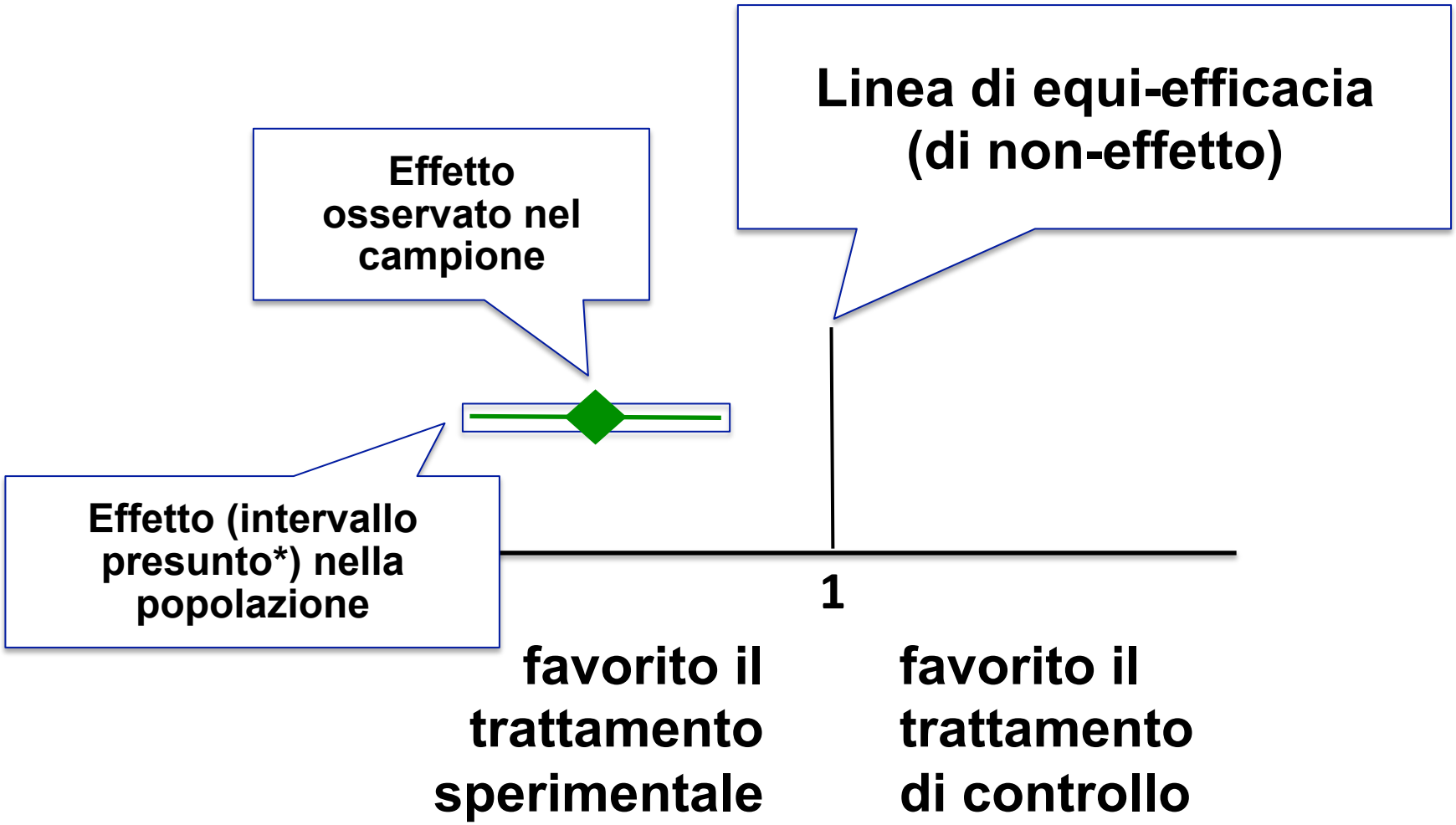


Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



* convenzionalm. 95%

Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



* convenzionalm. 95%

Statistical Vs Clinical Significance

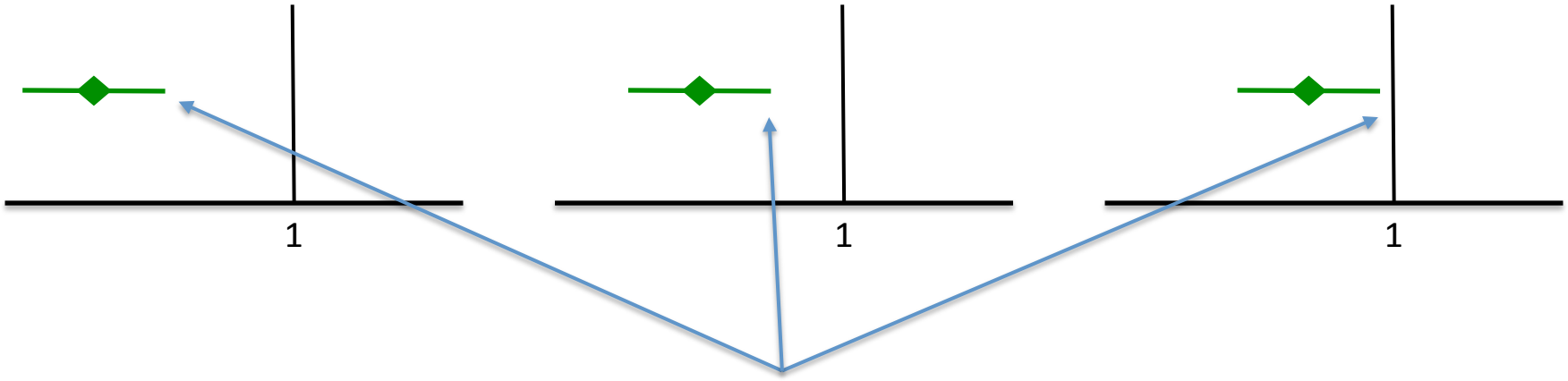
- **Statistical Significance**

  "Is an observed difference likely to be real"

  ✓ dependent on the magnitude of the number of patients and/or the **magnitude of the difference** NOT on whether the difference is meaningful for patients

Interpretazione statistica di uno Studio di Superiorità

Tutti e tre gli esempi indicano una differenza statisticamente significativa



L'estremo dx dell'intervallo di confidenza NON interseca la linea di non-effetto (P<0.05)

Statistical Vs Clinical Significance

- **Statistical Significance**

  "Is an observed difference likely to be real"

  ✓ dependent on the magnitude of the number of patients and/or the **magnitude of the difference** NOT on whether the difference is meaningful for patients

- **Clinical Significance**

  "Is an observed difference likely to be meaningful for patients"

  ✓ dependent on the **magnitude of the difference** NOT the number of patients

Rilevanza Clinica

Si ritiene che il trattamento in esame "A" abbia le potenzialità per migliorare il trattamento standard "B" almeno di una quantità Δ

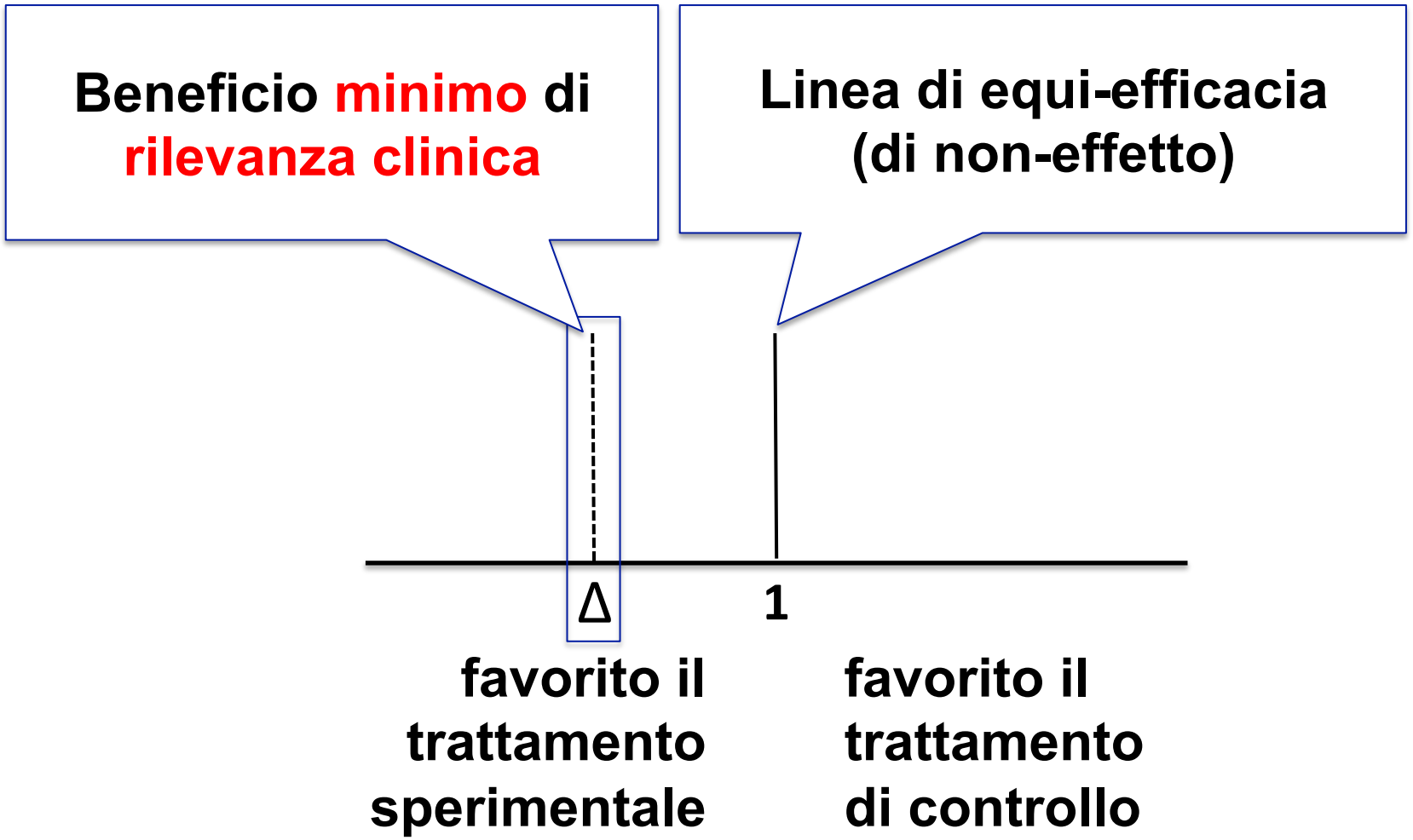
studio di superiorità

A > B di una quantità Δ di interesse clinico

studio di non inferiorità

A < B non oltre una quantità M di rilevanza clinica

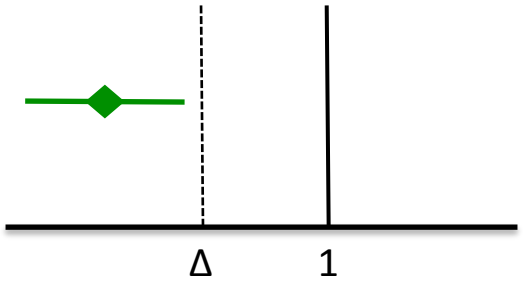
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione clinica di uno Studio di Superiorità

Effetto (sempre) clinicamente rilevante?

