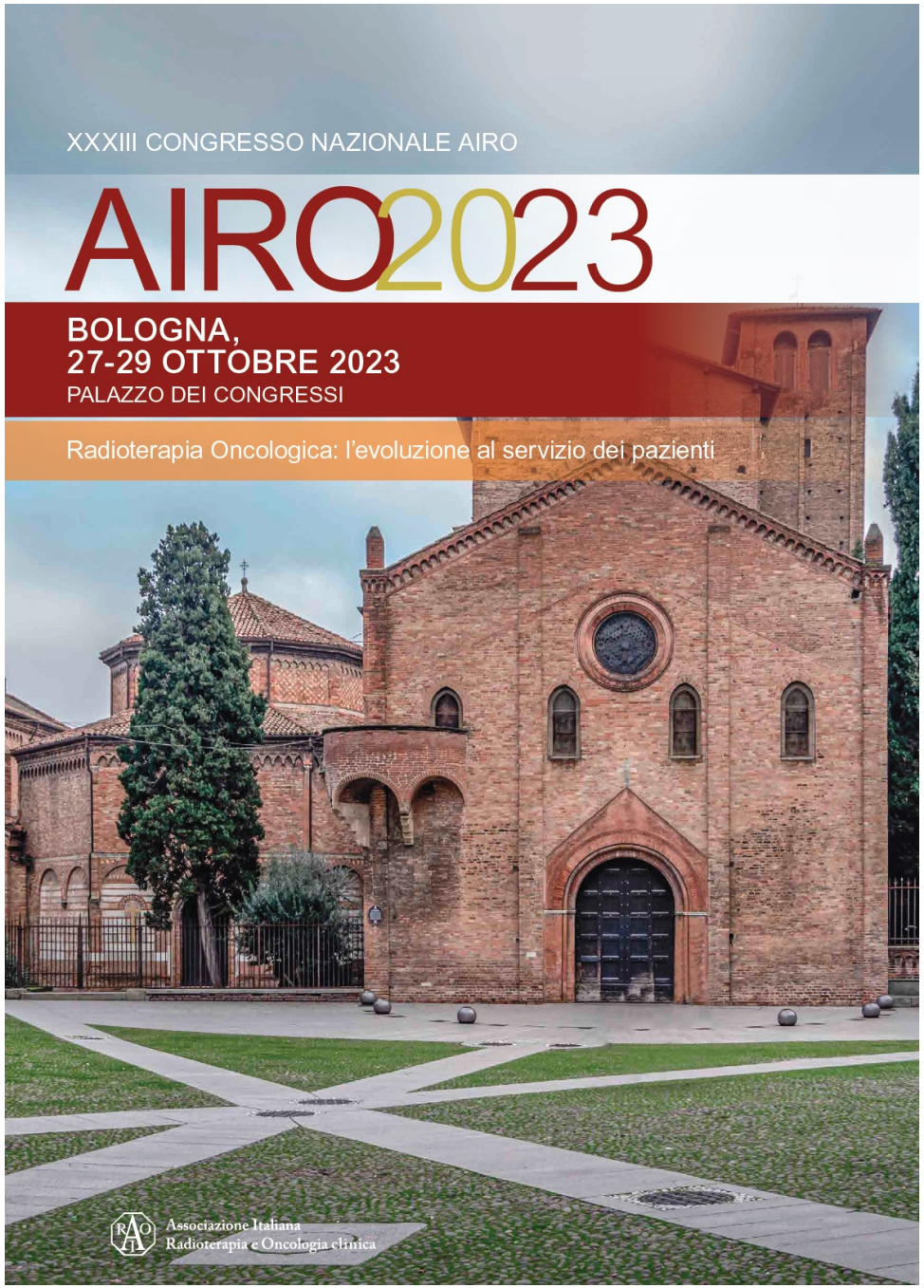


PAPER LAB 1

ECM

Speaker: **R.M. D'Angelillo, G. Pappagallo**

Valutazione delle evidenze disponibili

- Momento fondamentale (mai esclusivo!) nel processo di formazione della proposta terapeutica.
- Tre componenti fondamentali:
  - **rilevanza degli effetti (desiderabili e non desiderabili) osservati**
    - ✓ indicatori, significatività statistica Vs rilevanza clinica
  - **trasferibilità delle evidenze disponibili alla situazione presente**
    - ✓ aderenza al P.I.C.O.
  - **affidabilità (*confidence*) degli effetti osservati**
    - ✓ imprecisione, rischio di bias, eterogeneità

Valutazione delle evidenze disponibili

- Momento fondamentale (mai esclusivo!) nel processo di formazione della proposta terapeutica.
- Tre componenti fondamentali:
  - **rilevanza degli effetti (desiderabili e non desiderabili) osservati**
    - ✓ **indicatori**, significatività statistica *Vs* rilevanza clinica
  - trasferibilità delle evidenze disponibili alla situazione presente
    - ✓ aderenza al P.I.C.O.
  - affidabilità (*confidence*) degli effetti osservati
    - ✓ imprecisione, rischio di bias, eterogeneità

Variabili di comune riscontro negli studi clinici e indicatori comparativi di effetto

Endpoint	Tipo di variabile statistica	Indicatore comparativo di effetto	
		relativo	assoluto
Attività (CR, PR, SD, ecc.) Tossicità (eventi avversi)	nominale	RR (relative risk) OR (odds ratio)	RD (risk difference)
Punteggi di qualità di vita	intervallare	-	MD (mean difference)
OS (overall survival) PFS (progression-free survival DFS (disease-free survival) TTD (time to deterioration…)	tempo a evento	HR (hazard ratio) - - -	differenza tra mediane differenza tra stime al tempo t… RMST (restricted mean survival time) RD (risk difference)

