III Sessione: Approcio metodologico alla ricerca

Varese
10-11 Marzo 2016

Come interpretare i risultati degli studi di non inferiorità

(l'esempio dei DOACs)

Francesco Dentali
Dipartimento di Medicina Clinica
Università dell' Insubria, Varese

Conflitti di Interesse

Letture
Protocolli di Ricerca
Advisory Boards

- Bayer
- BMS/Pfizer
- Boheringer
- Daiichi
- Sanofi
- Alfa Wasserman

Antefatto

PubMed

PubMed | Superiority vs Non-inferiority

Filters removed.

PubMed

PubMed comprises more than 25 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

PubMed COMMONS

Featured comment - Mar 8

Metformin, sulphonylurea & type 2 diabetes: A Margulis & colleagues discuss challenges in survival study. 1.usa.gov/1VV6cVh

Using PubMed

- PubMed Quick Start Guide
- Full Text Articles
- PubMed FAQs
- PubMed Tutorials
- New and Noteworthy

PubMed Tools

- PubMed Mobile
- Single Citation Matcher
- Batch Citation Matcher
- Clinical Queries
- Topic-Specific Queries

More Resources

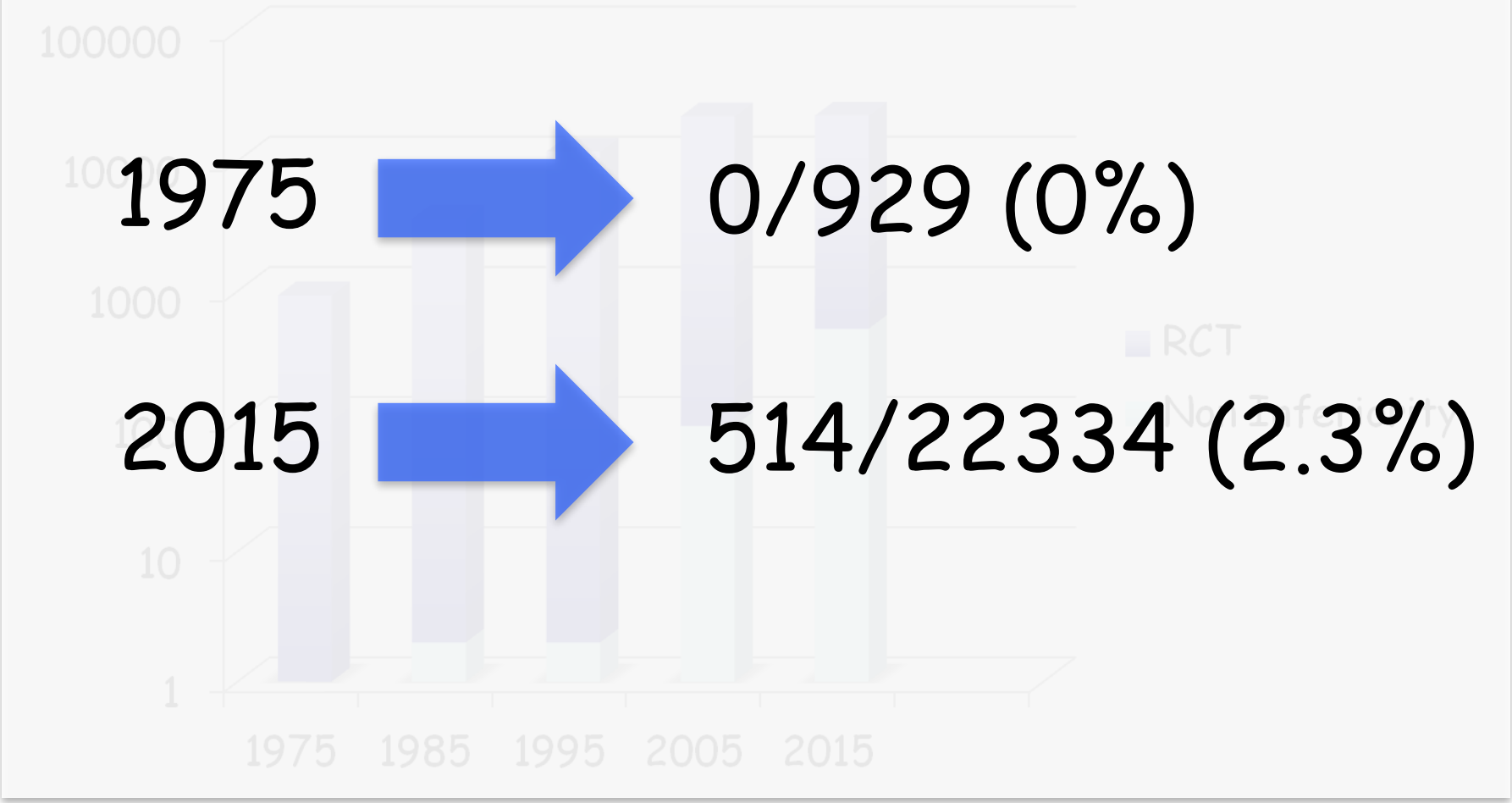
- MeSH Database
- Journals in NCBI Databases
- Clinical Trials
- E-Utilities (API)
- LinkOut

Superiorità vs non-inferiorità

Pubmed Research

1975 → 0/929 (0%)

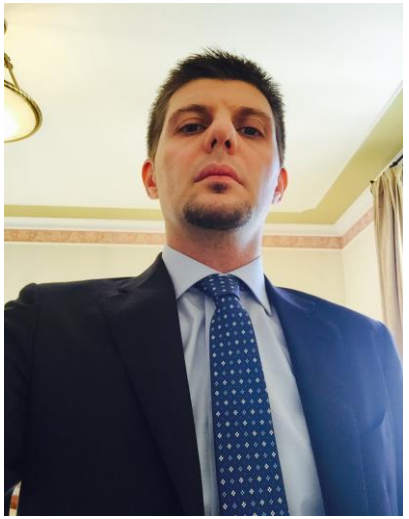
2015 → 514/22334 (2.3%)