(dato uno specifico Δ di interesse)



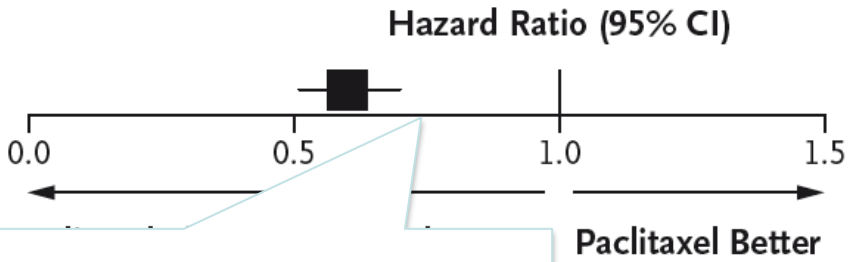
RILEVANTE e
(del tutto) **AFFIDABILE**

Paclitaxel plus Bevacizumab versus Paclitaxel Alone for Metastatic Breast Cancer

Kathy Miller, M.D., Molin Wang, Ph.D., Julie Gralow, M.D., Maura Dickler, M.D., Melody Cobleigh, M.D., Edith A. Perez, M.D., Tamara Shenkier, M.D., David Cella, Ph.D., and Nancy E. Davidson, M.D.

N Engl J Med 2007;357:2666-76.

Progression-free Survival (mo)		Hazard Ratio (95% CI)
Paclitaxel	Paclitaxel plus Bevacizumab	
5.9	11.8	0.60 (0.51–0.70)



Beneficio minimo preordinato (*Target* Δ) = 0.75

Monotherapy with the PCSK9 inhibitor alirocumab versus ezetimibe in patients with hypercholesterolemia: Results of a 24 week, double-blind, randomized Phase 3 trial

Eli M. Roth [a,*], Marja-Riitta Taskinen [b], Henry N. Ginsberg [c], John J.P. Kastelein [d], Helen M. Colhoun [e], Jennifer G. Robinson [f], Laurence Merlet [g], Robert Pordy [h], Marie T. Baccara-Dinet [i]

International Journal of Cardiology 176 (2014) 55–61

A sample size of 45 patients per treatment arm was calculated to have 95% power to detect a mean difference between alirocumab and ezetimibe of 20% in LDL-C percent change from baseline to week 24 using a 2-sided *t*-test with 5% significance, assuming a common standard deviation (SD) of 25% based on a previous alirocumab trial [1] and with an expected rate of exclusion of 5%.

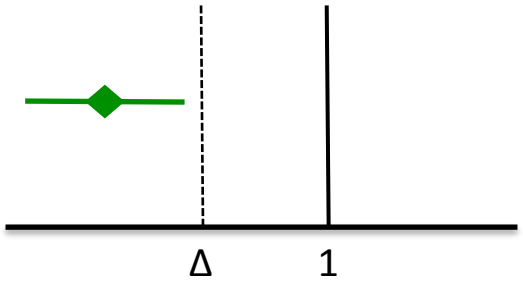
Percent change in LDL-C from baseline to week 24 (ITT and on-treatment analysis).

LDL-C	Alirocumab versus ezetimibe		
	LS mean difference (SE) %	95% CI	*p*-Value
ITT			
LS mean (SE) change from baseline (%)	−31.6 (4.3)	−40.2 to −23.0	<0.0001[a]

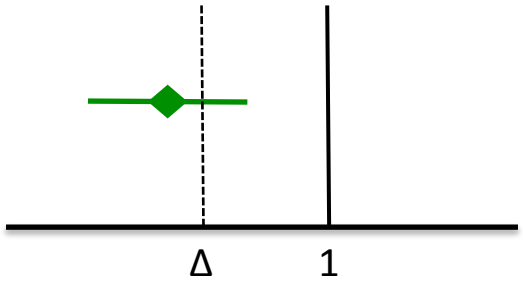
Interpretazione clinica di uno Studio di Superiorità

Effetto (sempre) clinicamente rilevante?

(dato uno specifico Δ di interesse)



RILEVANTE e
(del tutto) **AFFIDABILE**



RILEVANTE e
(ragionevolmente) **AFFIDABILE**

Oral teriflunomide for patients with relapsing multiple sclerosis (TOWER): a randomised, double-blind, placebo-controlled, phase 3 trial

Christian Confavreux, Paul O'Connor, Giancarlo Comi, Mark S Freedman, Aaron E Miller, Tomas P Olsson, Jerry S Wolinsky, Teresa Bagulho, Jean-Luc Delhay, Deborah Dukovic, Philippe Truffinet, Ludwig Kappos, for the TOWER Trial Group†*

Lancet Neurol **2014; 13: 247–56**

We estimated that 370 patients randomly assigned to each treatment group would provide 94% power to detect a 25% relative risk reduction in annualised relapse rate at the two-tailed significance level of $\alpha=0\cdot05$, assuming an annualised relapse rate of 0·74 in the placebo group.

	Placebo (n=388)	Teriflunomide 7 mg (n=407)	Teriflunomide 14 mg (n=370)
Annualised relapse rate (primary endpoint)			
Adjusted annualised relapse rate* (95% CI)	0·50 (0·43 to 0·58)	0·39 (0·33 to 0·46)	0·32 (0·27 to 0·38)
Relative risk (95% CI)	NA	0·78 (0·63 to 0·96)	0·64 (0·51 to 0·79)
Relative reduction versus placebo, % (95% CI)	NA	22·3% (4·2 to 37·0)	36·3% (20·7 to 48·8)
p value versus placebo	NA	0·0183	0·0001

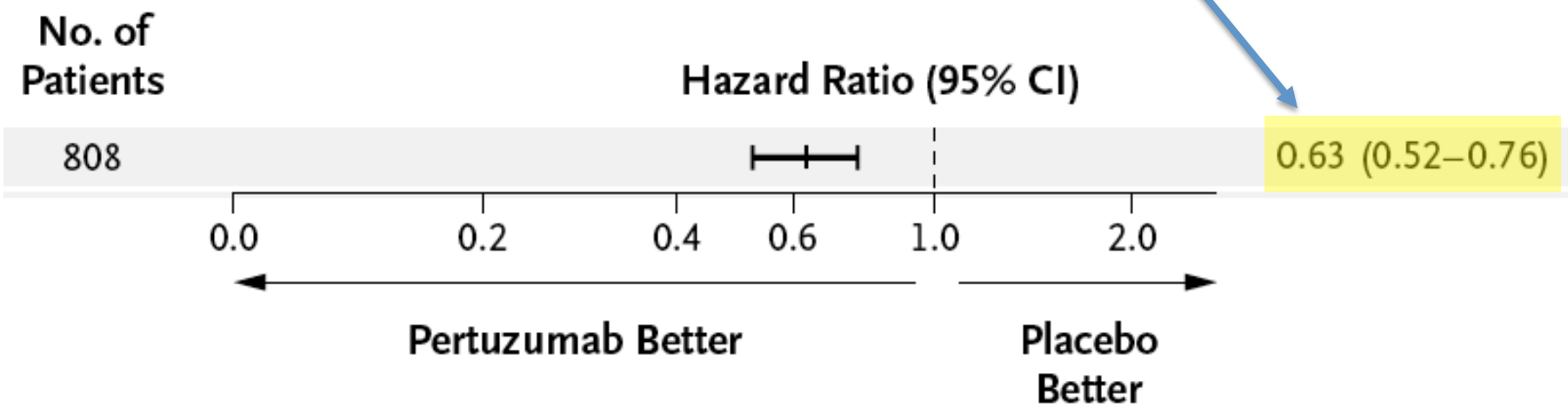
Pertuzumab plus Trastuzumab plus Docetaxel for Metastatic Breast Cancer

José Baselga, M.D., Ph.D., Javier Cortés, M.D., Sung-Bae Kim, M.D., Seock-Ah Im, M.D., Roberto Hegg, M.D., Young-Hyuck Im, M.D., Laslo Roman, M.D., José Luiz Pedrini, M.D., Tadeusz Pienkowski, M.D., Adam Knott, Ph.D., Emma Clark, M.Sc., Mark C. Benyunes, M.D., Graham Ross, F.F.P.M., and Sandra M. Swain, M.D., for the CLEOPATRA Study Group*

N Engl J Med 2012;366:109-19.

STATISTICAL ANALYSIS

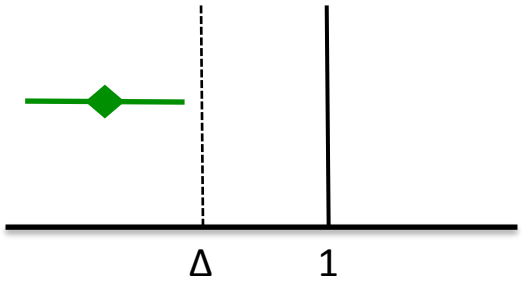
We planned to enroll 800 patients in the study and to perform the primary analysis of progression-free survival after the occurrence of approximately 381 events of independently assessed disease progression or death from any cause within 18 weeks after the last independent assessment of tumors; with that number of events, it was estimated that the study would have 80% power to detect a 33% improvement in median progression-free survival in the pertuzumab group (hazard ratio, 0.75) at a two-sided significance level of 5%.



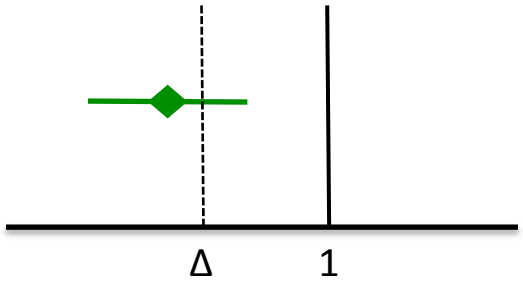
Interpretazione clinica di uno Studio di Superiorità

Effetto (sempre) clinicamente rilevante?