Risk, Odds…

Misure di effetto relativo

Risk 70 risposte / 100 pazienti = 0.70

Risk Ratio $\frac{\text{70 risposte / 100 pazienti (braccio A)}}{\text{30 risposte / 100 pazienti (braccio B)}} = \frac{0.70}{0.30} = 2.33$

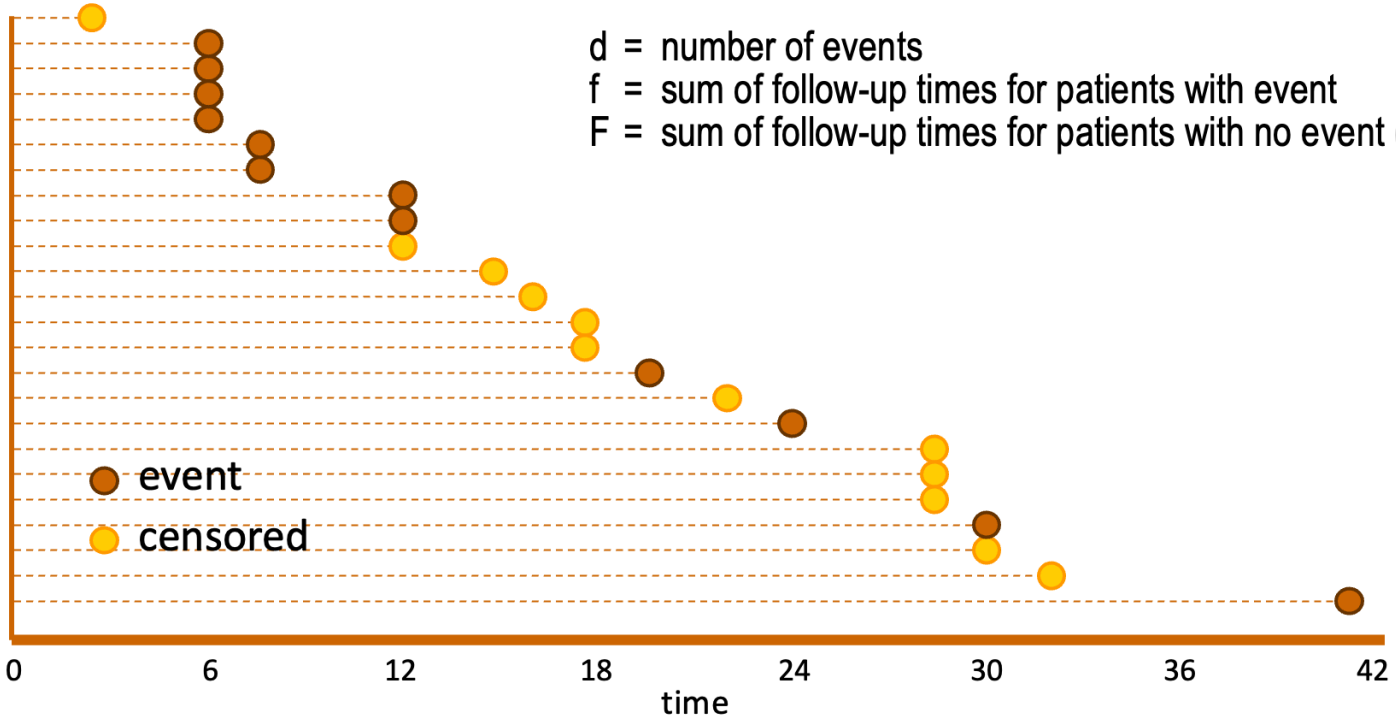
Odds 65 risposte / 35 non risposte = 1.86

Odds Ratio $\frac{\text{65 risposte / 35 non risposte (braccio A)}}{\text{25 risposte / 75 non risposte (braccio A)}} = \frac{1.86}{0.33} = 5.63$

Misura di effetto assoluto

Risk Difference 0.70 – 0.30 = 0.40, ovvero: 40 risposte *in più* ogni 100 pazienti trattati

… e *hazard rate*



d = 12

f = 6+6+6+6+8+8+12+12+20+24+30+42 = 180

F = 3+12+15+16+18+18+22+28+28+28+30+33 = 251

$$\textit{hazard rate} = \frac{12}{431} = 0.0278$$

$$\textbf{Hazard Ratio} = \frac{\textit{hazard rate}\ \text{trattamento sperimentale}}{\textit{hazard rate}\ \text{trattamento di controllo}}$$

Valutazione delle evidenze disponibili

- Momento fondamentale (mai esclusivo!) nel processo di formazione della proposta terapeutica.
- Tre componenti fondamentali:
  - **rilevanza degli effetti (desiderabili e non desiderabili) osservati**
    - ✓ indicatori, **significatività statistica *Vs* rilevanza clinica**
  - trasferibilità delle evidenze disponibili alla situazione presente
    - ✓ aderenza al P.I.C.O.
  - affidabilità (*confidence*) degli effetti osservati
    - ✓ imprecisione, rischio di bias, eterogeneità