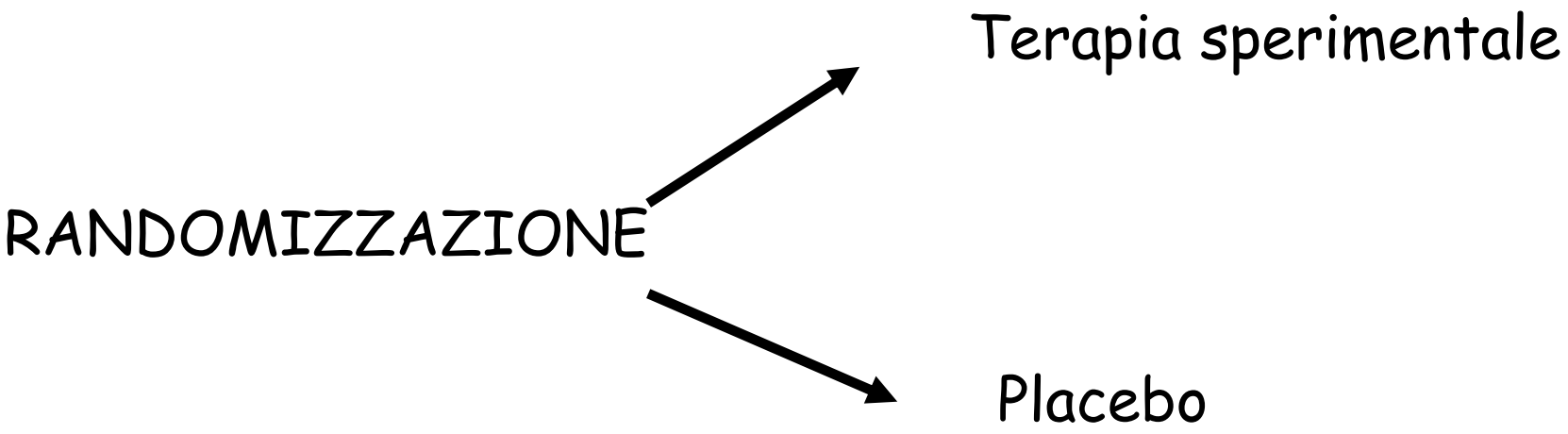
Premesse


LA
VEGA

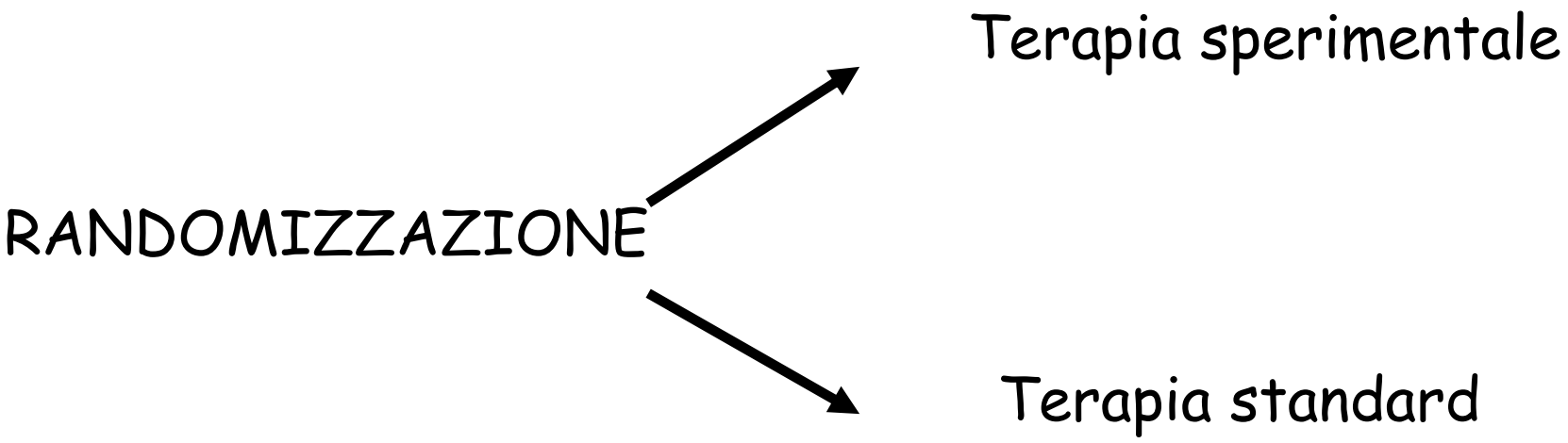

PERUGIATODAY



Non esiste un trattamento standard



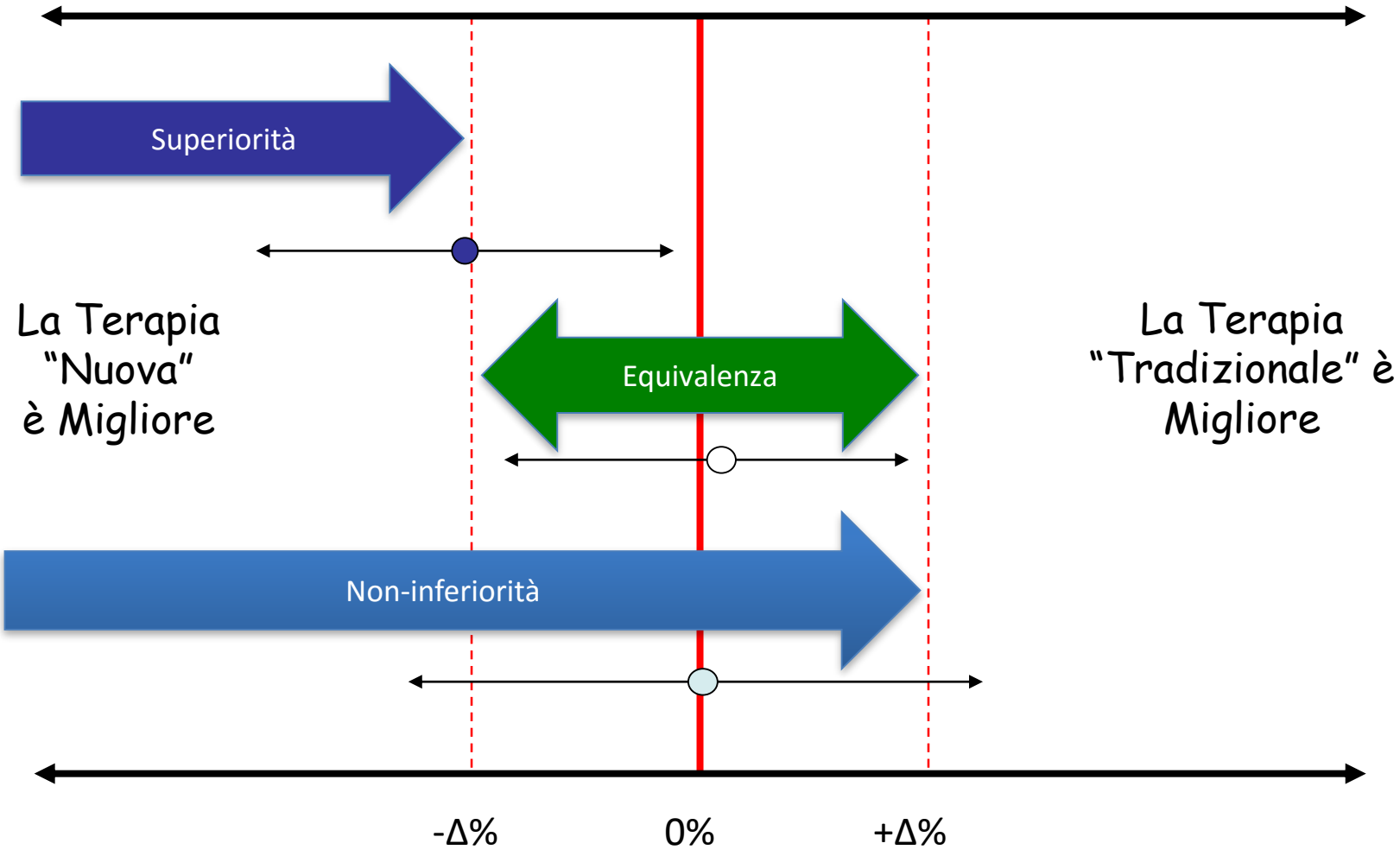
Esiste un trattamento standard



Studi Sperimentali

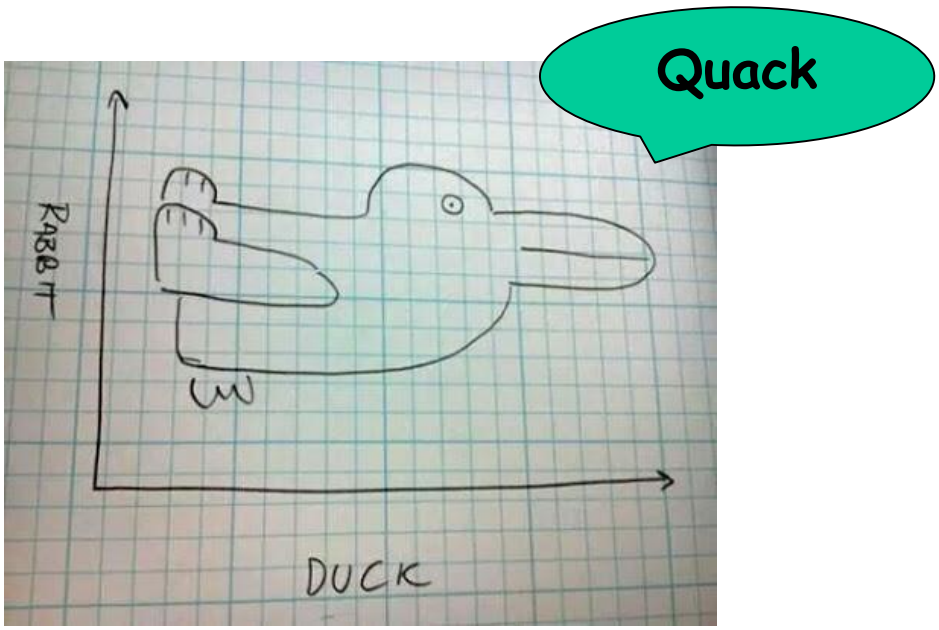
	Objective	Typical No. of Patients
Phase I	To explore possible toxic effects and determine tolerance of the intervention (and tolerated dose, if a drug study).	10 to 30
Phase II	To determine if treatment has a therapeutic effect or if there is any hope for benefits to outweigh the risks.	20 to 50
Phase III	To compare new treatment to the standard therapy or a control or placebo (if no standard therapy exists).	100 to 1000
Phase IV	To obtain long-term, large-scale information on morbidity and late effects (postmarketing study).	Hundreds or thousands