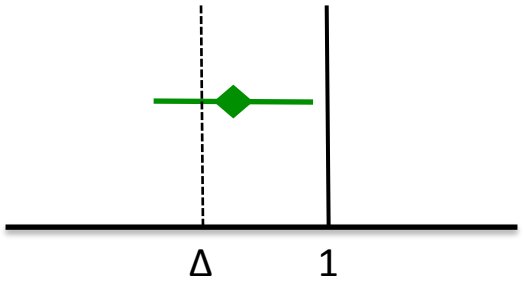
(dato uno specifico Δ di interesse)



RILEVANTE e
(del tutto) **AFFIDABILE**



RILEVANTE e
(ragionevolmente) **AFFIDABILE**



STATISTICAMENTE SIGNIFICATIVO

Oral teriflunomide for patients with relapsing multiple sclerosis (TOWER): a randomised, double-blind, placebo-controlled, phase 3 trial

Christian Confavreux, Paul O'Connor, Giancarlo Comi, Mark S Freedman, Aaron E Miller, Tomas P Olsson, Jerry S Wolinsky, Teresa Bagulho, Jean-Luc Delhay, Deborah Dukovic, Philippe Truffinet, Ludwig Kappos, for the TOWER Trial Group†*

Lancet Neurol **2014; 13: 247–56**

We estimated that 370 patients randomly assigned to each treatment group would provide 94% power to detect a 25% relative risk reduction in annualised relapse rate at the two-tailed significance level of $\alpha=0\cdot05$, assuming an annualised relapse rate of 0·74 in the placebo group.

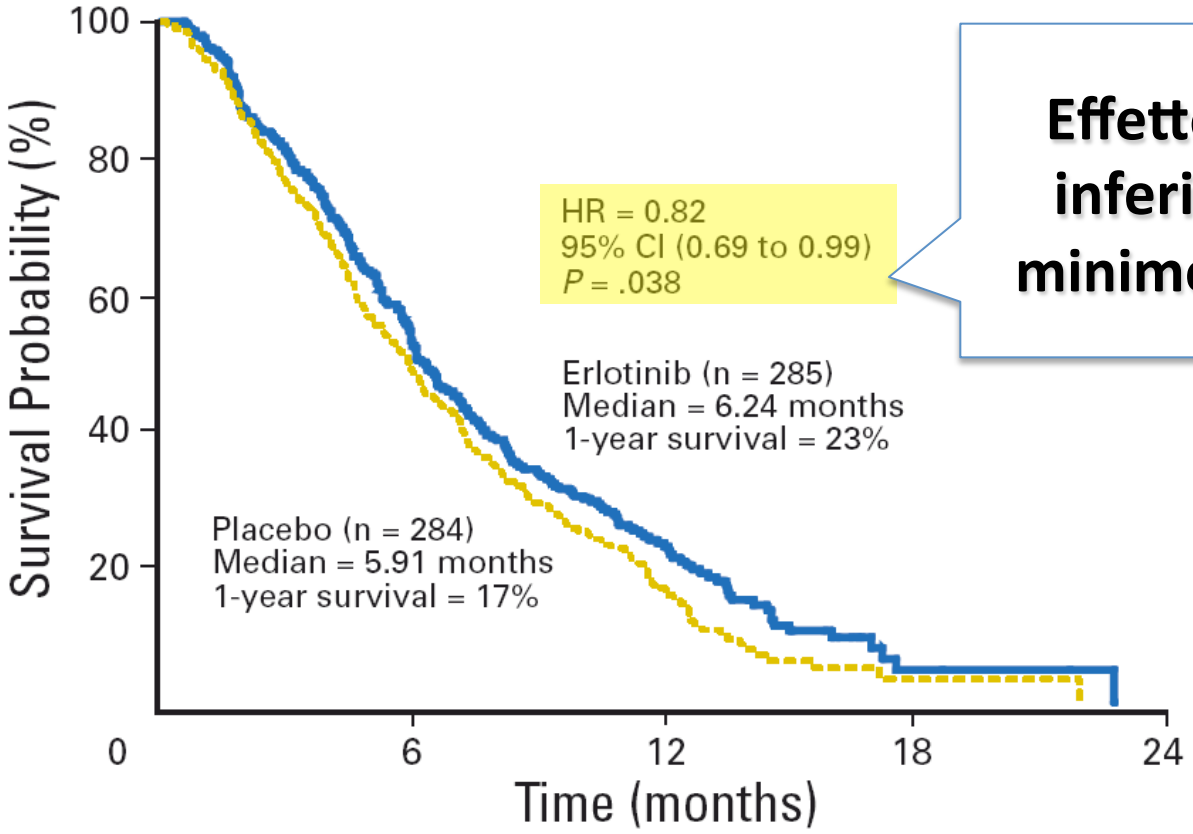
	Placebo (n=388)	Teriflunomide 7 mg (n=407)	Teriflunomide 14 mg (n=370)
Annualised relapse rate (primary endpoint)			
Adjusted annualised relapse rate* (95% CI)	0·50 (0·43 to 0·58)	0·39 (0·33 to 0·46)	0·32 (0·27 to 0·38)
Relative risk (95% CI)	NA	0·78 (0·63 to 0·96)	0·64 (0·51 to 0·79)
Relative reduction versus placebo, % (95% CI)	NA	22·3% (4·2 to 37·0)	36·3% (20·7 to 48·8)
p value versus placebo	NA	0·0183	0·0001

Erlotinib Plus Gemcitabine Compared With Gemcitabine Alone in Patients With Advanced Pancreatic Cancer: A Phase III Trial of the National Cancer Institute of Canada Clinical Trials Group

Malcolm J. Moore, David Goldstein, John Hamm, Arie Figer, Joel R. Hecht, Steven Gallinger, Heather J. Au, Pawel Murawa, David Walde, Robert A. Wolff, Daniel Campos, Robert Lim, Keyue Ding, Gary Clark, Theodora Voskoglou-Nomikos, Mieke Ptasynski, and Wendy Parulekar

J Clin Oncol 25:1960-1966. © 2007 by American Society of Clinical Oncology

Target Δ: HR erlotinib:placebo = 0.75 (**2 months** OS improvement)
Analysis after **381 events** (450 patients; α 5%, power 80%)



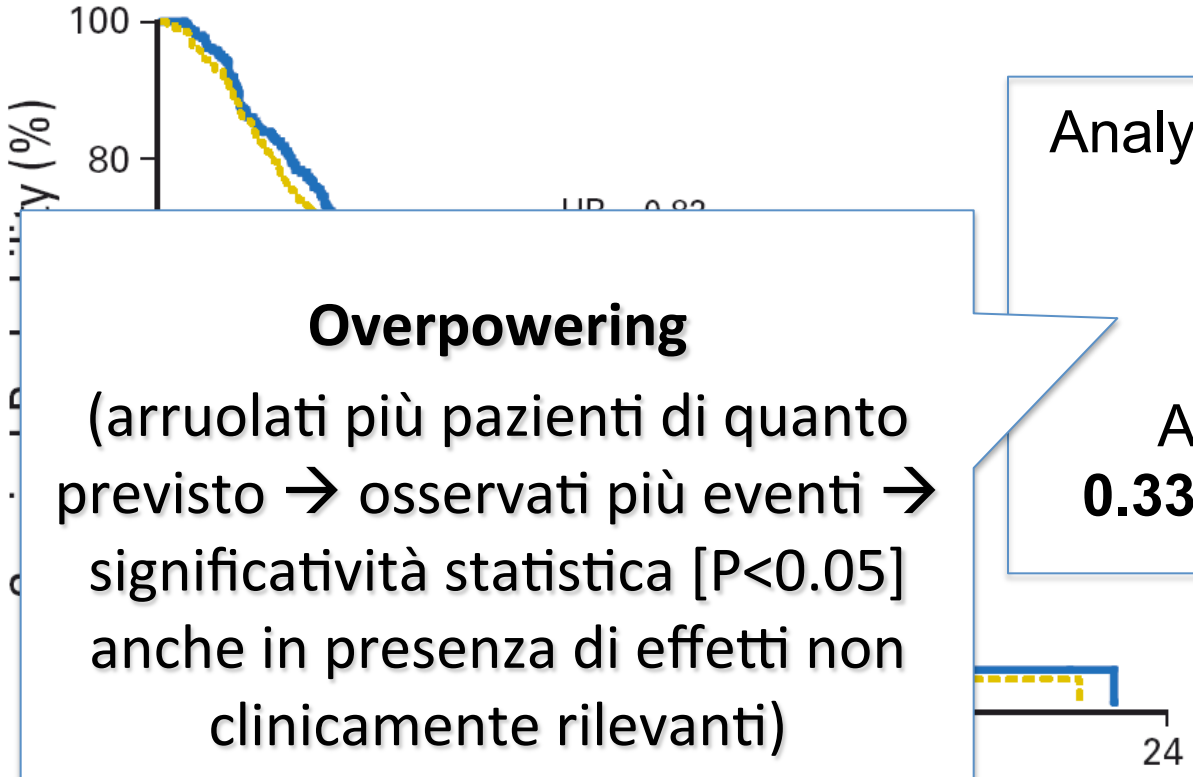
Effetto di dimensione inferiore al beneficio minimo preordinato (Δ)

Erlotinib Plus Gemcitabine Compared With Gemcitabine Alone in Patients With Advanced Pancreatic Cancer: A Phase III Trial of the National Cancer Institute of Canada Clinical Trials Group

Malcolm J. Moore, David Goldstein, John Hamm, Arie Figer, Joel R. Hecht, Steven Gallinger, Heather J. Au, Pawel Murawa, David Walde, Robert A. Wolff, Daniel Campos, Robert Lim, Keyue Ding, Gary Clark, Theodora Voskoglou-Nomikos, Mieke Ptasynski, and Wendy Parulekar

J Clin Oncol 25:1960-1966. © 2007 by American Society of Clinical Oncology

Target Δ: HR erlotinib:placebo = 0.75 (**2 months** OS improvement)
Analysis after **381 events** (450 patients; α 5%, power 80%)



Overpowering

(arruolati più pazienti di quanto previsto → osservati più eventi → significatività statistica [P<0.05] anche in presenza di effetti non clinicamente rilevanti)

Analysis after **486 events** (569 patients)

Actual difference: **0.33 months** (10 days)

Symptom Endpoints (Patient-Reported Outcomes)

- Blinding is often difficult
- Data are often missing or incomplete
- Clinical significance of small changes unknown
- Few validated instruments

Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma

Robert J. Motzer, M.D., Thomas E. Hutson, D.O., David Cella, Ph.D., James Reeves, M.D., Robert Hawkins, M.B., B.S., Ph.D., Jun Guo, Ph.D., Paul Nathan, M.B., B.S., Ph.D., Michael Staehler, M.D., Paul de Souza, M.B., B.S., Ph.D., Jaime R. Merchan, M.D., Ekaterini Boleti, M.D., Ph.D., Kate Fife, M.D., Jie Jin, M.D., Robert Jones, Ph.D., Hirotsugu Uemura, M.D., Ph.D., Ugo De Giorgi, M.D., Ulrika Harmenberg, M.D., Ph.D., Jinwan Wang, M.D., Cora N. Sternberg, M.D., Keith Deen, M.S., Lauren McCann, Ph.D., Michelle D. Hackshaw, Ph.D., Rocco Crescenzo, D.O., Lini N. Pandite, M.D., and Toni K. Choueiri, M.D.