Statistical Vs Clinical Significance

- **Statistical Significance**

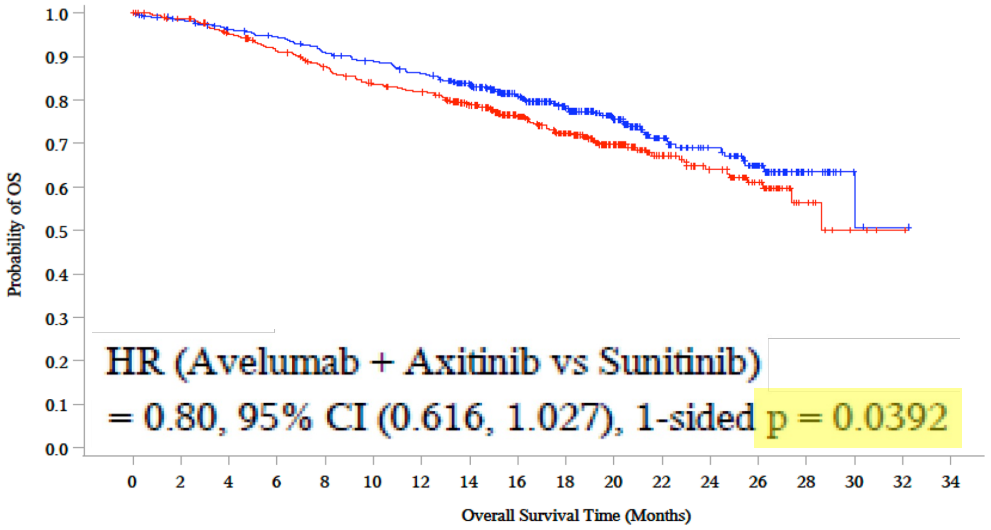
  "Is an observed difference **likely to be real**"

  ✓ dependent on the magnitude of the number of patients NOT on whether the difference is meaningful for patients

- quando il valore di **p** risultante dal test di significatività è più piccolo del valore soglia (usualmente 5%), si considera lo studio (statisticamente) positivo;
- se il valore di **p** è maggiore del 5%, si considera lo studio (statisticamente) negativo

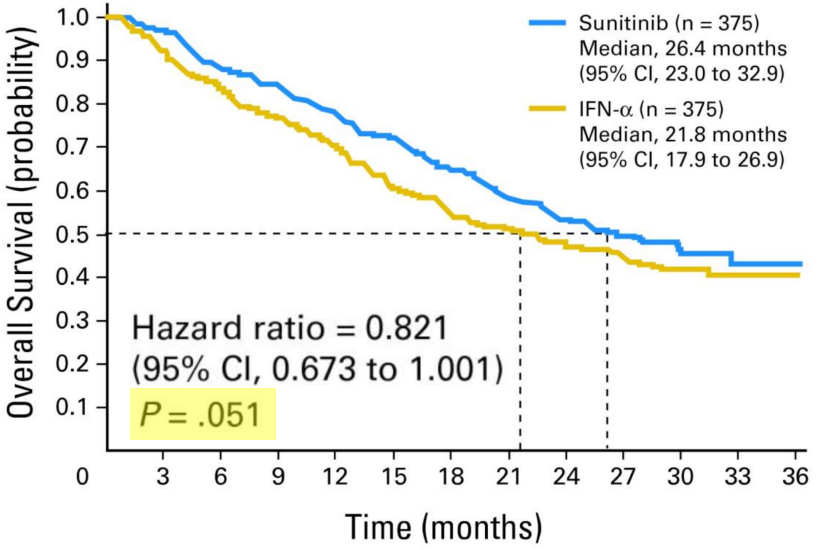
P-value is a measure of the probability that an observed difference could have occurred just by random chance.

- Don’t base your conclusions solely on whether an association or effect was found to be “statistically significant”.



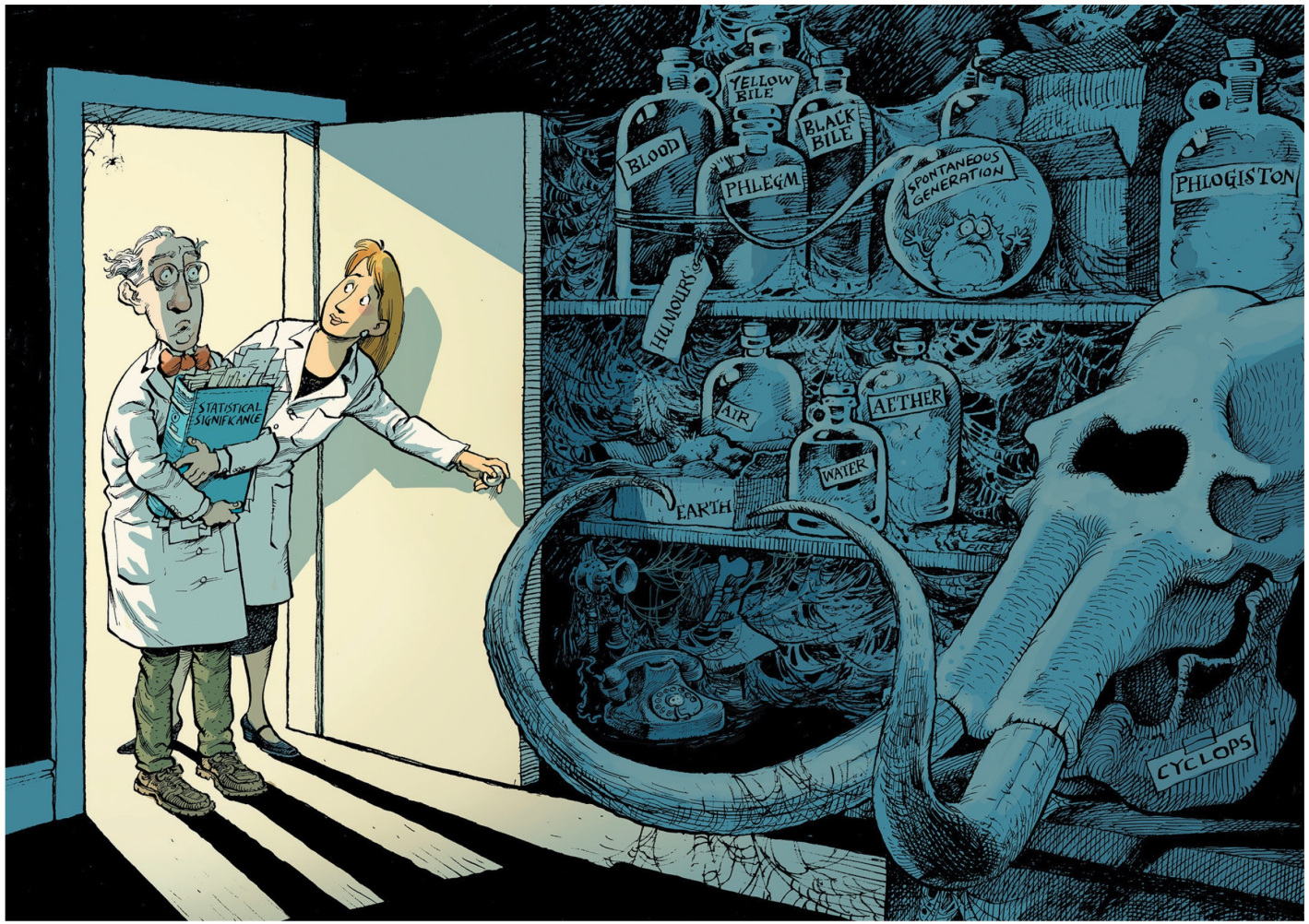
P-value is a measure of the probability that an observed difference could have occurred just by random chance.