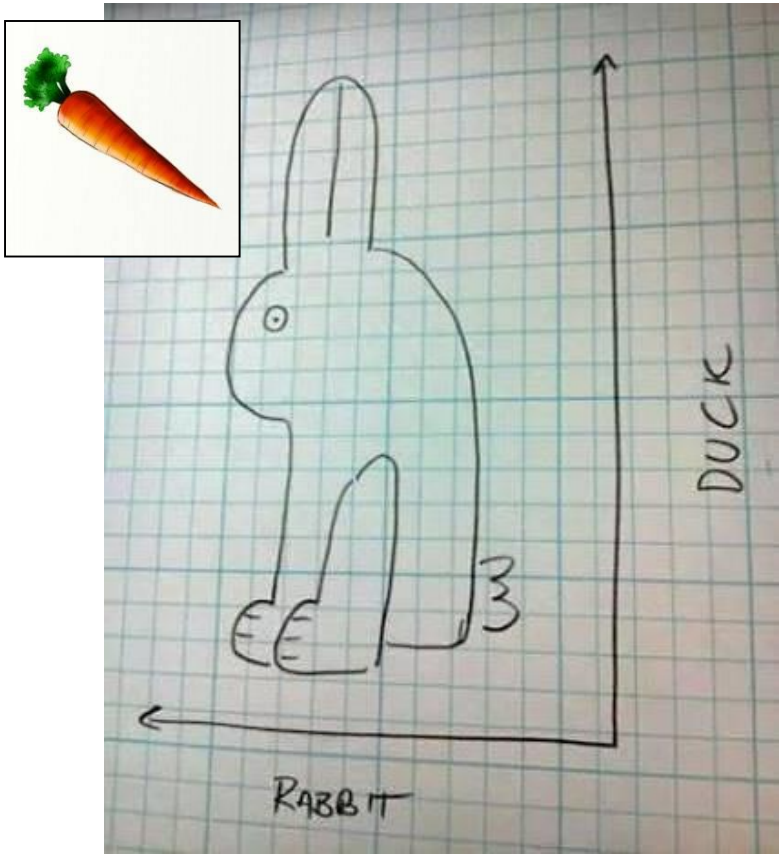
Studi Sperimentali



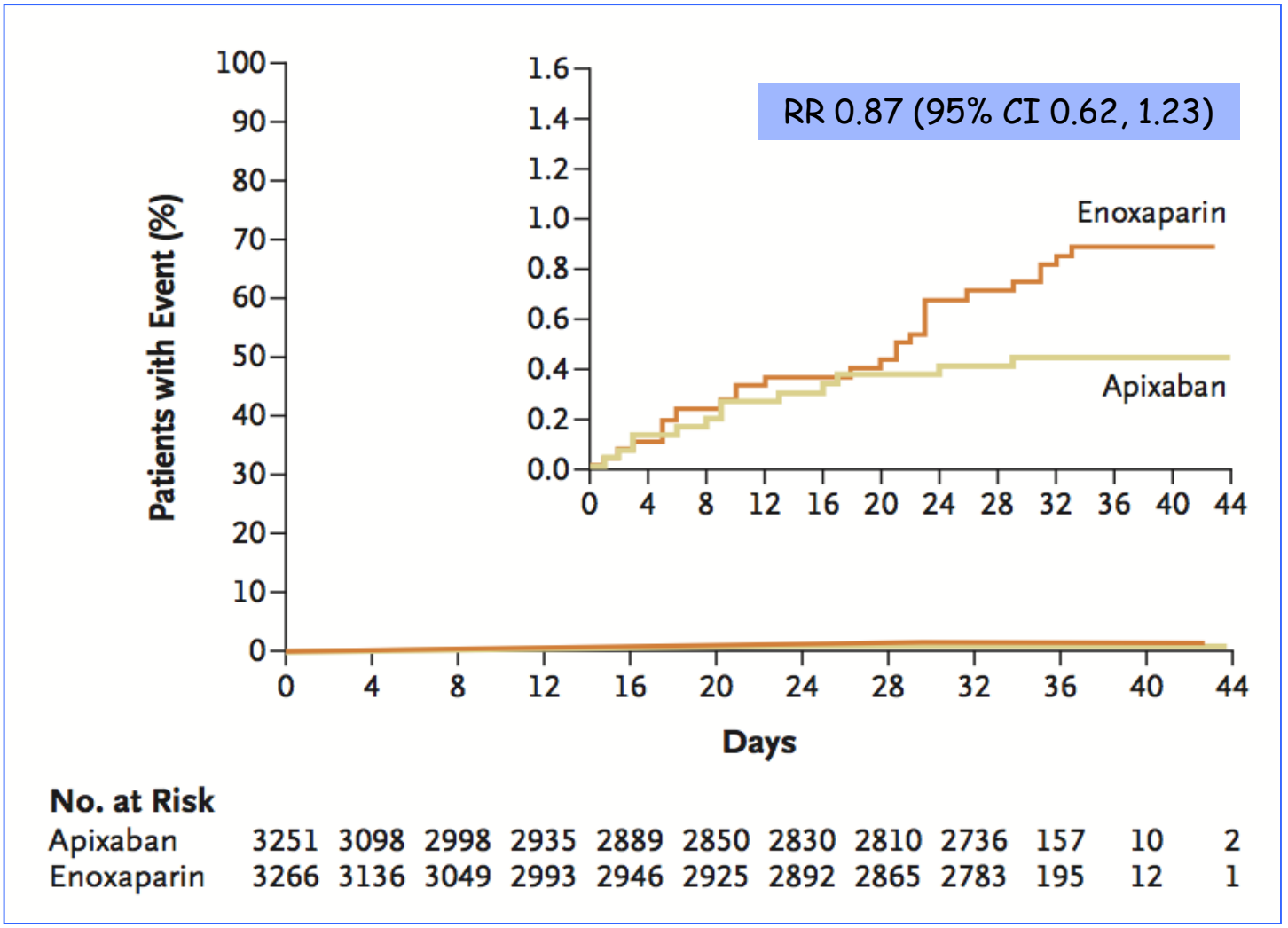
Differenza nel rischio di eventi

How do we change data interpretation?





Non superiority or Non Inferiority?

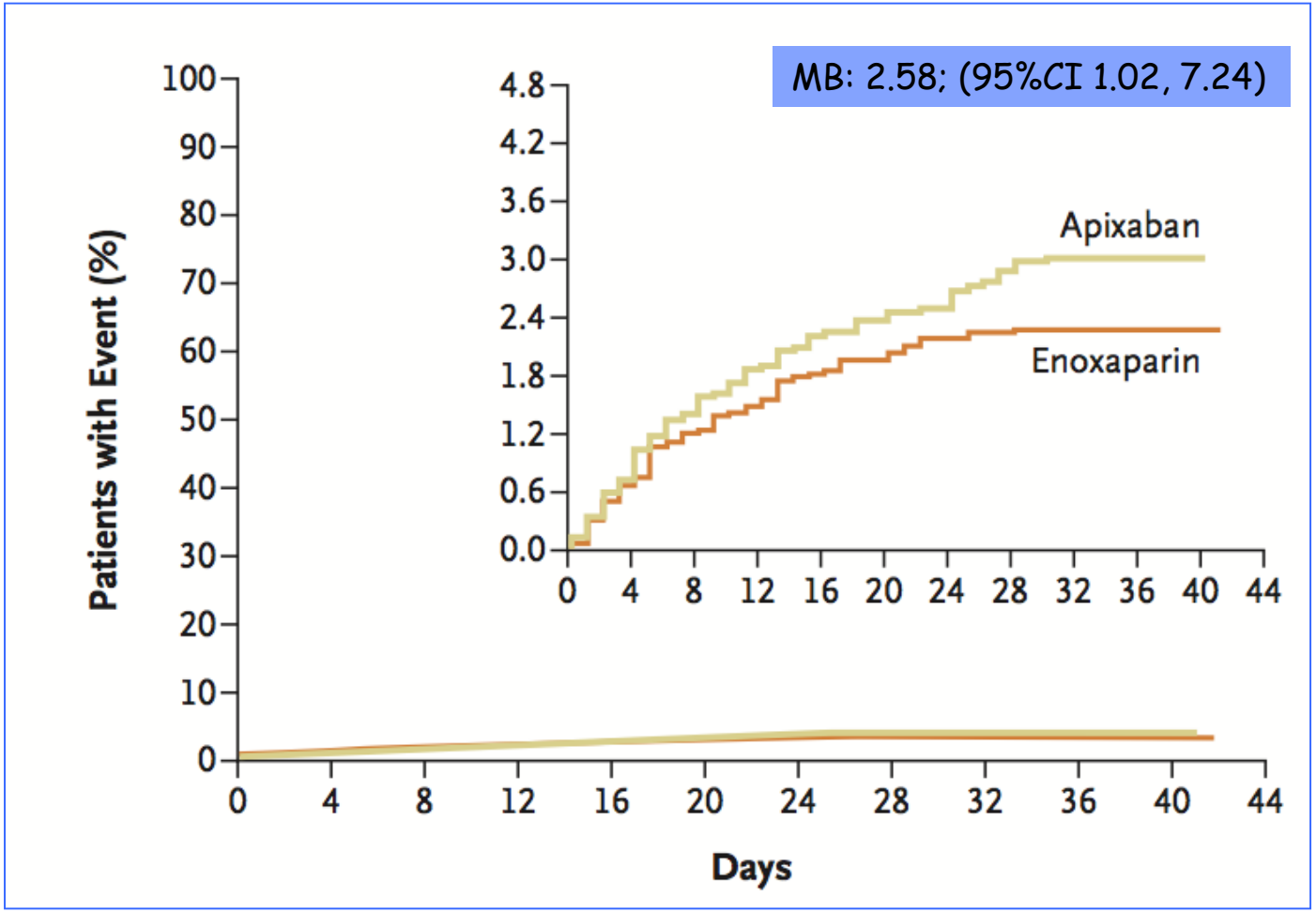


No. at Risk

Apixaban	3251	3098	2998	2935	2889	2850	2830	2810	2736	157	10	2
Enoxaparin	3266	3136	3049	2993	2946	2925	2892	2865	2783	195	12	1

Goldhaber et al. NEJM 2011

Non superiority or Non Inferiority?



Goldhaber et al. NEJM 2011

La posizione dell'AIFA

"Condurre uno studio di non inferiorità o di equivalenza potrebbe rivelarsi utile quando si voglia valutare se un trattamento sia **più sicuro** rispetto a quello di riferimento, possa offrire dei vantaggi in termini di **compliance** o di **costi**, o ancora quando si vogliano mettere a confronto diversi dosaggi, **formulazioni o vie di somministrazione** di uno stesso farmaco".

www.agenziafarmaco.gov.it/it/content/st

The ethics of non-inferiority trials

Silvio Garattini and Vittorio Bertele'[1] rightly caution about the traps to be avoided in doing non-inferiority trials, but go too far in suggesting that such trials should be uniformly banned "because they are unethical". A good example of the appropriate use of a non-inferiority trial is the current research programmes to develop new drugs for treatment of drug-sensitive tuberculosis.

Internationally recommended regimens are highly effective, curing 95% or more of patients in clinical trials in a wide variety of settings.[2,3] However, they require a minimum of three drugs which have significant side-effects and need to be given for at least 6 months. Improving on such high cure rates is almost impossible, but shortening treatment duration would improve completion rates and reduce both the time that patients are exposed to potentially toxic drugs and the cost of delivering tuberculosis chemotherapy in the developing world where resources are severely stretched—important medical and public-health goals.

It is impossible to prove that two treatment regimens have the same effect; there will always be some uncertainty surrounding estimates, and a small difference in effect size can never be excluded. However, the risks to patients in a properly done non-inferiority trial are no greater than those in a superiority trial. If non-inferiority designs were banned, there would be no prospect of shortening the duration of chemotherapy for patients with tuberculosis. And that would surely be unethical.

We are part of a consortium that is doing a non-inferiority trial.

*Andrew J Nunn, Sarah K Meredith, Melvin K Spigelman, Ann M Ginsberg, Stephen H Gillespie
ajn@ctu.mrc.ac.uk

MRC Clinical Trials Unit, London NW1 2DA, UK (AJN, SKM); Global Alliance for TB Drug Development, New York, NY, USA (MKS, AMG); and Centre for Medical Microbiology, Royal Free and University College Medical School, London, UK (SHG)

1 Garattini S, Bertele' V. Non-inferiority trials are unethical because they disregard patients' interests. *Lancet* 2007; 370: 1875–77.
2 Fox W, Ellard GA, Mitchison DA. Studies on the treatment of tuberculosis undertaken by the British Medical Research Council Tuberculosis Units, 1946–1986, with relevant subsequent publications. *Int J Tuberc Lung Dis* 1999; 3: S231–79.
3 Jindani A, Nunn AJ, Enarson DA. Two 8-month regimens of chemotherapy for treatment of newly diagnosed pulmonary tuberculosis: international multicentre randomised trial. *Lancet* 2004; 364: 1244–51.