N Engl J Med 2013;369:722-31

Table 2. Change in Health-Related Quality of Life during the First 6 Months for 927 Patients Treated in the Study.*

Instrument	Pazopanib	Sunitinib	Difference in Mean Change from Baseline Score with Pazopanib vs. Sunitinib†	P Value‡	Drug Favored According to Significant Difference§	Effect Size¶
	number of patients					
FACIT-F**	377	403	2.32	<0.001	Pazopanib	0.24
FKSI-19**						
Treatment side effects	351	382	0.31	0.03	Pazopanib	0.14
Disease-related physical symptoms	378	407	0.78	0.03	Pazopanib	0.13
Disease-related emotional symptoms	370	402	–0.05	0.41	Neither	–0.04
Functional well-being	378	403	0.31	0.10	Neither	0.09
Total score	377	408	1.41	0.02	Pazopanib	0.14

Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma

Robert J. Motzer, M.D., Thomas E. Hutson, D.O., David Cella, Ph.D., James Reeves, M.D., Robert Hawkins, M.B., B.S., Ph.D., Jun Guo, Ph.D.,

Probabilità che l'effetto osservato (*difference in mean score*) sia dovuto al caso

(non esprime l'entità dell'effetto!)

Table 2. Change in Health-R … dy.*

Instrument	Pazopanib	Sunitinib	Difference in Mean Change from Baseline Score with Pazopanib vs. Sunitinib‡	P Value§	Drug Favored According to Significant Difference¶	Effect Size‖
	number of patients					
FACIT-F**	377	403	2.32	<0.001	Pazopanib	0.24
FKSI-19**						
Treatment side effects	351	382	0.31	0.03	Pazopanib	0.14
Disease-related physical symptoms	378	407	0.78	0.03	Pazopanib	0.13
Disease-related emotional symptoms	370	402	–0.05	0.41	Neither	–0.04
Functional well-being	378	403	0.31	0.10	Neither	0.09
Total score	377	408	1.41	0.02	Pazopanib	0.14

Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma

Robert J. Motzer, M.D., Thomas E. Hutson, D.O., David Cella, Ph.D., James Reeves, M.D., Robert Hawkins, M.B., B.S., Ph.D., Jun Guo, Ph.D., Paul Nathan, M.B., B.S., Ph.D., Michael Staehler, M.D., Paul de Souza, M.B., B.S., Ph.D., Jaime R. Merchan, M.D., Ekaterini Boleti, M.D., Ph.D., Kate Fife, M.D., Jie Jin, M.D., Robert Jones, Ph.D., Hirotsugu Uemura, M.D., Ph.D., Ugo De Giorgi, M.D., Ulrika Harmenberg, M.D., Ph.D., Jinwan Wang, M.D., Cora N. Sternberg, M.D., Keith Deen, M.S., Lauren McCann, Ph.D., Michelle D. Hackshaw, Ph.D., Rocco Crescenzo, D.O., Lini N. Pandite, M.D., and Toni K. Choueiri, M.D.

N Engl J Med 2013;369:722-31

Table 2. Change in Health-Related Quality of Life during the First 6 Months for 927 Patients Treated in the Study.*

Instrument	Pazopanib	Sunitinib	Difference in Mean Change from Baseline Score with Pazopanib vs. Sunitinib†	P Value‡	Drug Favored According to Significant Difference§	Effect Size¶
	number of patients					
FACIT-F**	377	403	2.32	<0.001	Pazopanib	0.24
FKSI-19**						
Treatment side effects				0.03	Pazopanib	0.14
Disease-related physical symptoms				0.03	Pazopanib	0.13
Disease-related emotional symptoms				0.41	Neither	–0.04
Functional well-being	378	403	0.31	0.10	Neither	0.09
Total score	377	408	1.41	0.02	Pazopanib	0.14

Rilevanza dell'effetto da rapportare alla M.I.D. specifica

The Minimal (Clinical) Interesting Difference (M.I.D. / M.C.I.D.)

- **Easily understood by clinicians as a key concept in the interpretability of PRO scores.**
- **The smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management**

The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation

Kimberly Webster, David Cella* and Kathleen Yost

Health and Quality of Life Outcomes 2003, **1**:79

Table 1: Minimally important differences for select FACIT scales

Instrument	Scale/Subscale	MID (points)	Reference
FACT-G	PWB	2–3	[28]
	SWB	NA	
	EWB	2*	[28,29]
	FWB	2–3	[28]
	Total FACT-G	3–7	[27,28,30,31]
FACT-Anemia	Fatigue Subscale	3–4	[27,31]
	TOI-Fatigue	5	[27]
	TOI-Anemia	6	
	Total FACT-Anemia	7	
FACT-Breast	Breast cancer subscale	2–3	[30]
	TOI-Breast	5–6	
	Total FACT-Breast	7–8	
FACT-Colorectal	Colorectal cancer subscale	2–3	[32]
	TOI-Colorectal	4–6	
	Total FACT-Colorectal	5–8	
FACT-Head & Neck	Total FACT-Head & Neck	6–12	[33]
FACT-Lung	Lung cancer subscale	2–3	[34]
	TOI-Lung	5–6	

*This MID should be considered tentative as it may be revised based on future research.

Vista la **migliore tollerabilità** del trattamento in esame “A”, si è disposti ad accettarne una eventuale minore efficacia rispetto al trattamento standard “B” purché questa non vada oltre un **margine M**