- Don't base your conclusions solely on whether an association or effect was found to be "statistically significant".
- Don't believe that an association or effect is absent just because it was not statistically significant.



P-value is a measure of the probability that an observed difference could have occurred just by random chance.

- Don't base your conclusions solely on whether an association or effect was found to be "statistically significant".
- Don't believe that an association or effect is absent just because it was not statistically significant.
- Don't conclude anything about scientific or practical importance based on statistical significance (or lack thereof).



Retire statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

21 MARCH 2019 | VOL 567 | NATURE | 305

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

Obiettivi di uno studio comparativo

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

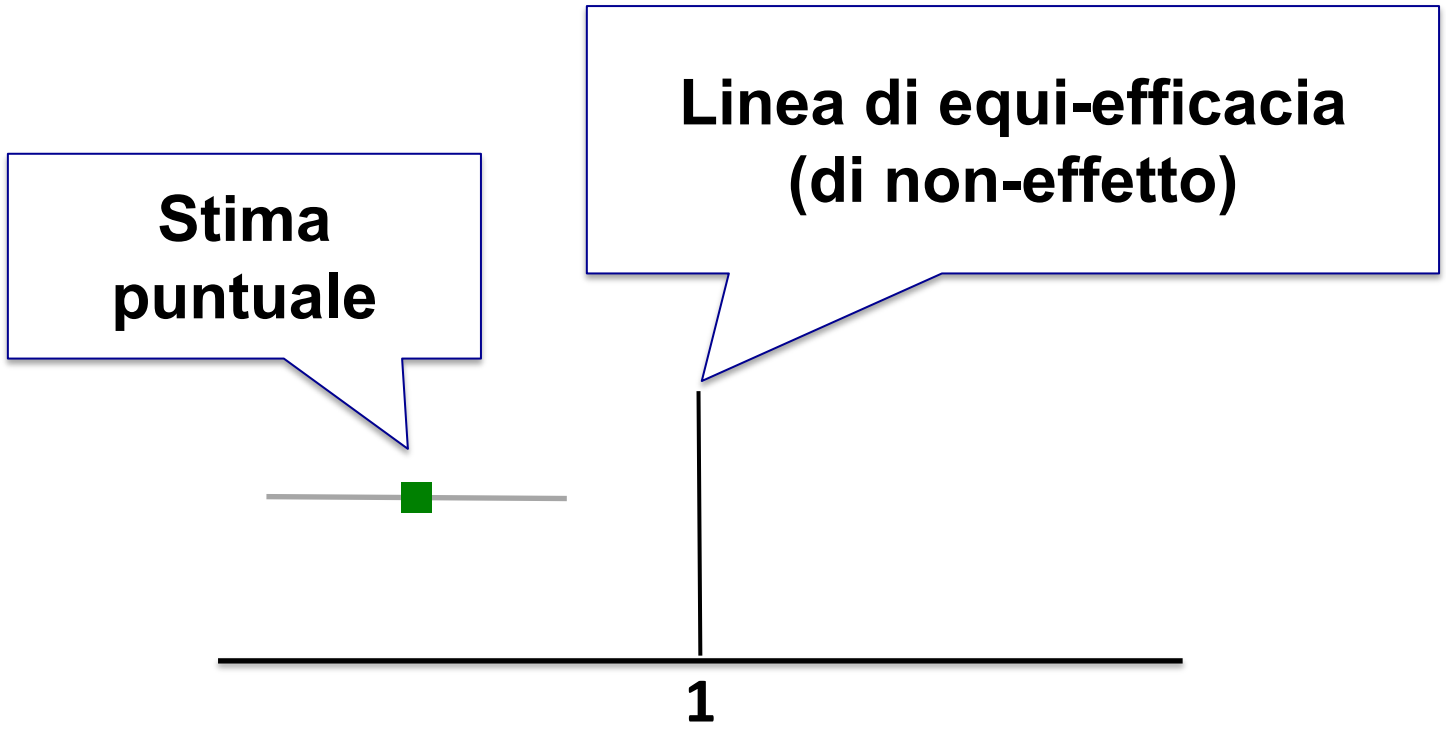
studio di superiorità

A > B di una quantità Δ di interesse clinico

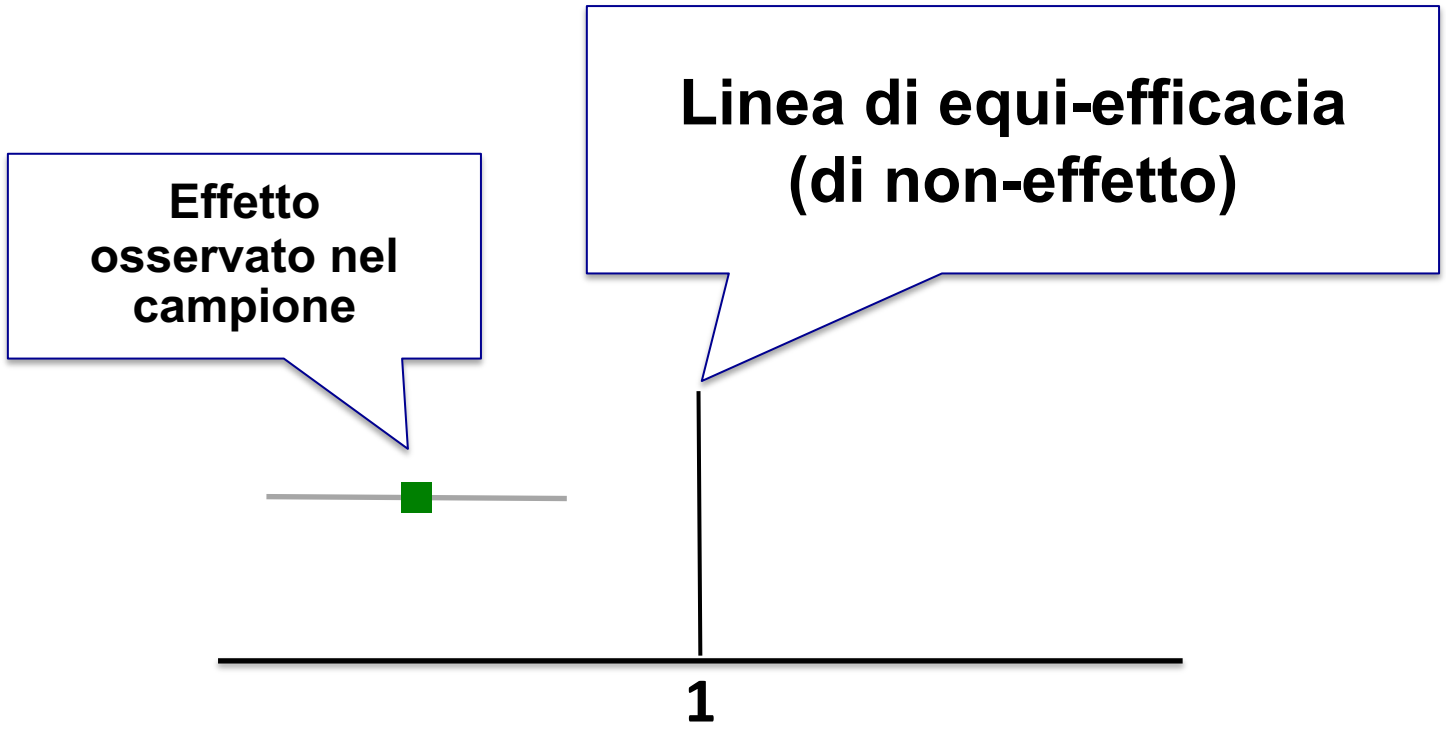
studio di non inferiorità

A < B non oltre una quantità M di rilevanza clinica

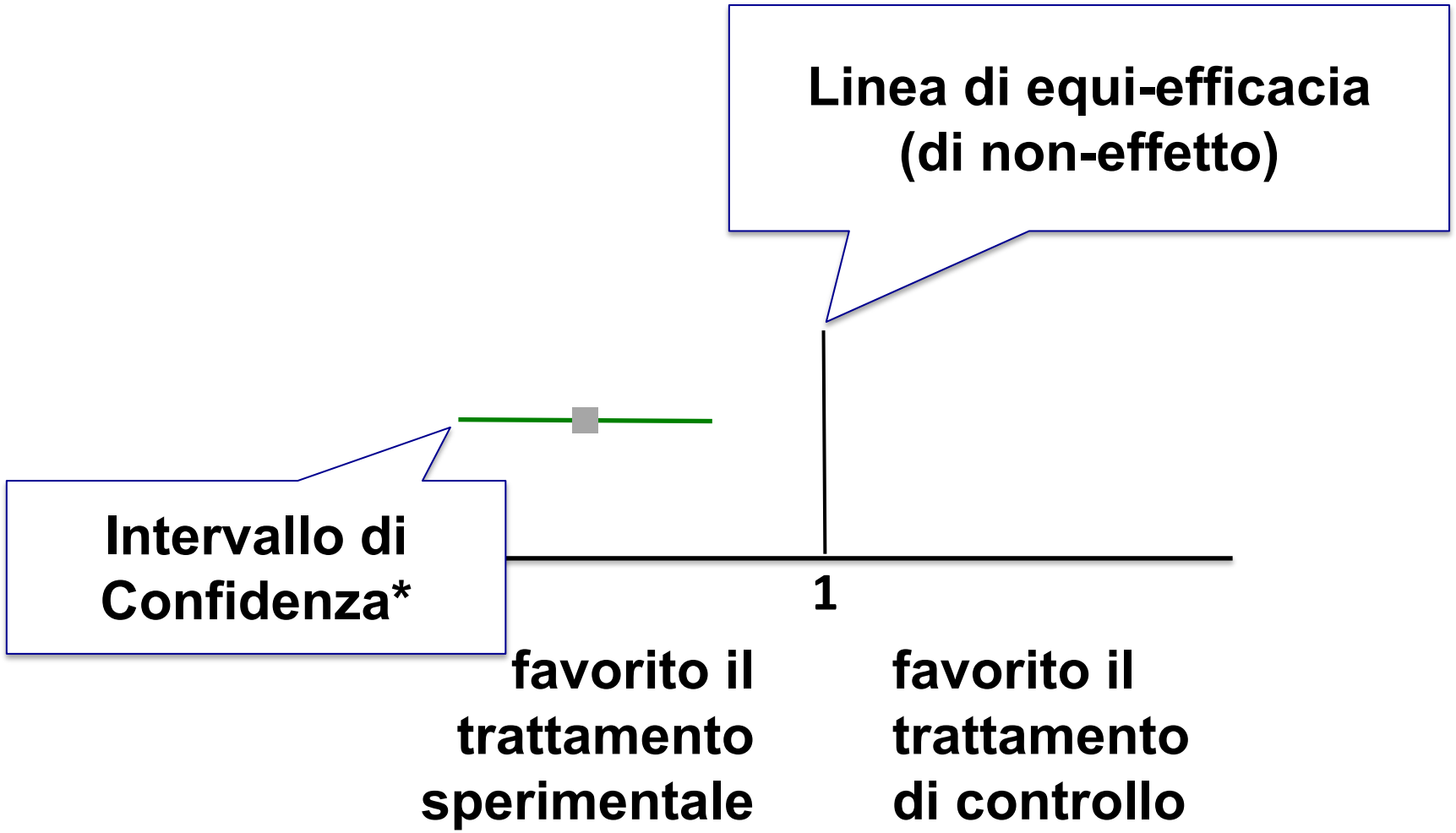
Interpretazione degli studi clinici mediante Forest Plot