Silvio Garattini and Vittorio Bertele' (Dec 1, p 1875)[1] argue that non-inferiority trials "have no ethical justification, since they do not offer any possible advantage...to pa[illegible]nts". Their conclusion is based [illegible] ethic of physicia[illegible] for each pa[illegible] [illegible]tions.[2] However, this ethic only applies to a world where resources for health care are endless and hence do not matter. Yet in a world of limited resources, this ethic can lead to unfair practices, with some patients getting full access to services and others getting none.[2]

Under resource scarcity, non-inferiority trials can therefore be ethical: if the tested treatment is cheaper, savings can be used to treat patients with other diseases who would otherwise be denied treatment. Thus, non-inferiority trials can help to increase population health. Even providing patients with slightly inferior interventions can be ethically justified if savings are substantial and help to treat other patients for a larger benefit.

Finally, it is true that the definition of the inferiority margin is arbitrary, but this certainly also applies to the significance level of α=0·05 in superiority trials.

I declare that I have no conflict of interest.

Afschin Gandjour
afschin.gandjour@uk-koeln.de

Institute for Health Economics and Clinical Epidemiology, University of Cologne, 50935 Cologne, Germany

1 Garattini S, Bertele' V. Non-inferiority trials are unethical because they disregard patients' interests. *Lancet* 2007; 370: 1875–77.
2 Pearson SD. Caring and cost: the challenge for physician advocacy. *Ann Intern Med* 2000; 133: 148–53.

In their Viewpoint, Silvio Garattini and Vittorio Bertele'[1] call on the scientific community to ban non-inferiority trials because they are unethical. If efficacy were the only advantage patients might get from clinical trials, I would agree. However, there is no doubt that increasing drug adherence by providing a combination pill, for example, or increasing avai[illegible] and affordability by [illegible] drugs [illegible]. Garattini and Bertele' take examples [illegible]dies that used non-inferiority designs to argue for their call to ban them. I believe that they confuse non-inferiority as a design with investigators' errors in using it. Setting wide inferiority limits or using statistical rather than clinical difference as a basis for concluding that a drug is non-inferior are investigators' decisions, which are not necessarily right. The appropriate action would be to ban the improper application of the non-inferiority design, not the design itself.

I declare that I have no conflict of interest.

Elsayed Z Soliman
esoliman@wfubmc.edu

Epidemiological Cardiology Research Center (EPICARE), Department of Epidemiology and Prevention, Wake Forest University School of Medicine, Winston Salem, NC 27157, USA

1 Garattini S, Bertele' V. Non-inferiority trials are unethical because they disregard patients' interests. *Lancet* 2007; 370: 1875–77.

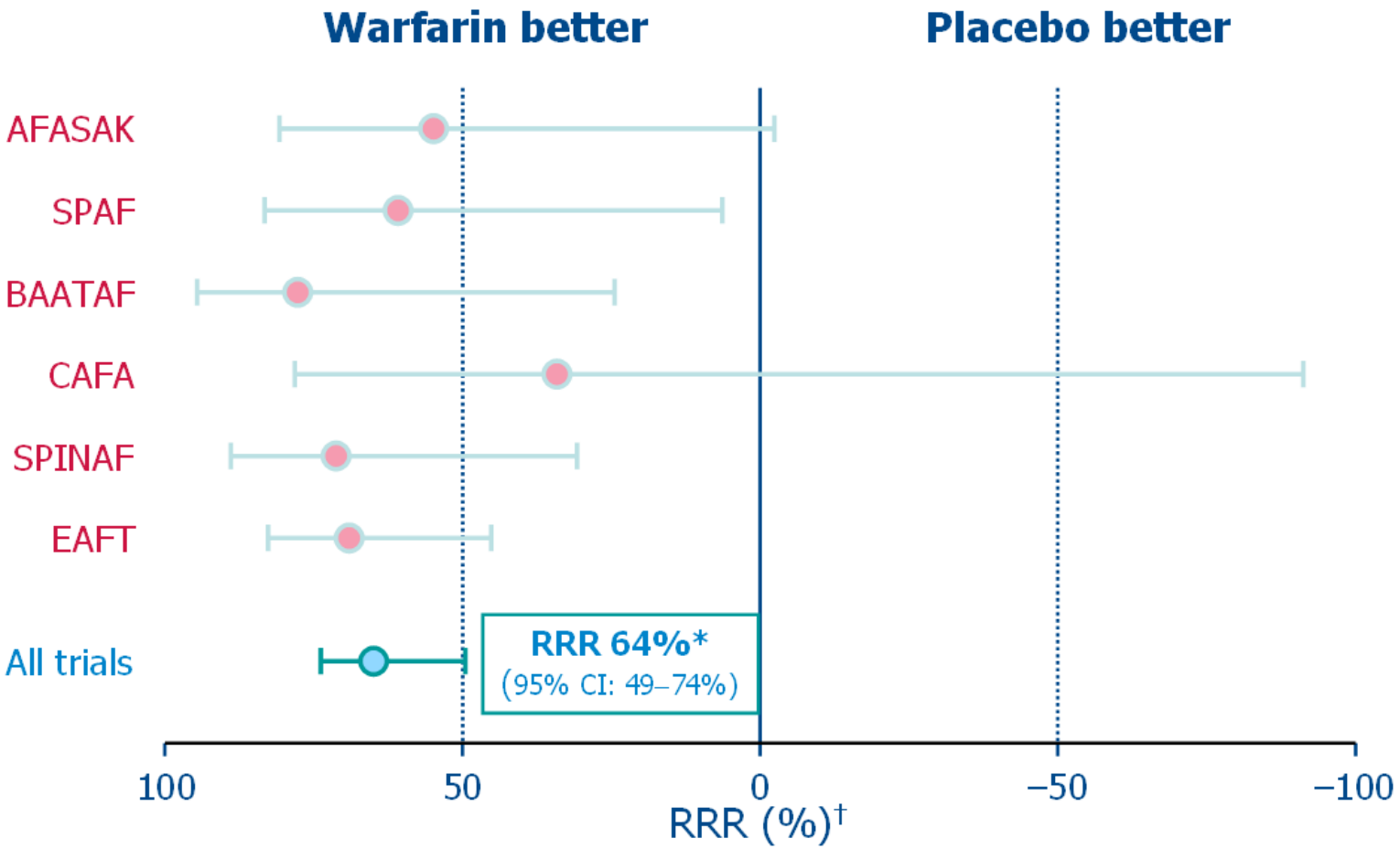
Silvio Garattini and Vittorio Bertele'[1] assert that "non-inferiority trials are unethical because they disregard patients' interest". This ignores the demonstrable and continuing value of non-inferiority trials. For example, since the introduction of cytotoxic drugs against solid tumours and

The printed journal includes an image merely for illustration

Submissions should be made via our electronic submission system at http://ees.elsevier.com/thelancet/

Dimostrare che un trattamento antitubercolare di durata < 6 mesi è "non inferiore" ad un trattamento di durata 6 mesi permetterebbe di:
1) Evitare molti effetti collaterali
2) Ridurre i costi, utilizzando le (scarse) risorse per curare altri pazienti.

Warfarin in Atrial Fibrillation



Hart et al. Ann Int Med 2007

Warfarin in VTE

WARFARIN SODIUM VERSUS LOW-DOSE HEPARIN IN THE LONG-TERM TREATMENT OF VENOUS THROMBOSIS

RUSSELL HULL, M.B., B.S., TERRY DELMORE, M.Sc., EDWARD GENTON, M.D., JACK HIRSH, M.D., MICHAEL GENT, M.Sc., DAVID SACKETT, M.D., M.Sc., DERMOT MCLOUGHLIN, M.B., B.S., AND PETER ARMSTRONG, M.D.

Abstract Acute deep-vein thrombosis is usually treated with intravenous heparin for a number of days, then with oral anticoagulants for weeks to months. We have compared adjusted-dose warfarin sodium with fixed low-dose subcutaneous heparin in the prevention of recurrent deep-vein thrombosis. Sixty-eight patients with acute deep-vein thrombosis confirmed by venography were treated with intravenous heparin and then randomized to secondary prophylaxis. Nine of 35 patients receiving subcutaneous heparin, but none of 33 receiving warfarin sodium, had new episodes of objectively documented venous thromboembolism (P = 0.001). Seven patients on warfarin sodium experienced bleeding complications (of which four were major), as compared with no patients receiving subcutaneous heparin (P<0.005). Thus, adjusted-dose warfarin sodium is more effective than low-dose subcutaneous heparin in preventing recurrent venous thromboembolism, but its use is accompanied by a significant risk of bleeding. (N Engl J Med 301:855-858, 1979)

9/35 (25.7%) vs 0/33 (0%)

anticoagulant therapy to prevent delayed recurrence of venous thromboembolism. The evidence that oral anticoagulants are effective for this purpose is limited to one retrospective study[1] in which it was reported that the frequency of recurrence was lower in patients treated with oral anticoagulants than in a group receiving no treatment. However, since both the initial diagnosis of venous thrombosis and the diagnosis of recurrent thrombosis were based on clinical criteria and the patients were not randomized, only limited conclusions can be drawn from this study.

It would be desirable to obtain more definitive information about the value of oral anticoagulant drugs in the prevention of recurrent venous thromboembolism, since their use is associated with bleeding complications in 5 to 10 per cent of patients.[2] In addition, an alternative form of secondary prevention with a lower risk of bleeding would be attractive if it was effective in preventing recurrent venous thromboembolism. Low-dose subcutaneous heparin has proved to be effective in the primary prevention of venous thromboembolism in a number of well defined high-risk groups without inducing major bleeding,[3] and this approach was therefore compared with oral anticoagulant therapy for the secondary prevention of venous thromboembolism. Our study sought to establish the relative effectiveness and safety of these treatments in patients with acute venous thrombosis.