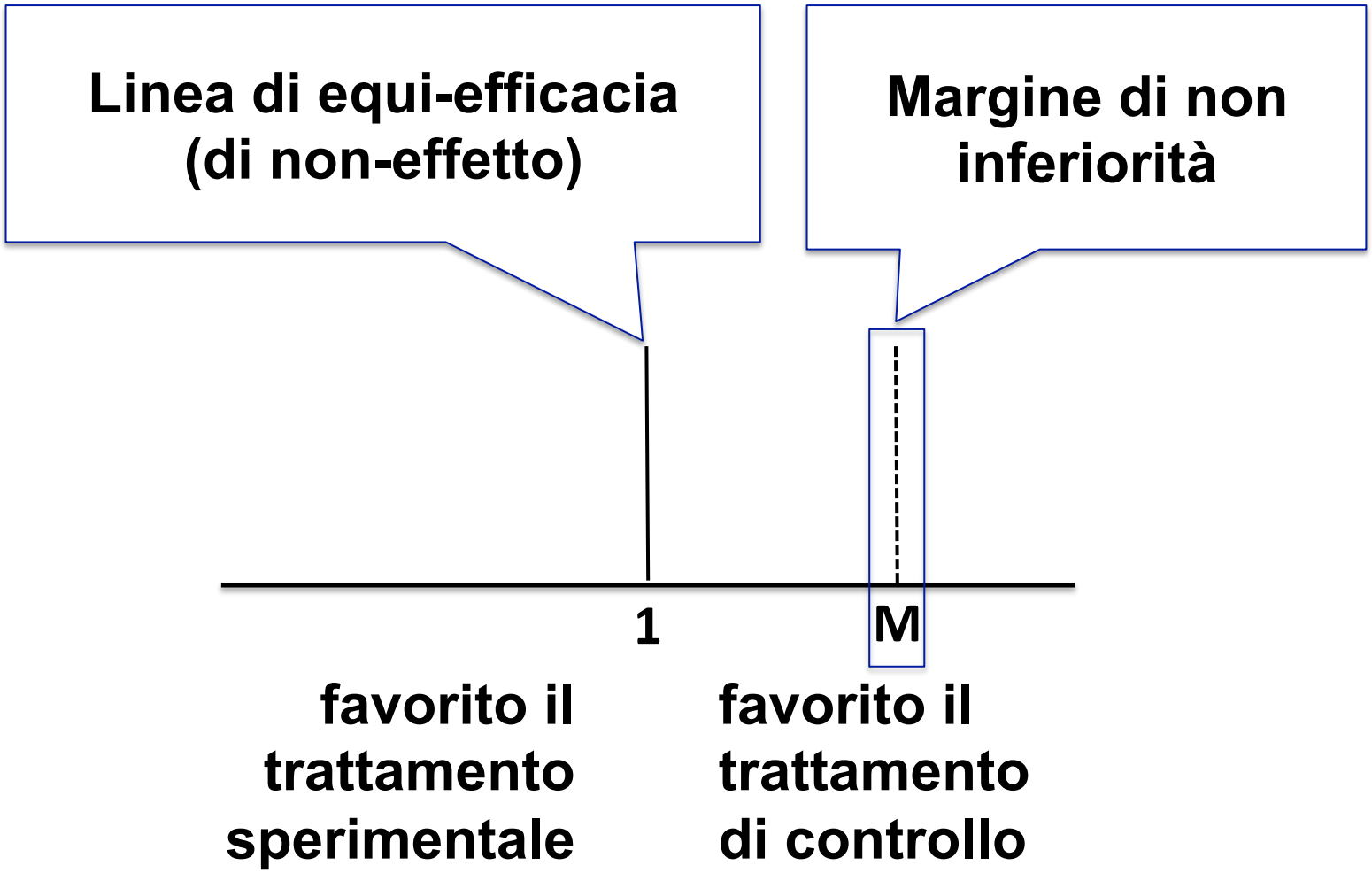
studio di superiorità

A > B di una quantità Δ di interesse clinico

studio di non inferiorità

A < B non oltre una quantità M di rilevanza clinica

Interpretazione degli studi clinici mediante Forest (Forrest?) Plot

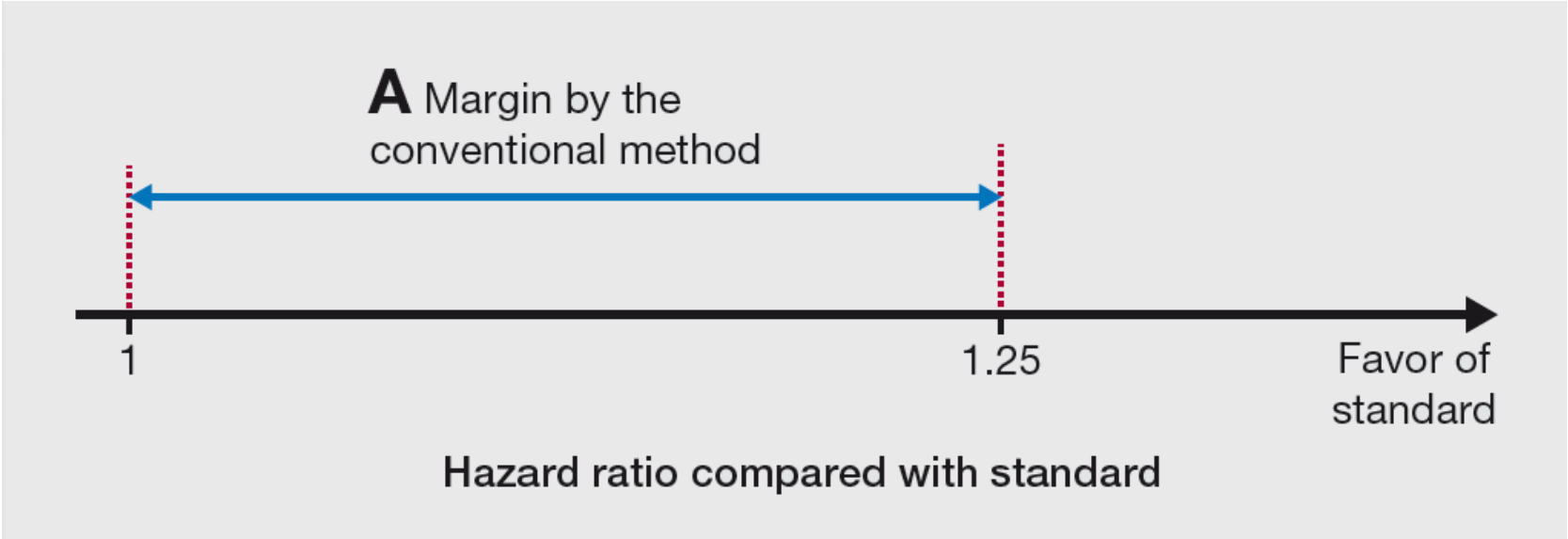


Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review

Shiro Tanaka[1], Yousuke Kinjo[2], Yoshiki Kataoka[2], Kenichi Yoshimura[1], and Satoshi Teramukai[1]

Clin Cancer Res; 18(7); 1837–47. ©2012 AACR.

Fixed Margin Method



Through the looking glass: understanding non-inferiority

Jennifer Schumi* and Janet T Wittes

Trials 2011, **12**:106

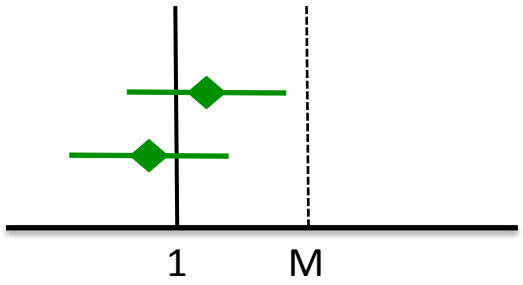
One focuses on the upper bound for this non-inferiority comparison; what happens at the lower end of the CI is not the primary concern.

The purpose of the trial is to estimate the upper bound of the CI, not to establish a point estimate of the treatment effect.

Bear in mind that the opposite of 'non-inferior' is not 'inferior'; it is 'not non-inferior'.

Interpretazione clinica di uno Studio di Non-Inferiorità

(dato uno specifico M di interesse)



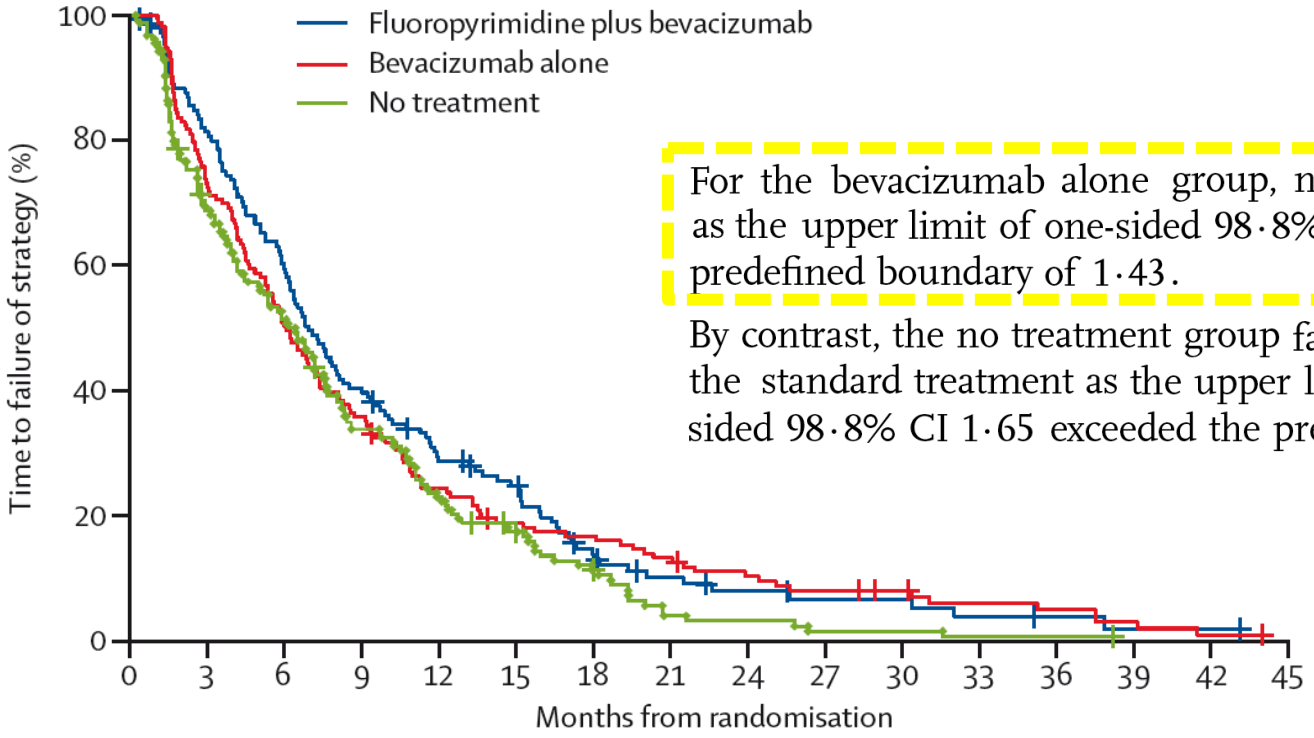
Dimostrazione di Non-Inferiorità

Il limite superiore dell'intervallo di confidenza non interseca la linea di non-effetto …indipendentemente da dove si colloca la stima puntuale dell'effetto

Maintenance strategies after first-line oxaliplatin plus fluoropyrimidine plus bevacizumab for patients with metastatic colorectal cancer (AIO 0207): a randomised, non-inferiority, open-label, phase 3 trial

Susanna Hegewisch-Becker, Ullrich Graeven*, Christian A Lerchenmüller, Birgitta Killing, Reinhard Depenbusch, Claus-Christoph Steffens, Salah-Eddin Al-Batran, Thoralf Lange, Georg Dietrich, Jan Stoehlmacher, Andrea Tannapfel, Anke Reinacher-Schick, Julia Quidde, Tanja Trarbach, Axel Hinke, Hans-Joachim Schmoll, Dirk Arnold*

Lancet Oncol **2015; 16: 1355–69**

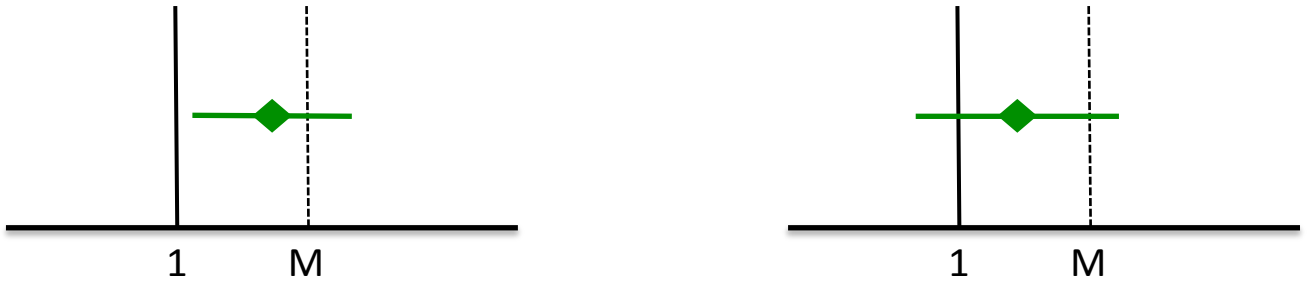


For the bevacizumab alone group, non-inferiority can be claimed as the upper limit of one-sided 98·8% CI of 1·42 did not exceed the predefined boundary of 1·43.

By contrast, the no treatment group failed to show non-inferiority to the standard treatment as the upper level of non-inferiority of one-sided 98·8% CI 1·65 exceeded the predefined boundary.