Interpretazione degli studi clinici mediante Forest Plot

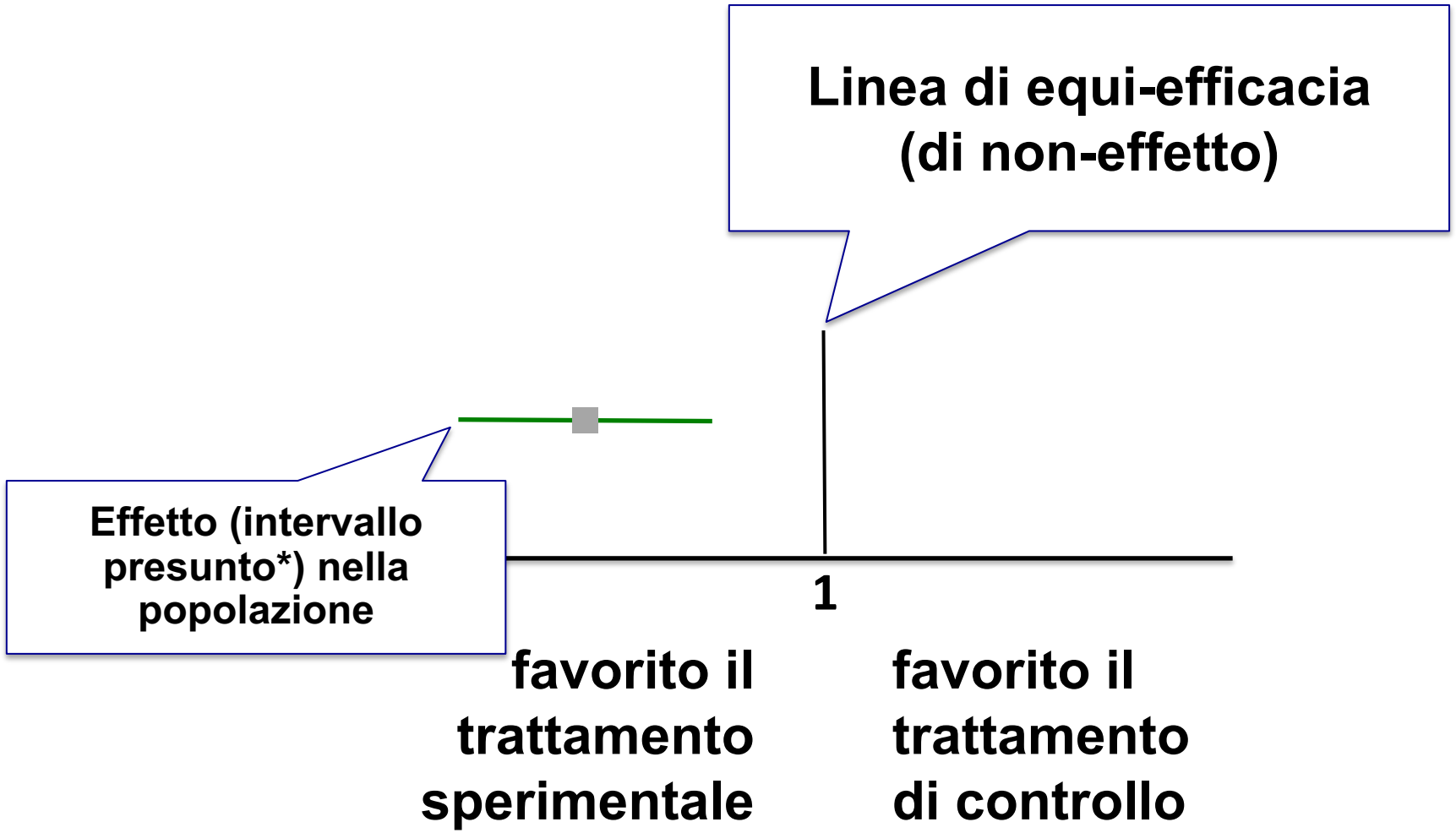


Interpretazione degli studi clinici mediante Forest Plot



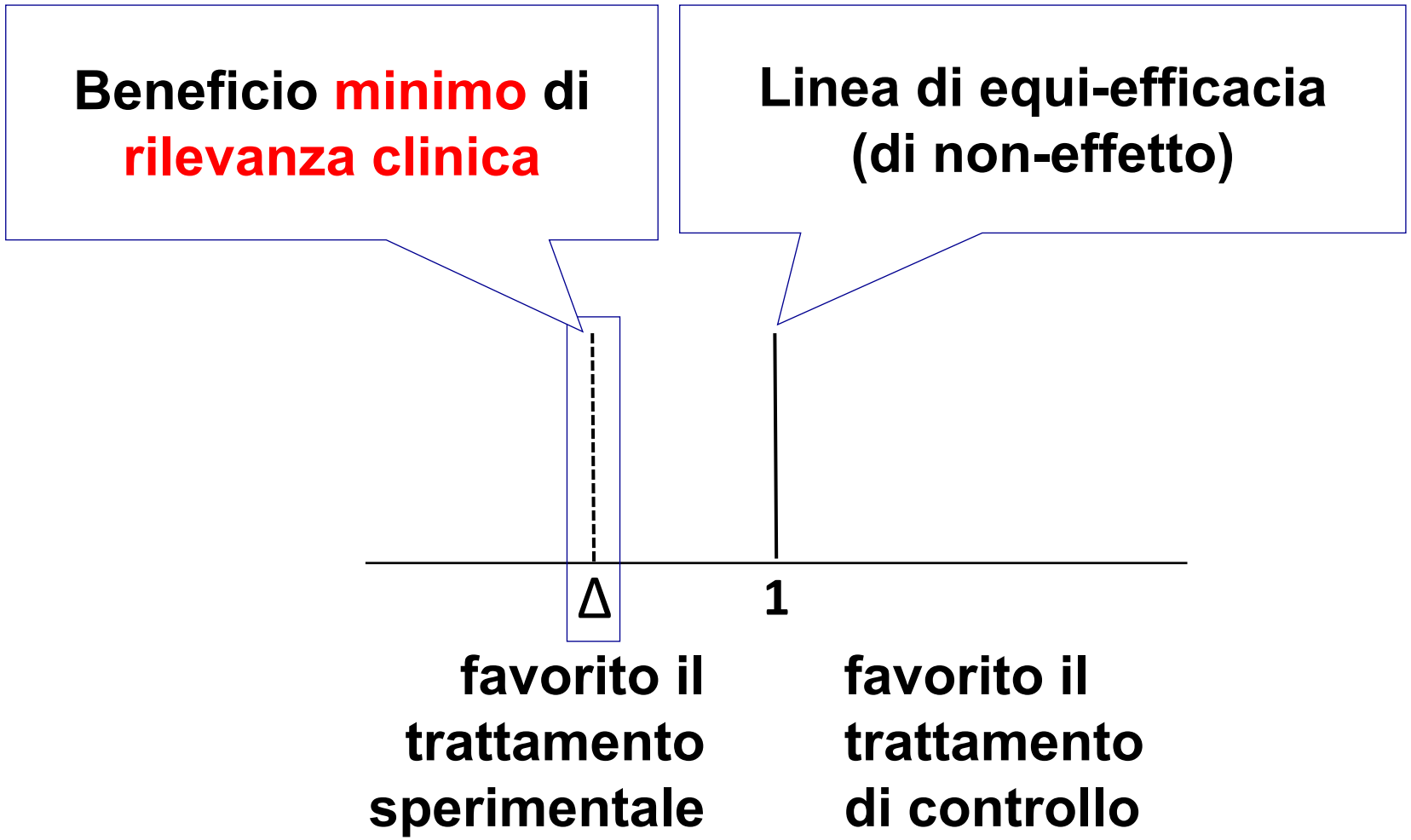
* convenzionalm. 95%

Interpretazione degli studi clinici mediante Forest Plot

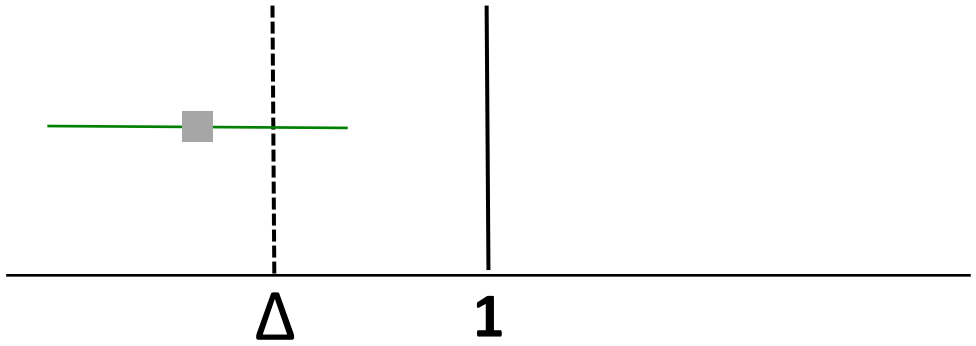


* convenzionalm. 95%

Interpretazione degli studi clinici mediante Forest Plot