From the departments of Pathology, Medicine and Clinical Epidemiology and Biostatistics, McMaster University Medical Centre and Chedoke Hospital, Hamilton, Ontario (address reprint requests to Dr. Hull at the Department of Medicine, Room 3V39, McMaster University Medical Centre, 1200 Main St. W., Hamilton, Ontario L8S 4J9, Canada).

This work was supported by grants from the Province of Ontario and from the Ontario and Canadian Heart Foundations.

strated acute deep-vein thrombosis. Venography was performed because deep-vein thrombosis was suspected or because ^{125}I-fibrinogen leg scanning or impedance plethysmography (IPG) was positive in patients who were being screened postoperatively. Venography was performed by the method of Rabinov and Paulin,[4] using diagnostic criteria previously described.[5]

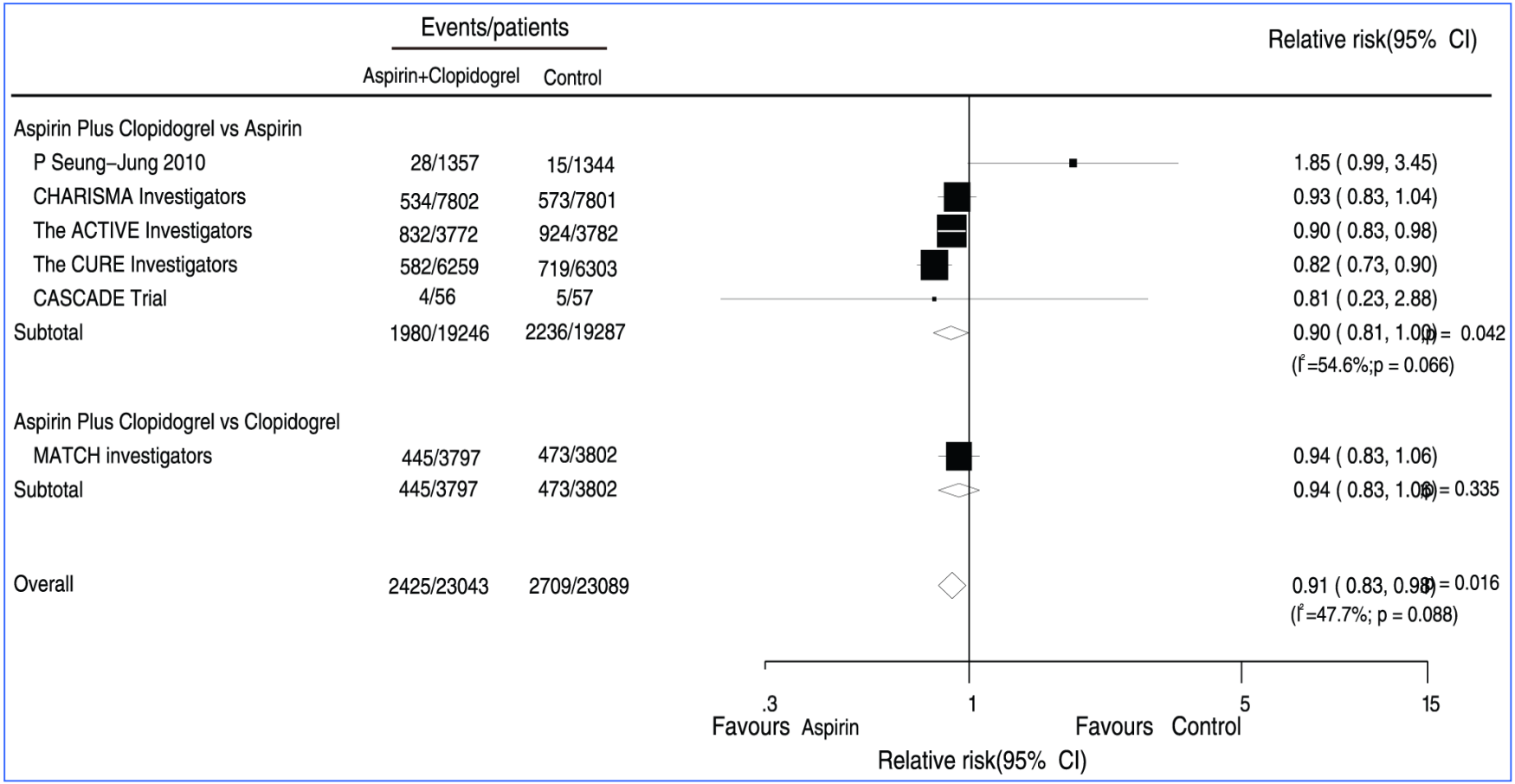
Patients were excluded if they were pregnant, had active peptic-ulcer disease, had a history of allergy to the dye used in venography or could not be followed as outpatients because of geographic inaccessibility; all exclusions were documented. Patients were stratified according to the site of thrombosis (proximal-vein thrombosis or calf only) and according to history of previous venous thromboembolism.

Regimens

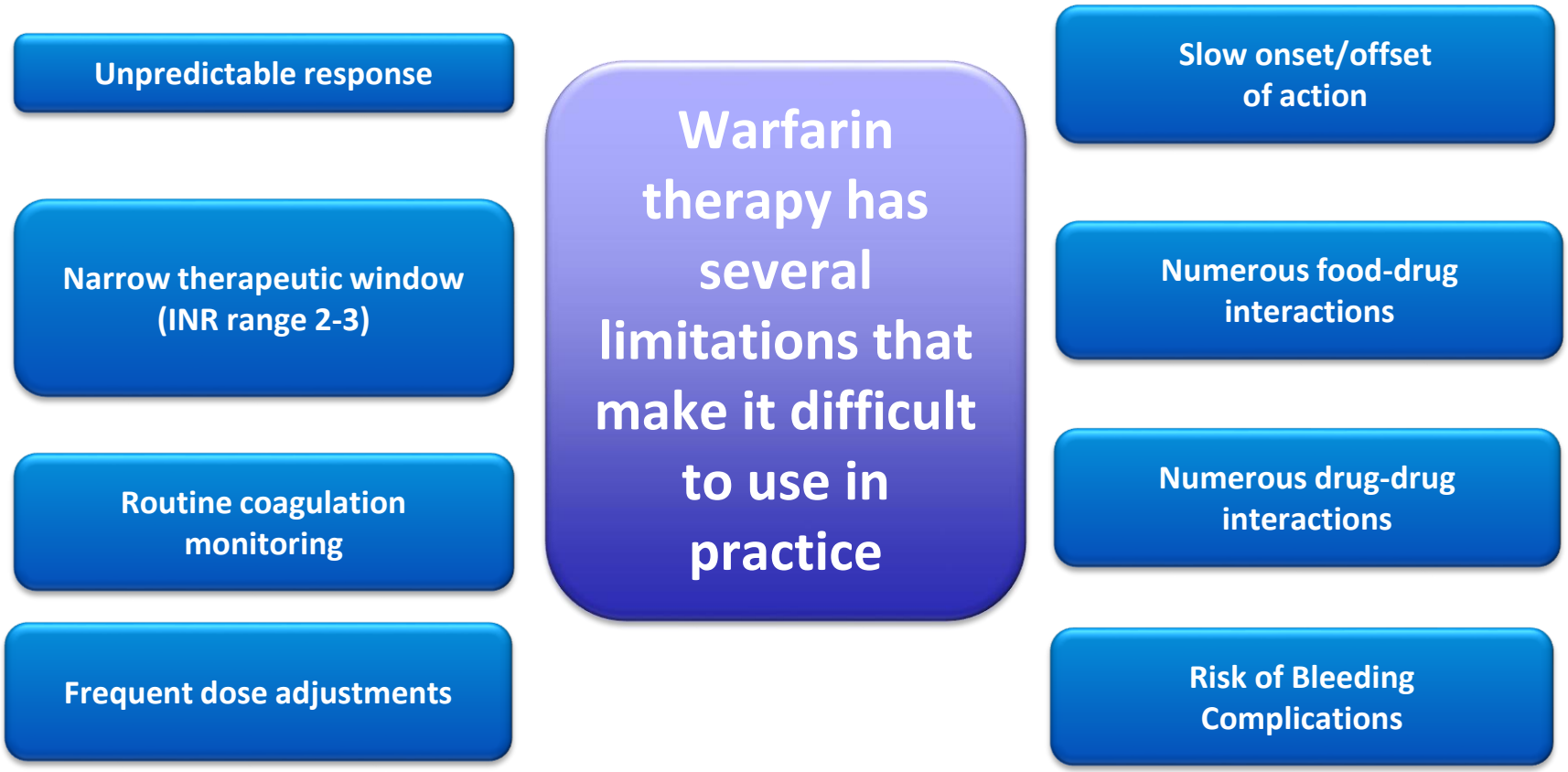
All patients were treated for 14 days with continuous intravenous heparin that was adjusted to maintain the activated partial thromboplastin time at 1½ to two times the normal control. Informed consent was obtained, and the patients were allocated at random on Day 10 (by use of a system of sealed envelopes) to receive either warfarin sodium or low-dose heparin. Warfarin sodium treatment began four days before stopping heparin in an initial dose of 10 mg per day. To avoid intervention by study physicians who might keep the warfarin sodium patients under closer supervision, anticoagulant control was monitored weekly with Simplastin (General Diagnostics) by each patient's family physician, who agreed to adjust the warfarin sodium dose to maintain the prothrombin time at 1½ to two times the control value. Fixed, low-dose subcutaneous heparin was begun on Day 14 at a dose of 5000 units every 12 hours. The initial injections were carried out by the nursing staff and then by the patients under direct supervision. Upon discharge from hospital, patients were given a three- to six-week supply of heparin, which was subsequently administered either by the patients themselves or by visiting nurses in the occasional patient who was unable to measure accurately the heparin dose.