Interpretazione clinica di uno Studio di Non-Inferiorità

(dato uno specifico M di interesse)

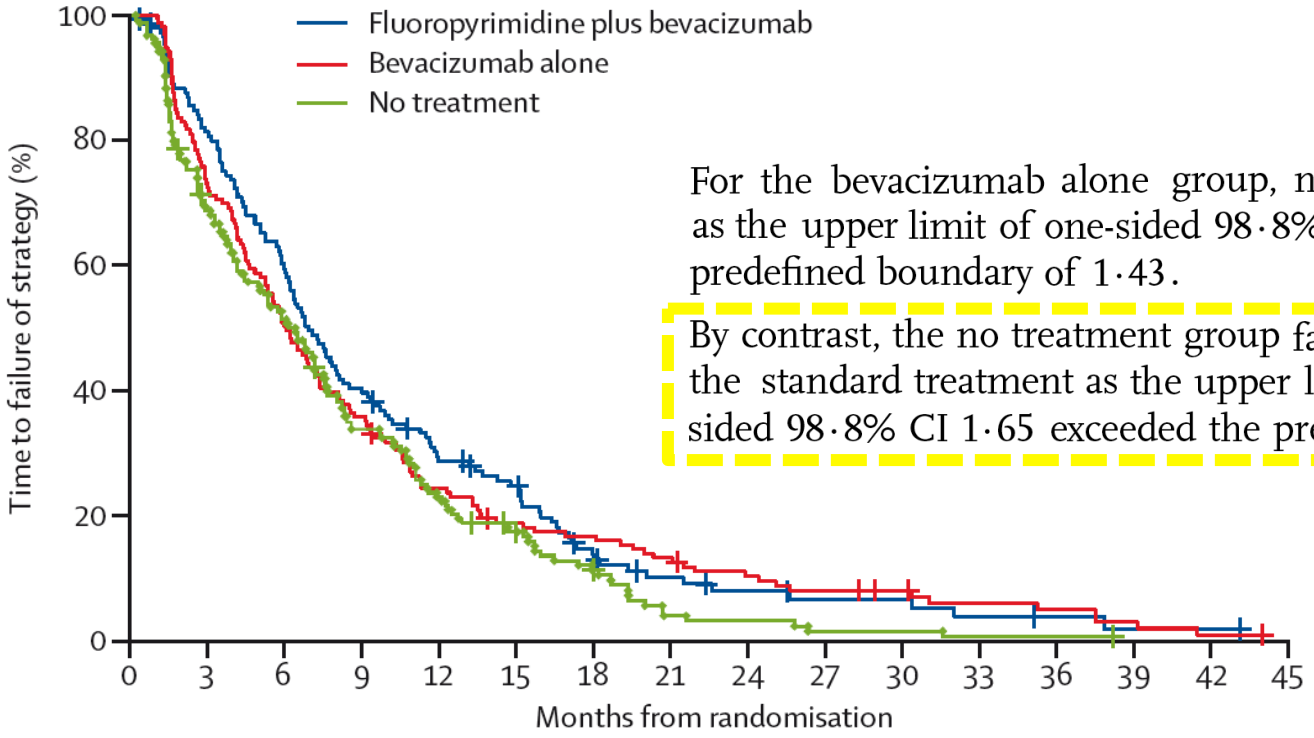


NON Dimostrazione di Non-Inferiorità

Maintenance strategies after first-line oxaliplatin plus fluoropyrimidine plus bevacizumab for patients with metastatic colorectal cancer (AIO 0207): a randomised, non-inferiority, open-label, phase 3 trial

Susanna Hegewisch-Becker, Ullrich Graeven*, Christian A Lerchenmüller, Birgitta Killing, Reinhard Depenbusch, Claus-Christoph Steffens, Salah-Eddin Al-Batran, Thoralf Lange, Georg Dietrich, Jan Stoehlmacher, Andrea Tannapfel, Anke Reinacher-Schick, Julia Quidde, Tanja Trarbach, Axel Hinke, Hans-Joachim Schmoll, Dirk Arnold*

Lancet Oncol **2015; 16: 1355–69**



For the bevacizumab alone group, non-inferiority can be claimed as the upper limit of one-sided 98·8% CI of 1·42 did not exceed the predefined boundary of 1·43.

By contrast, the no treatment group failed to show non-inferiority to the standard treatment as the upper level of non-inferiority of one-sided 98·8% CI 1·65 exceeded the predefined boundary.

Through the looking glass: understanding non-inferiority

Jennifer Schumi* and Janet T Wittes

Trials 2011, **12**:106

One focuses on the upper bound for this non-inferiority comparison; what happens at the lower end of the CI is not the primary concern.

The purpose of the trial is to estimate the upper bound of the CI, not to establish a point estimate of the treatment effect.

Bear in mind that the opposite of 'non-inferior' is not 'inferior'; it is 'not non-inferior'.

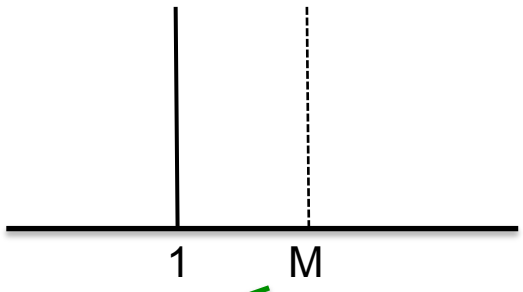
European Medicines Agency
Pre-authorisation Evaluation of Medicines for Human Use

London, 27 July 2005

MEA/CPMP/EWP/2158/99

N USE

MARGIN



1.15 ("Phare" Trastuzumab Study)

1.20 (Intergroup, intermittent ADT Study)

1.25 ("COMPARZ", Pazopanib Study)

re-specify a margin of non-

lence interval (or one-sided 97.5% interval) for the true difference between t will be constructed. This interval should lie entirely on the positive side of the non-inferiority .

The choice of delta must always be justified on both clinical and statistical grounds. It always needs to be tailored specifically to the particular clinical context and no rule can be provided that covers all clinical situations.

Highlights from the Tenth World Conference on Lung Cancer

TRACEY L. EVANS

The University of Pennsylvania Medical Center—Presbyterian, Division of Hematology/Oncology, Philadelphia, Pennsylvania, USA

The Oncologist 2004;9:232-238

	Pemetrexed (n = 283)	Docetaxel (n = 288)
Median survival	8.3 months	7.9 months
Hazard ratio* (95% CI)	0.99 (0.8-1.2)	
1-year survival rate	29.7%	29.7%

*This study was initially designed to be a noninferiority trial using the fixed margin method. Using this statistical test, pemetrexed would be declared no worse than 10% less efficacious than docetaxel if the upper limit of the 95% CI of the overall survival hazard ratio was ≤1.11. The upper limit of the hazard ratio observed in the study was 1.2. By another retrospectively applied test of noninferiority, the percent retention method, the comparator must demonstrate preservation of no less than 50% of the benefit of the standard arm. Pemetrexed in this study had a 102% retention of efficacy (95% CI = 52%-157%, p value = 0.047). Therefore, by this test, pemetrexed would be declared noninferior.

Highlights from the Tenth World Conference on Lung Cancer

TRACEY L. EVANS

The University of Pennsylvania Medical Center—Presbyterian, Division of Hematology/Oncology, Philadelphia, Pennsylvania, USA

The Oncologist 2004;9:232-238

	Pemetrexed (n = 283)	Docetaxel (n = 288)
Median survival	8.3 months	7.9 months
Hazard ratio* (95% CI)	0.99 (0.8-1.2)	
1-year survival rate	29.7%	29.7%

*This study was initially designed to be a noninferiority trial using the fixed margin method. Using this statistical test, pemetrexed would be declared no worse than 10% less efficacious than docetaxel if the upper limit of the 95% CI of the overall survival hazard ratio was ≤1.11. The upper limit of the hazard ratio observed in the study was 1.2. By another retrospectively applied test of noninferiority, the percent retention method, the comparator must demonstrate preservation of no less than 50% of the benefit of the standard arm. Pemetrexed in this study had a 102% retention of efficacy (95% CI = 52%-157%, p value = 0.047). Therefore, by this test, pemetrexed would be declared noninferior.

Statistical Review Addendum #1

Medical Division: Oncology Drug Products (HFD-150)
Biometrics Division: Division of Biometrics I (HFD-710)
NDA NUMBER: NDA 21-677 / N000
DRUG NAME: Alimta (Pemetrexed, LY231514)
INDICATION: Locally advanced or metastatic non-small cell lung cancer
SPONSOR: Eli Lilly and Company

3. Based on our internal discussion, we have determined that the lower limit of the 95% confidence interval is probably too conservative for estimating the control treatment (docetaxel) effect; thus, the resulting statistical test used in the statistical review dated 6/29/04 is also probably too conservative.

Design and analysis of non-inferiority mortality trials in oncology

Mark Rothmann[1,*,†], Ning Li[1], Gang Chen[1], George Y. H. Chi[1], Robert Temple[2] and Hsiao-Hui Tsou[1]

Statist. Med. 2003; **22**:239–264 (DOI: 10.1002/sim.1400)

Good Enough: A Primer on the Analysis and Interpretation of Noninferiority Trials

Sanjay Kaul, MD, and George A. Diamond, MD

Ann Intern Med. 2006;145:62-69

Active-control noninferiority trials are being performed with increasing frequency, especially in cardiovascular and oncologic applications when placebo-controlled trials are considered unethical.

Because of this uncertainty, the noninferiority margin is typically defined in terms of some fraction (f) of the standard treatment effect to be preserved (9, 11).

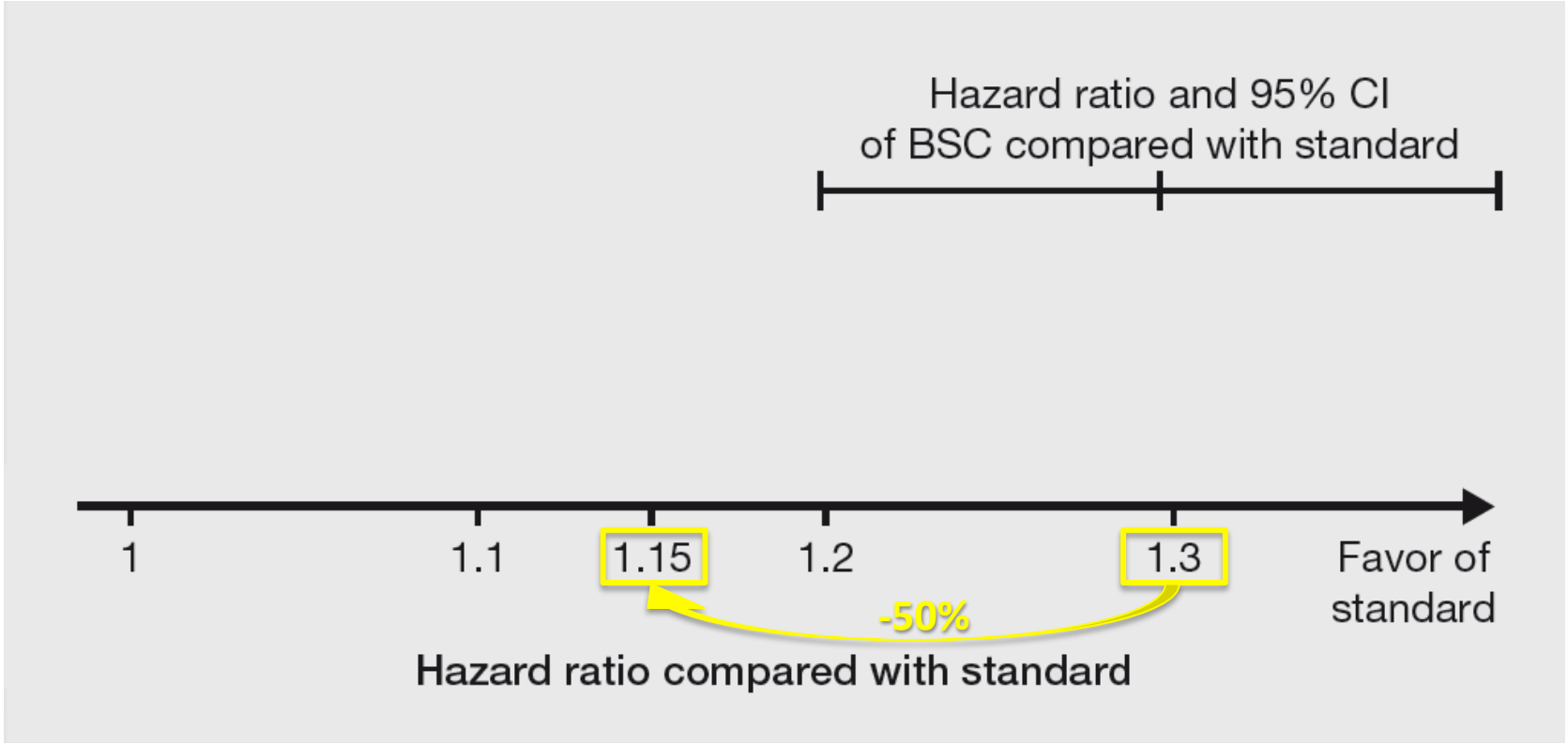
In the context of oncologic and thrombolytic trials, when mortality is evaluated, the U.S. Food and Drug Administration has suggested an f value of 0.5 (9, 11, 16).

Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review

Shiro Tanaka[1], Yousuke Kinjo[2], Yoshiki Kataoka[2], Kenichi Yoshimura[1], and Satoshi Teramukai[1]

Clin Cancer Res; 18(7); 1837–47. ©2012 AACR.

Percent Retention Method

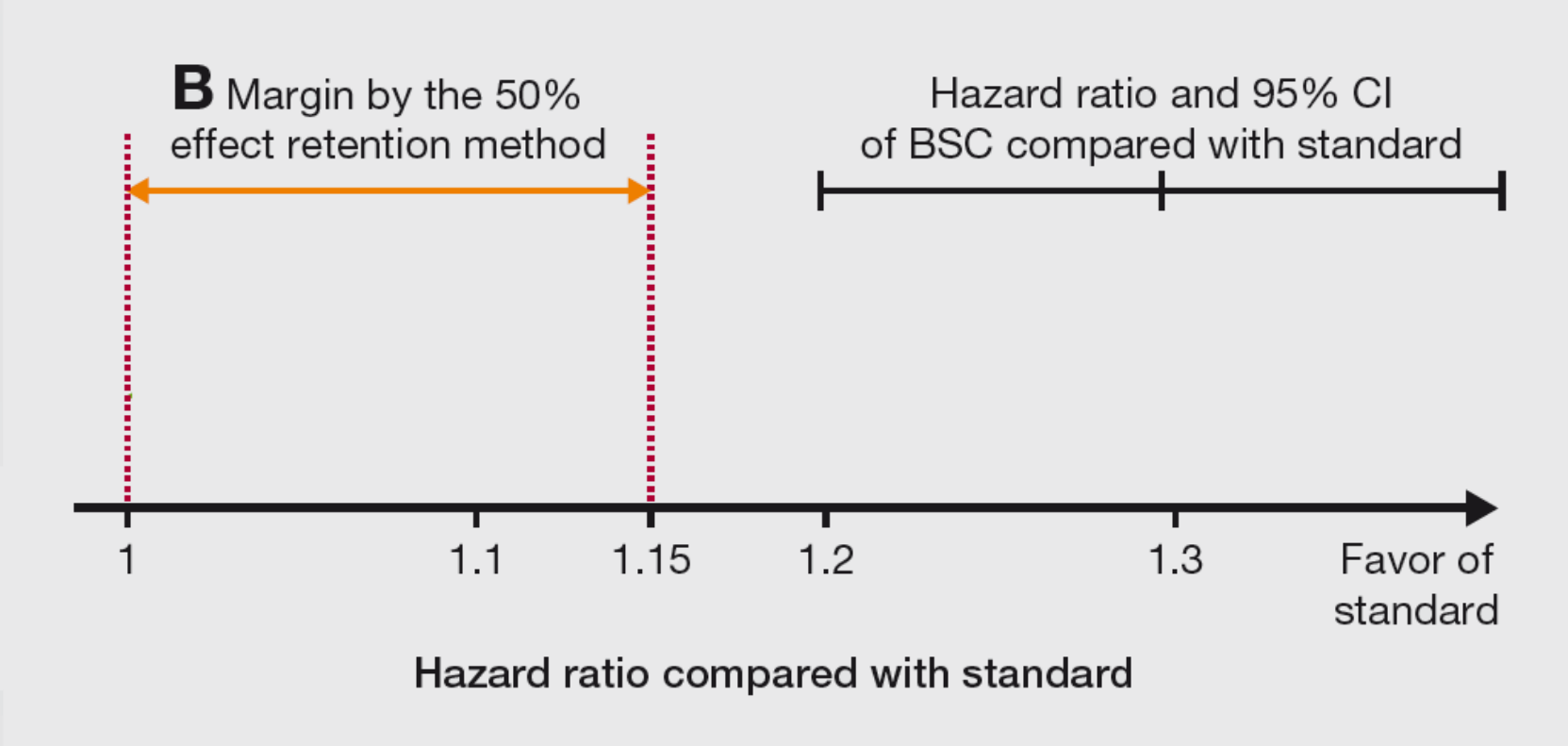


Statistical Issues and Recommendations for Noninferiority Trials in Oncology: A Systematic Review

Shiro Tanaka[1], Yousuke Kinjo[2], Yoshiki Kataoka[2], Kenichi Yoshimura[1], and Satoshi Teramukai[1]

Clin Cancer Res; 18(7); 1837–47. ©2012 AACR.

Percent Retention Method



Randomized Phase III Trial of Pemetrexed Versus Docetaxel in Patients With Non–Small-Cell Lung Cancer Previously Treated With Chemotherapy

Nasser Hanna, Frances A. Shepherd, Frank V. Fossella, Jose R. Pereira, Filippo De Marinis, Joachim von Pawel, Ulrich Gatzemeier, Thomas Chang Yao Tsao, Miklos Pless, Thomas Muller, Hong-Liang Lim, Christopher Desch, Klara Szondy, Radj Gervais, Shaharyar, Christian Manegold, Sofia Paul, Paolo Paoletti, Lawrence Einhorn, and Paul A. Bunn Jr.

J Clin Oncol 22:1589-1597. © 2004 by American Society of Clinical Oncology

the hypothesis that pemetrexed retained ≥ 50% of the survival benefit of docetaxel over BSC using data from the randomized comparative trial of docetaxel versus BSC by Shepherd et al[5] was prospectively planned (percent retention method).[14]

Apixaban versus Warfarin in Patients with Atrial Fibrillation

Christopher B. Granger, M.D., John H. Alexander, M.D., M.H.S., John J.V. McMurray, M.D., Renato D. Lopes, M.D., Ph.D., Elaine M. Hylek, M.D., M.P.H., Michael Hanna, M.D., Hussein R. Al-Khalidi, Ph.D., Jack Ansell, M.D., Dan Atar, M.D., Alvaro Avezum, M.D., Ph.D., M. Cecilia Bahit, M.D., Rafael Diaz, M.D., J. Donald Easton, M.D., Justin A. Ezekowitz, M.B., B.Ch., Greg Flaker, M.D., David Garcia, M.D., Margarida Geraldes, Ph.D., Bernard J. Gersh, M.D., Sergey Golitsyn, M.D., Ph.D., Shinya Goto, M.D., Antonio G. Hermosillo, M.D., Stefan H. Hohnloser, M.D., John Horowitz, M.D., Puneet Mohan, M.D., Ph.D., Petr Jansky, M.D., Basil S. Lewis, M.D., Jose Luis Lopez-Sendon, M.D., Prem Pais, M.D., Alexander Parkhomenko, M.D., Freek W.A. Verheugt, M.D., Ph.D., Jun Zhu, M.D., and Lars Wallentin, M.D., Ph.D., for the ARISTOTLE Committees and Investigators*

10.1056/NEJMoa1107039 NEJM.ORG

The primary noninferiority hypothesis required that apixaban preserve at least 50% of the relative reduction in the risk of stroke or systemic embolism associated with warfarin (62%) in six previous, major randomized, controlled trials.[10] This hypothesis provided a lower 95% confidence interval of 1.88 for the relative risk with placebo as compared with warfarin, and one half of this value was 1.44 (or 1.38 on a log scale).

Coordinate di riferimento nell'analisi di una S.C.C.

- **Statistica**: Tutte le analisi statistiche devono essere esplicitamente predeterminate (endpoint, trasformazione, test, tempi, sottogruppi)

- **Metodologica:** Tutti i pazienti randomizzati devono essere inseriti nell'analisi statistica in base al trattamento assegnato (**intention to treat**)

P. Bruzzi, 1989

Some essential considerations in the design and conduct of non-inferiority trials

Thomas R Fleming[a,b], *Katherine Odem-Davis*[a,b], *Mark D Rothmann*[c] *and Yuan Li Shen*[c]

Clinical Trials 2011; **8**: 432–439

Irregularities in quality of the conduct of the non-inferiority trial induce increased risk of both bias and variability.

While such irregularities are of concern in superiority trials, they are even more problematic in a non-inferiority trial since they often dilute the sensitivity to true differences between the experimental intervention and Standard regimens, leading to an increased risk of falsely declaring non-inferiority in settings where the test treatment truly is clinically inferior to Standard.

Studi di non-inferiorità:
Analisi ITT *Vs* analisi PP

The European Agency for the Evaluation of Medicinal Products
Evaluation of Medicines for Human Use
London, 27 July 2000 CPMP/EWP/482/99

IV.2.3 Choice of analysis set

In a superiority trial the full analysis set, based on the ITT (intention-to-treat) principle, is the analysis set of choice, with appropriate support provided by the PP (per protocol) analysis set. In a non-inferiority trial the full analysis set and the PP analysis set have equal importance and their use should lead to similar conclusions for a robust interpretation.