Treatment with either warfarin sodium or low-dose heparin continued for 12 weeks in patients who had proximal-vein thrombosis and for six weeks in patients who had calf-vein thrombosis. To avoid contamination, aspirin-containing drugs, sulfinpyrazone and dipyridamole were prohibited during the period of trial.

Combined therapy and major CV events



Zhou et al. Plos One 2012



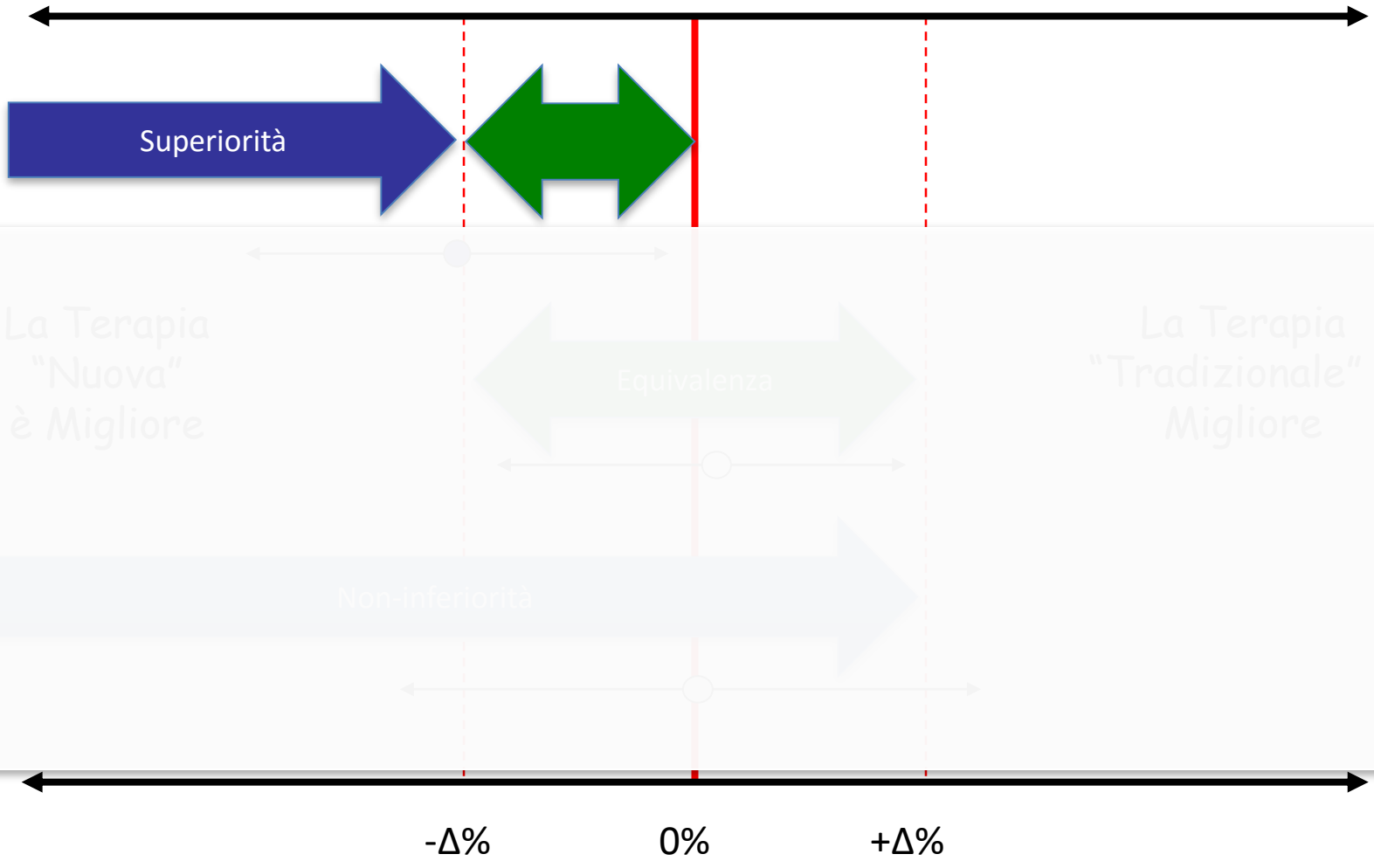
- Warfarin was #1 in 2003 and 2004 in the number of mentions of "deaths for drugs causing adverse effects in therapeutic use"
- Warfarin caused 6% of the 702,000 ADEs treated in the ED/year; 17% required hospitalization

J Thromb Thrombolysis 2008

Studi Sperimentali di superiorità
Calcolo del Sample Size

- Least relevant difference or the clinical significance
- Type 1 error (false positive) -> Statistical significance
- Type 2 error (false positive) -> Power

Studi Sperimentali



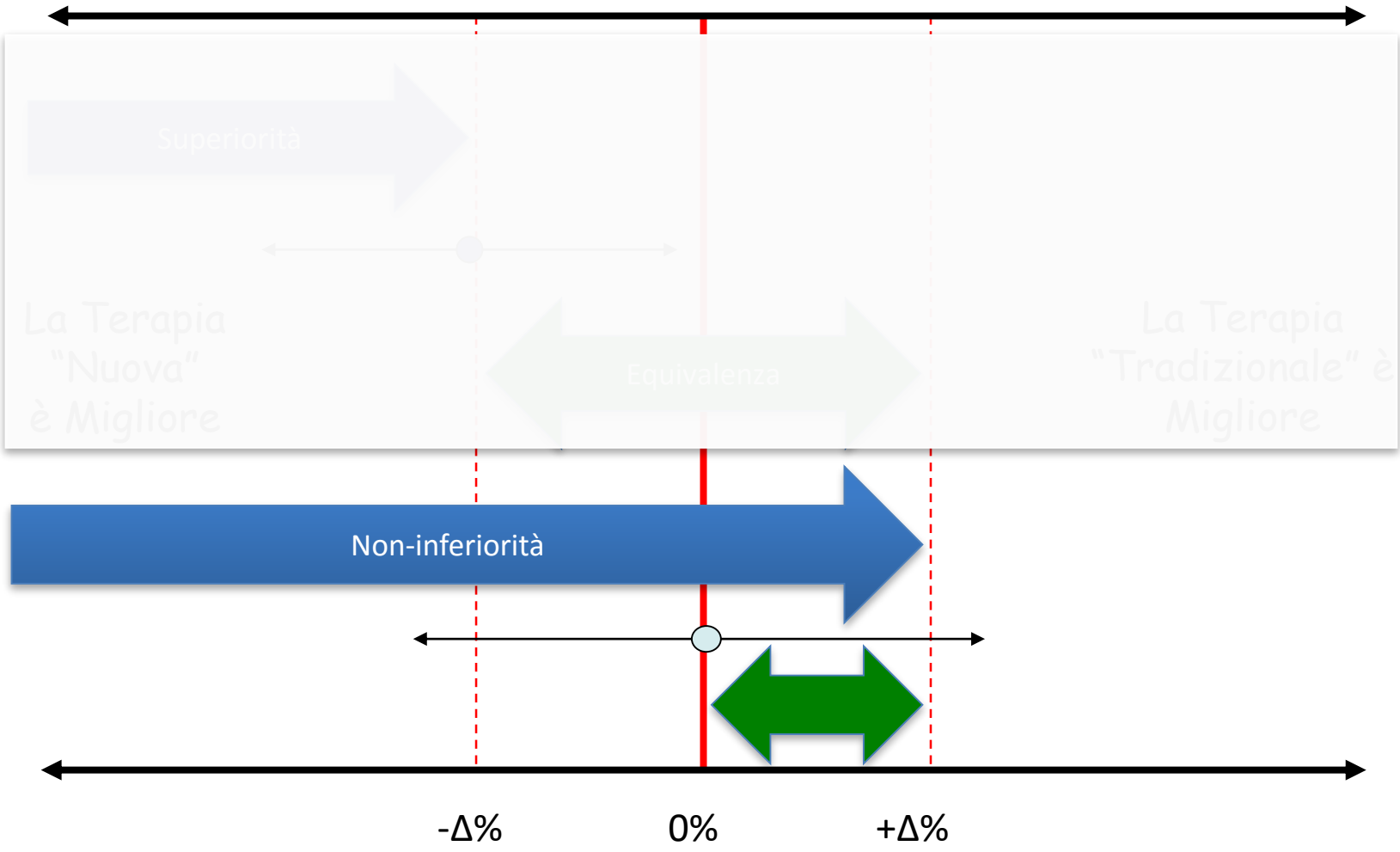
Differenza nel rischio di eventi

Studi Sperimentali di non inferiorità

Calcolo del Sample Size

- Maximum relevant difference
- Type 1 error (false positive) -> Statistical significance
- Type 2 error (false positive) -> Power