Guidance for Industry
Non-Inferiority Clinical Trials

March 2010
Clinical/Medical

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

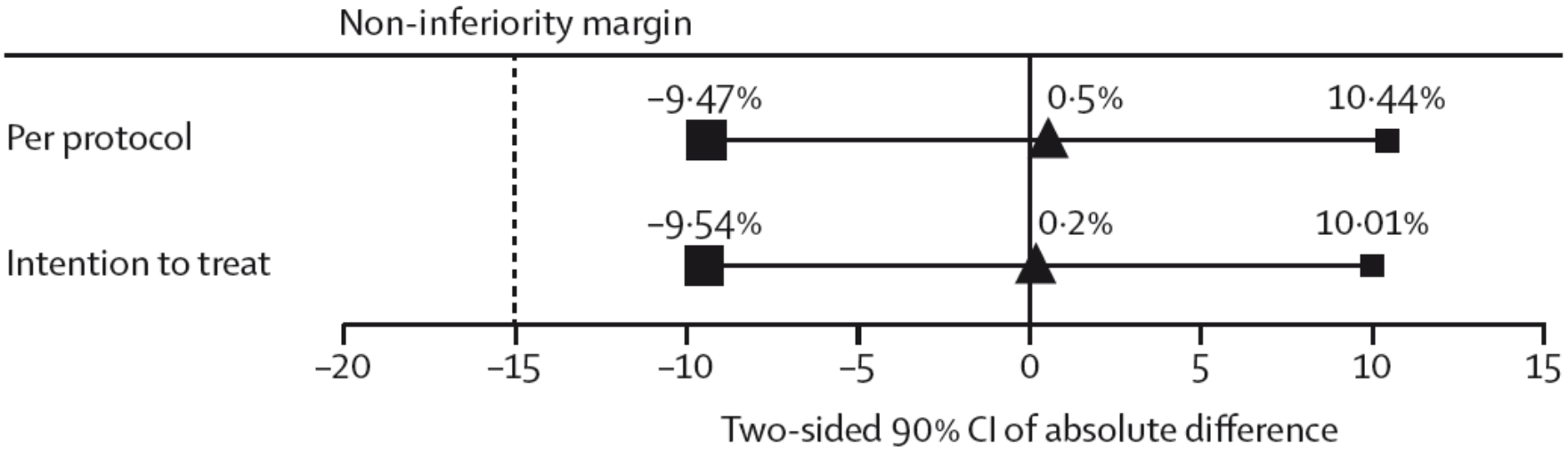
Although an "as-treated" analysis is therefore often suggested as the primary analysis for NI studies, there are also significant concerns with the possibility of informative censoring in an as-treated analysis. It is therefore important to conduct both ITT and as-treated analyses in NI studies.
Differences in results using the two analyses will need close examination.

Oral versus intravenous high-dose methylprednisolone for treatment of relapses in patients with multiple sclerosis (COPOUSEP): a randomised, controlled, double-blind, non-inferiority trial

Emmanuelle Le Page, David Veillard, David A Laplaud, Stéphanie Hamonic, Rasha Wardi, Christine Lebrun, Fabien Zagnoli, Sandrine Wiertlewski, Véronique Deburghgraeve, Marc Coustans, Gilles Edan, for the COPOUSEP investigators and the West Network for Excellence in Neuroscience*

Lancet 2015; 386: 974–81

The non-inferiority margin of 15% was set only for the primary endpoint. Oral methylprednisolone efficacy was to be judged non-inferior to intravenous methylprednisolone when the lower limit of the 90% CI (computed using Dunnett and Gent's[23,24] continuity corrected χ^2 for non-inferiority) of the absolute difference between the proportions of patients improved at day 28 was higher than $-\delta=-15\%$.

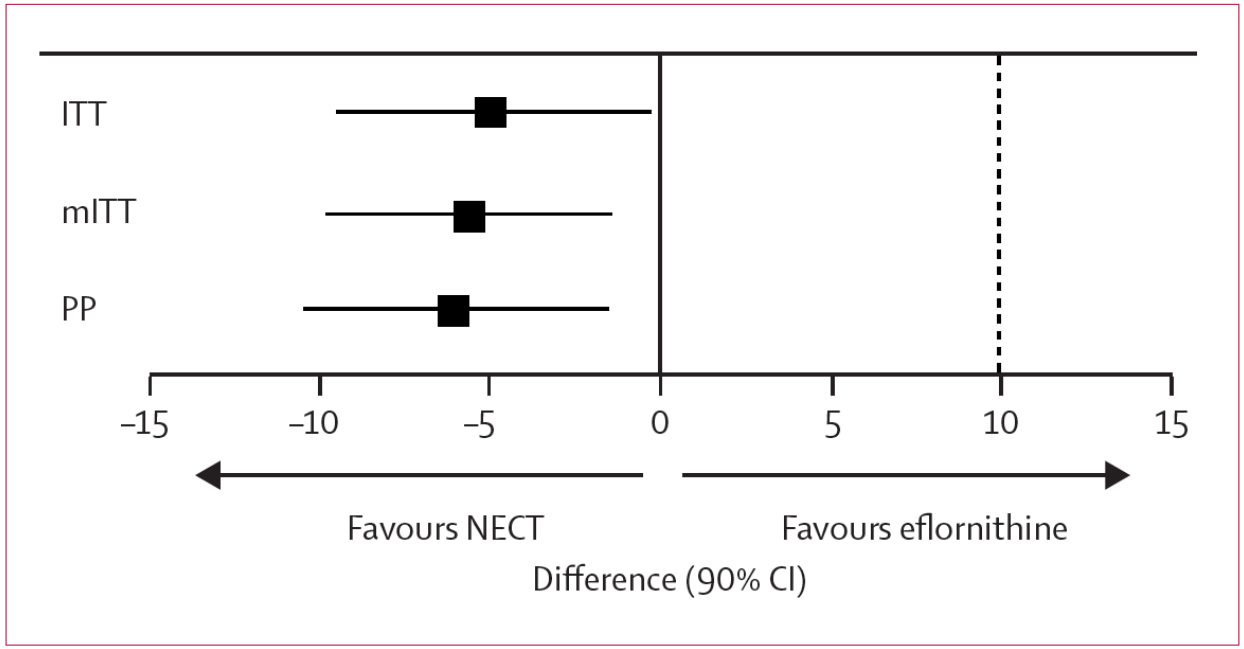


Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial

Gerardo Priotto, Serena Kasparian, Wilfried Mutombo, Daniel Ngouama, Sara Ghorashian, Ute Arnold, Salah Ghabri, Elisabeth Baudin, Vincent Buard, Serge Kazadi-Kyanza, Médard Ilunga, Willy Mutangala, Gabriele Pohlig, Caecilia Schmid, Unni Karunakara, Els Torreele, Victor Kande

Lancet 2009; 374: 56–64

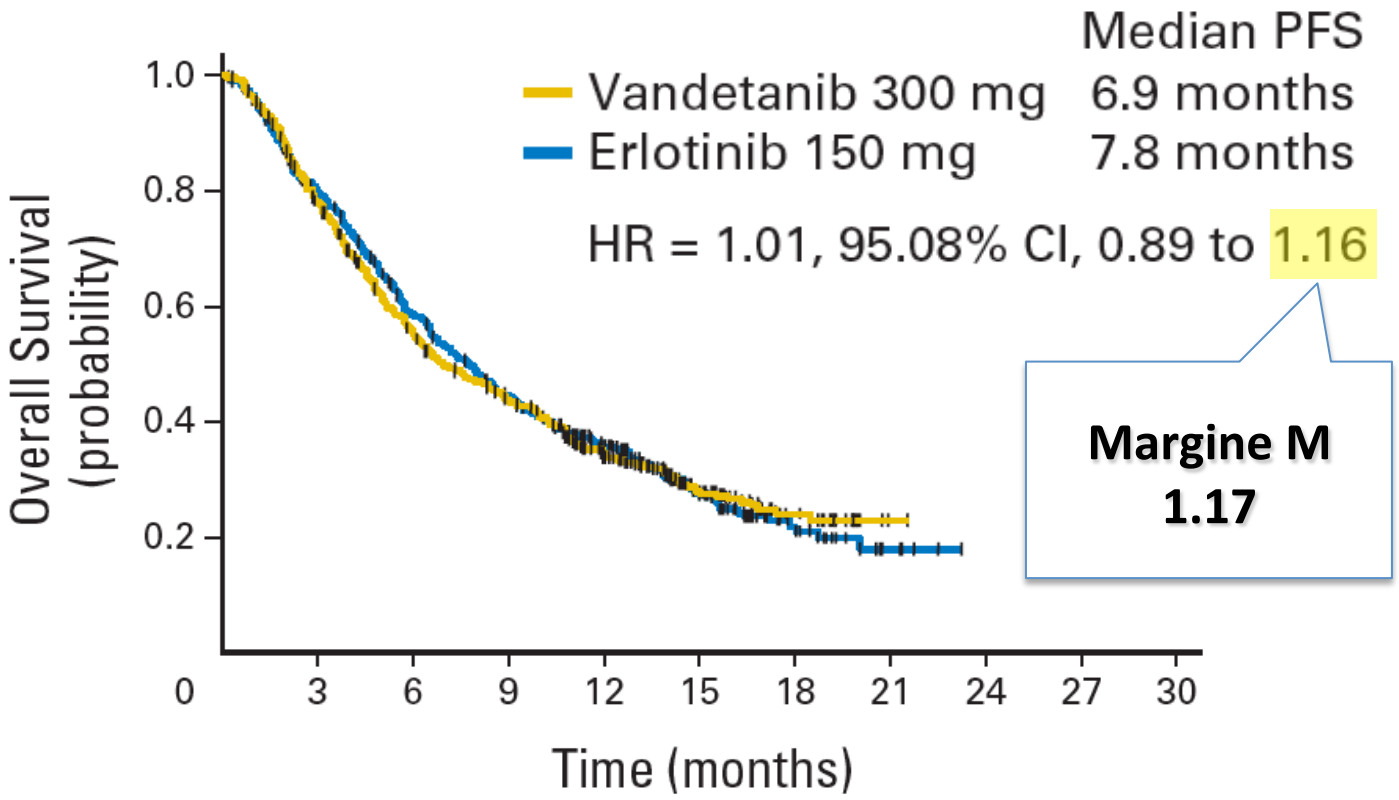
The non-inferiority margin for the difference in cure rates between groups was defined as 10% by the scientific committee, considering the high efficacy of the comparator and the crucial need of therapeutic alternatives.



Phase III Trial of Vandetanib Compared With Erlotinib in Patients With Previously Treated Advanced Non–Small-Cell Lung Cancer

Ronald B. Natale, Sumitra Thongprasert, F. Anthony Greco, Michael Thomas, Chun-Ming Tsai, Patrapim Sunpaweravong, David Ferry, Clive Mulatero, Robert Whorf, Joyce Thompson, Fabrice Barlesi, Peter Langmuir, Sven Gogov, Jacqui A. Rowbottom, and Glenwood D. Goss

J Clin Oncol 29:1059-1066. © 2011 by American Society of Clinical Oncology



Margine M 1.17

The overall incidence of grade ≥ 3 AEs was higher with vandetanib than erlotinib (50% *v* 40%, respectively)