Studi Sperimentali



Differenza nel rischio di eventi

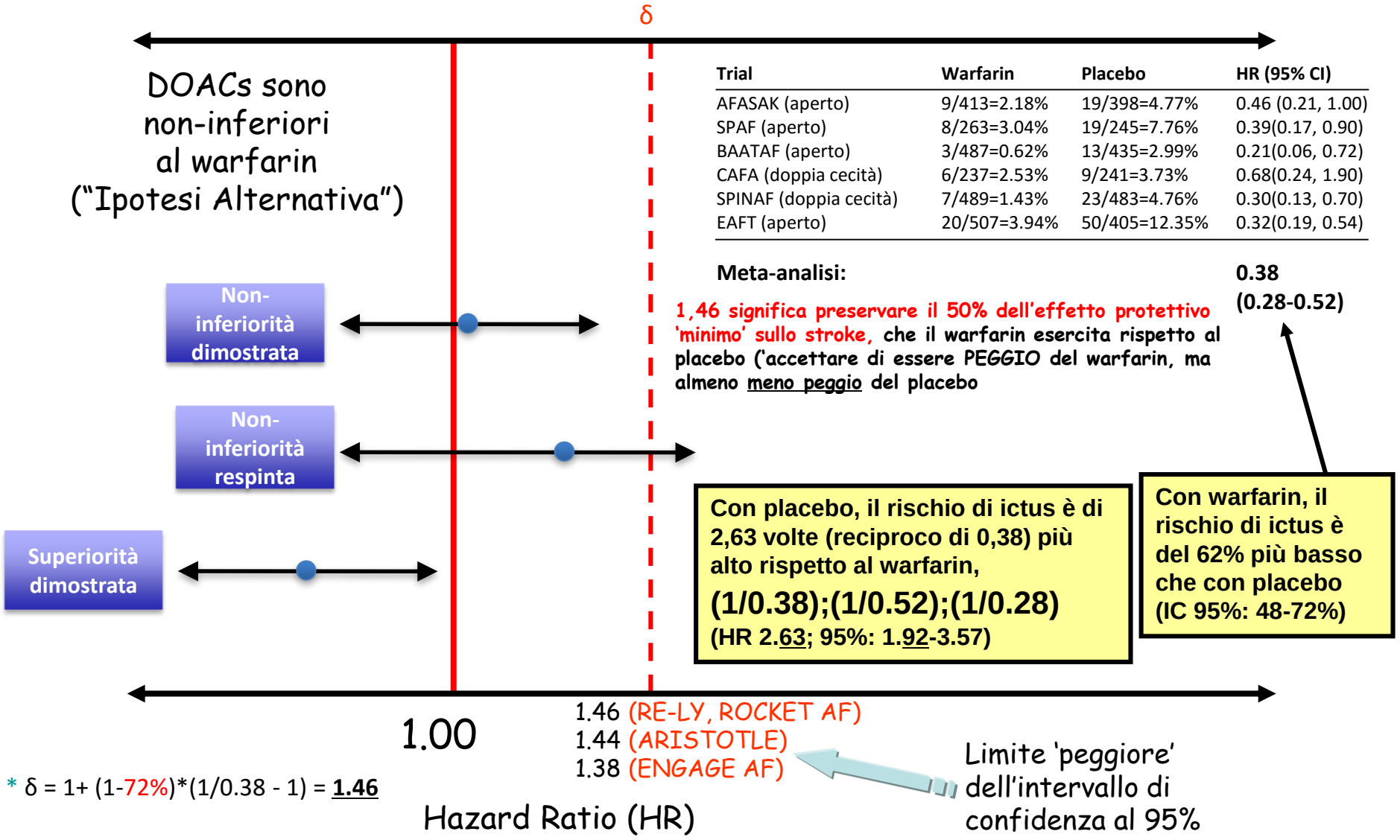
Come si fa a calcolare questo benedetto margine +Δ% ??

Margine di non inferiorità

- Clinicamente rilevante ("utile")
- Pre-specificato
- Scelto in riferimento all' efficacia (o effetto) del trattamento attivo (di riferimento) in studi precedenti controllati da placebo
- Il margine può essere espresso come media, rapporto...

EMEA-CPMP (2001) Points to consider on switching between superiority and non-inferiority. British Journal of Clinical Pharmacology. 52(3):223, 2001

Scelta del limite peggiore dell'intervallo di confidenza per testare la non inferiorità



Trial	Warfarin	Placebo	HR (95% CI)
AFASAK (aperto)	9/413=2.18%	19/398=4.77%	0.46 (0.21, 1.00)
SPAF (aperto)	8/263=3.04%	19/245=7.76%	0.39(0.17, 0.90)
BAATAF (aperto)	3/487=0.62%	13/435=2.99%	0.21(0.06, 0.72)
CAFA (doppia cecità)	6/237=2.53%	9/241=3.73%	0.68(0.24, 1.90)
SPINAF (doppia cecità)	7/489=1.43%	23/483=4.76%	0.30(0.13, 0.70)
EAFT (aperto)	20/507=3.94%	50/405=12.35%	0.32(0.19, 0.54)
Meta-analisi:			**0.38 (0.28-0.52)**

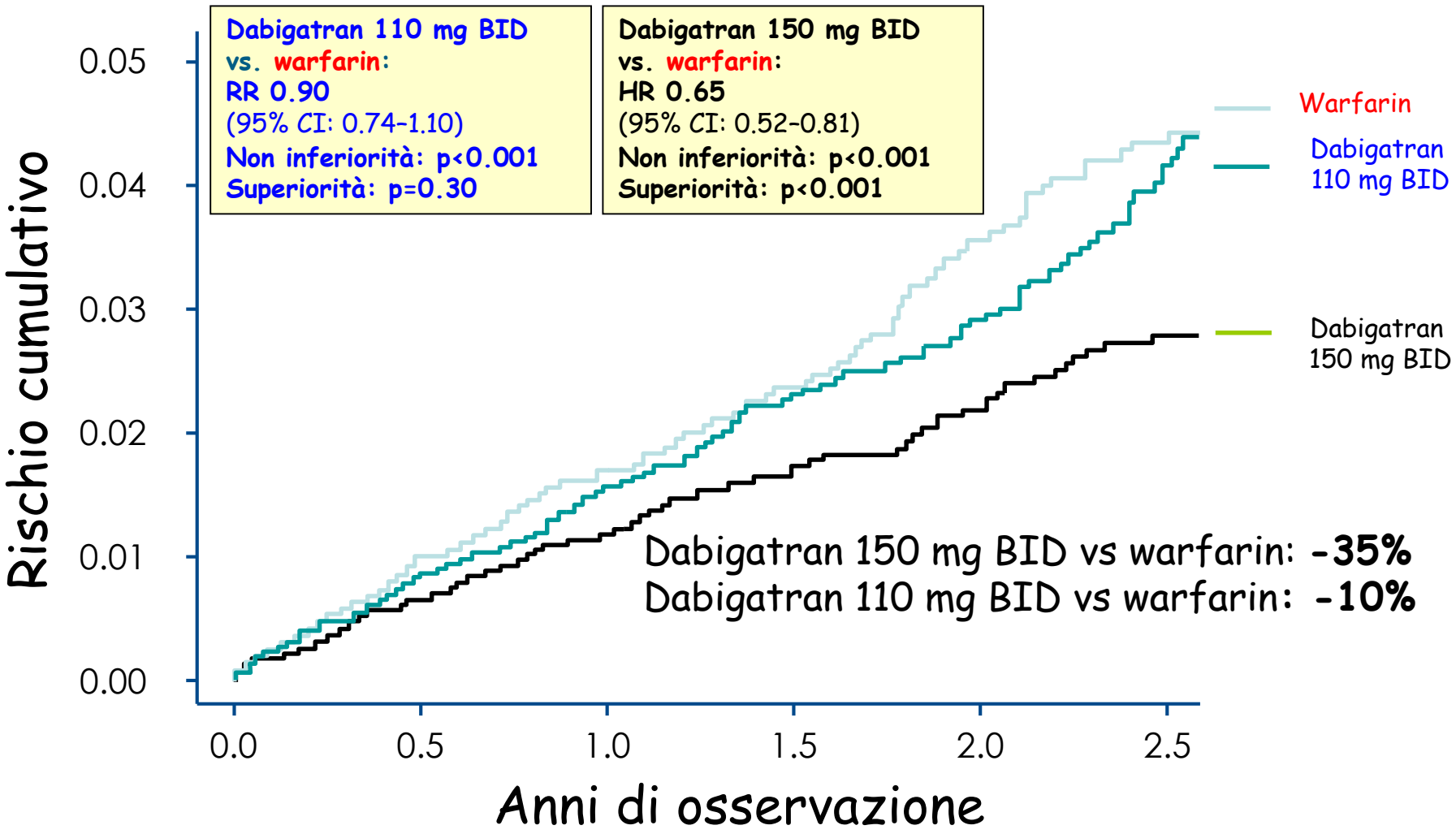
1,46 significa preservare il 50% dell'effetto protettivo 'minimo' sullo stroke, che il warfarin esercita rispetto al placebo ('accettare di essere PEGGIO del warfarin, ma almeno meno peggio del placebo

Con placebo, il rischio di ictus è di 2,63 volte (reciproco di 0,38) più alto rispetto al warfarin,
(1/0.38);(1/0.52);(1/0.28)
(HR 2.63; 95%: 1.92-3.57)

Con warfarin, il rischio di ictus è del 62% più basso che con placebo (IC 95%: 48-72%)

* δ = 1+ (1-72%)*(1/0.38 - 1) = **1.46**

RELY: Stroke or systemic embolism

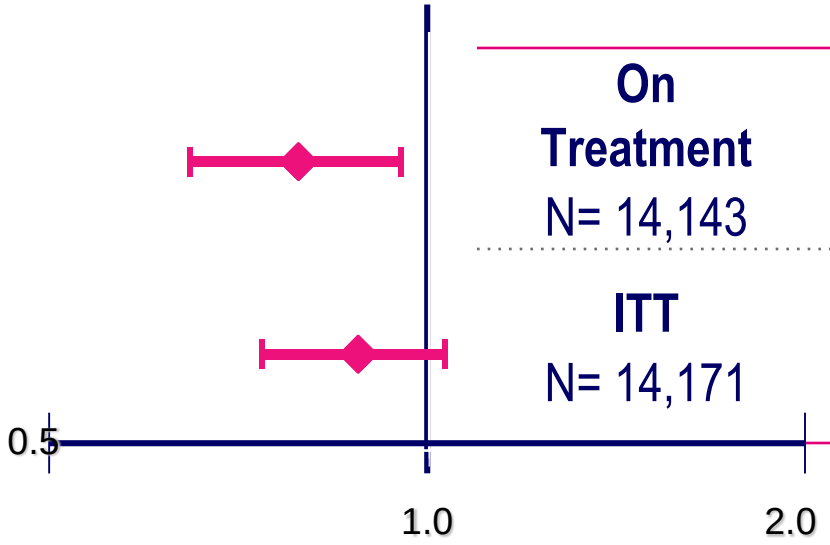


BID = 1 somministrazione ogni 12 ore

Connolly et al. NEJM 2010

ITT vs on-treatment

Stroke and non-CNS Embolism



	Rivaroxaban Event Rate	Warfarin Event Rate	HR (95% CI)	P-value
On Treatment N= 14,143	1.70	2.15	0.79 (0.65,0.95)	0.015
ITT N= 14,171	2.12	2.42	0.88 (0.74,1.03)	0.117

Patel et al. NEJM 2011

ITT or on-treatment?

- In a superiority trial, (aim is to decide if two treatments are different), an ITT analysis is generally conservative (inclusion of protocol violators and withdrawals usually tend to make the results from the 2 treatments more similar).

- For an equivalence or non-inferiority trial this effect is no longer conservative (any blurring of the difference between the treatments will increase the chance of finding non-inferiority).

ITT or on-treatment?

- In a superiority trial per protocol analysis this approach may tend to enhance any difference between the treatments rather than diminishing it, because uninformative "noise" is removed.

- In an equivalence or non- inferiority trial both types of analysis should be per- formed and equivalence or non-inferiority can only be established if both analyses support it.

Per valutare la non inferiorità è necessario:

- **Rigoroso disegno sperimentale:**
  - –includere un popolazione **sovrapponibile** a quella dello studio di riferimento
  - –un trattamento di provata efficacia con la **stessa dose e per un analogo periodo di osservazione**
  - –**Alta adesione** (compliance) alla terapia (se nessuno prende la terapia sia nel gruppo A che nel gruppo B, è facile che il gruppo A risulti 'non inferiore' al gruppo B)
  - –Una **adeguata potenza statistica** per ridurre la probabilità di escludere differenze clinicamente rilevanti

Quality of reports of Equivalence or Noninferiority RCT

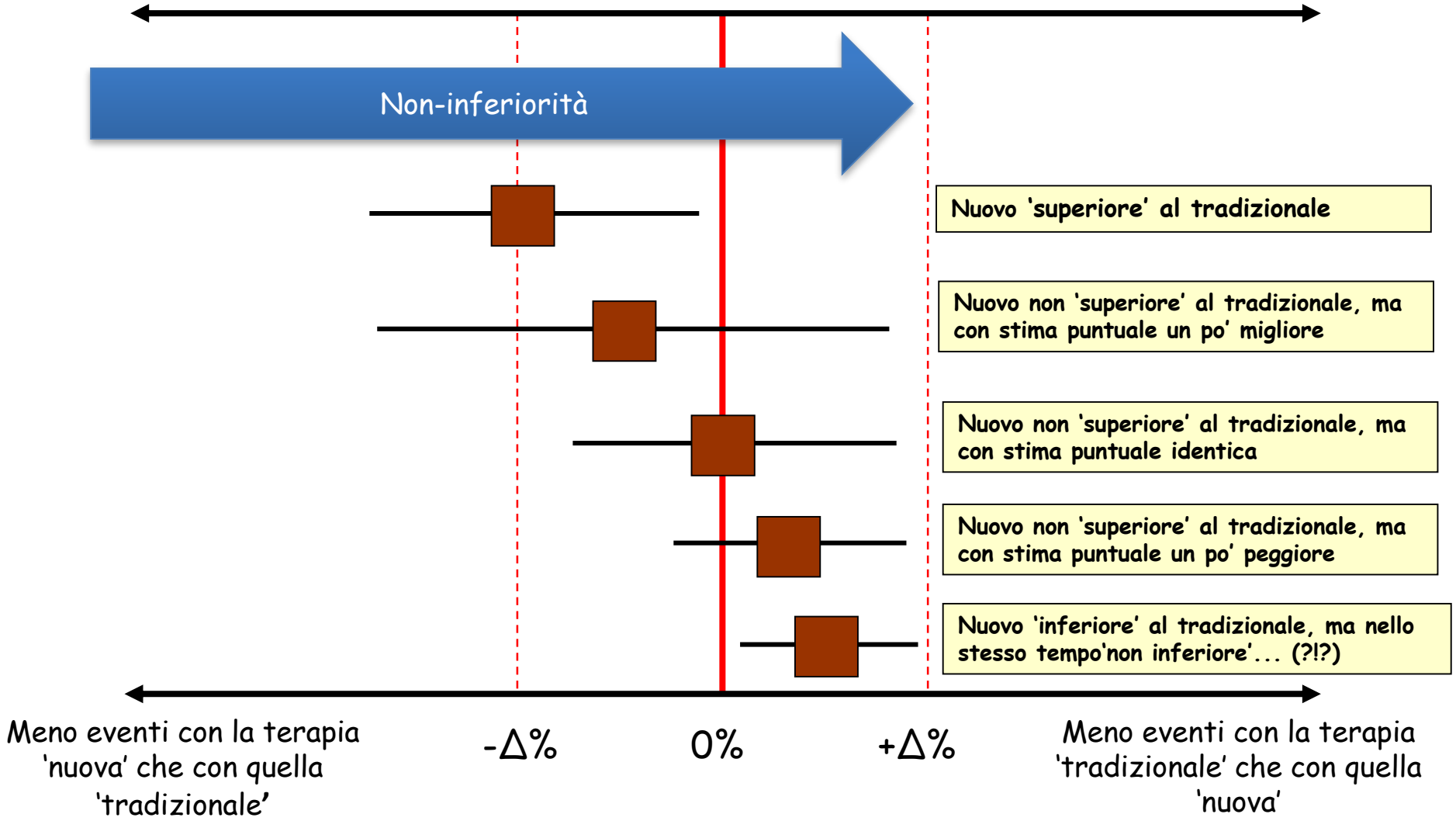
Characteristics of the Trial	No. (%) Noninferiority Reports (n = 116)	Equivalence Reports (n = 46)	All Reports (N = 162)
Noninferiority or equivalence margin reported	112 (96.6)	44 (95.7)	156 (96.3)
Justification of the choice of the margin	24 (20.7)	9 (19.6)	33 (20.4)
Statistical considerations	3 (2.6)	2 (4.3)	5 (3.1)
Clinical considerations or results of a previous study	13 (11.2)	5 (10.9)	18 (11.1)
Both statistical considerations and clinical considerations or results of a previous study	8 (6.9)	2 (4.3)	10 (6.2)
Sample size calculated	92 (79.3)	35 (76.1)	127 (78.4)
Sample size taking into account the noninferiority or equivalence margin	81 (69.8)	35 (76.1)	116 (71.6)
Presence of all the elements needed for recalculation	63 (54.3)	24 (52.2)	87 (53.7)
Analysis			
ITT or modified ITT	90 (77.6)	27 (58.7)	117 (72.2)
Per-protocol	74 (63.8)	35 (76.1)	109 (67.3)
Both per-protocol and ITT, or per-protocol and modified ITT	51 (44.0)	18 (39.1)	69 (42.6)
Statistical methods			
Results with a CI	95 (81.9)	41 (89.1)	136 (84.0)

Le Henanff et al. JAMA 2006

Consort Check-list for noninferiority trials

PAPER SECTION And topic	Item	Descriptor	Reported on Page #
TITLE & ABSTRACT	1	How participants were allocated to interventions (*e.g.*, "random allocation", "randomized", or "randomly assigned"), *specifying that the trial is a non-inferiority or equivalence trial.*	
INTRODUCTION Background	2	Scientific background and explanation of rationale, *including the rationale for using a non-inferiority or equivalence design.*	
METHODS Participants	3	Eligibility criteria for participants *(detailing whether participants in the non-inferiority or equivalence trial are similar to those in any trial(s) that established efficacy of the reference treatment)* and the settings and locations where the data were collected.	
Interventions	4	Precise details of the interventions intended for each group *detailing whether the reference treatment in the non-inferiority or equivalence trial is identical (or very similar) to that in any trial(s) that established efficacy,* and how and when they were actually administered.	
Objectives	5	Specific objectives and hypotheses, *including the hypothesis concerning non-inferiority or equivalence.*	
Outcomes	6	Clearly defined primary and secondary outcome measures *detailing whether the outcomes in the non-inferiority or equivalence trial are identical (or very similar) to those in any trial(s) that established efficacy of the reference treatment* and, when applicable, any methods used to enhance the quality of measurements (*e.g.*, multiple observations, training of assessors).	
Sample size	7	How sample size was determined *detailing whether it was calculated using a non-inferiority or equivalence criterion and specifying the margin of equivalence with the rationale for its choice.* When applicable, explanation of any interim analyses and stopping rules (*and whether related to a non-inferiority or equivalence hypothesis*).	
Randomization -- Sequence generation	8	Method used to generate the random allocation sequence, including details of any restrictions (*e.g.*, blocking, stratification)	
Randomization -- Allocation concealment	9	Method used to implement the random allocation sequence (*e.g.*, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.	
Randomization -- Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.	
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated.	
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s), *specifying whether a one or two-sided confidence interval approach was used.* Methods for additional analyses, such as subgroup analyses and adjusted analyses.	
RESULTS Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.	
Recruitment	14	Dates defining the periods of recruitment and follow-up.	
Baseline data	15	Baseline demographic and clinical characteristics of each group.	
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was *"intention-to-treat" and/or alternative analyses were conducted.* State the results in absolute numbers when feasible (*e.g.*, 10/20, not 50%).	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (*e.g.*, 95% confidence interval). *For the outcome(s) for which non-inferiority or equivalence is hypothesized, a figure showing confidence intervals and margins of equivalence may be useful.*	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.	
Adverse events	19	All important adverse events or side effects in each intervention group.	
DISCUSSION Interpretation	20	Interpretation of the results, taking into account the *non-inferiority or equivalence hypothesis and any other* study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.	

C'è 'non inferiorità' e 'non inferiorità'.....



Differenza nel rischio di eventi

Conclusioni

- Disegno molto utilizzato ultimamente
- Attenzione al Delta!
- Attenzione alla qualità!

Grazie per l'attenzione!?