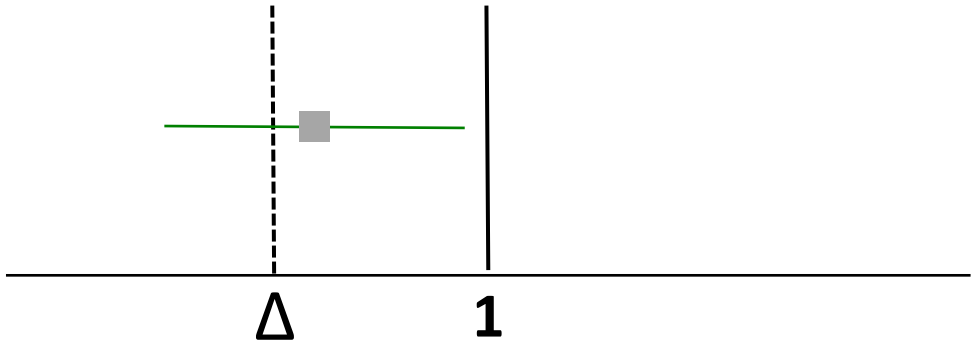


Interpretazione degli studi clinici mediante Forest Plot



favorito il trattamento sperimentale

favorito il trattamento di controllo



favorito il trattamento sperimentale

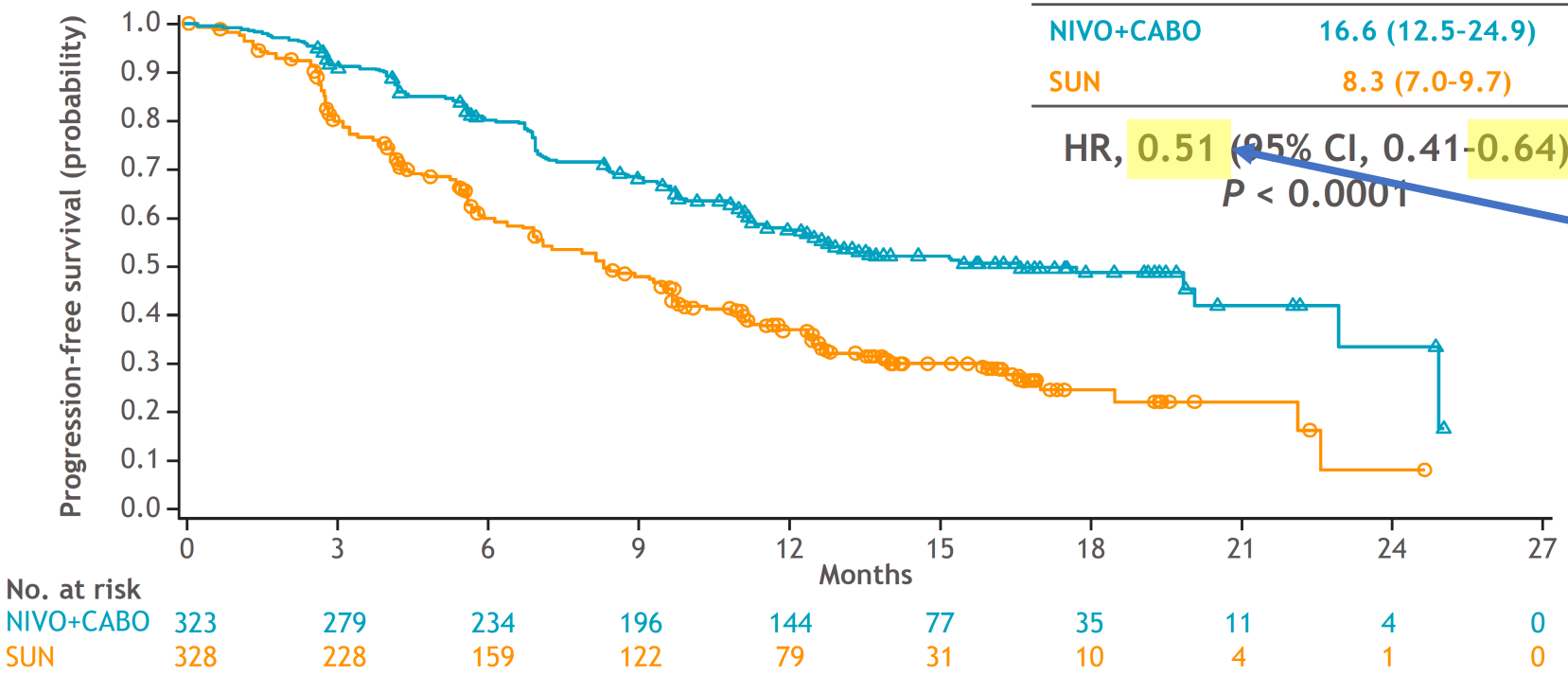
favorito il trattamento di controllo

VIRTUAL 2020 ESMO congress

Nivolumab plus cabozantinib versus sunitinib in first-line treatment for advanced renal cell carcinoma: first results from the randomized phase 3 CheckMate 9ER trial

Toni K. Choueiri,[1] Thomas Powles,[2] Mauricio Burotto,[3] Maria T. Bourlon,[4] Bogdan Zurawski,[5] Víctor Manuel Oyervides Juárez,[6] James J. Hsieh,[7] Umberto Basso,[8] Amishi Y. Shah,[9] Cristina Suarez,[10] Alketa Hamzaj,[11] Carlos Barrios,[12] Martin Richardet,[13] David Pook,[14] Yoshihiko Tomita,[15] Bernard Escudier,[16] Joshua Zhang,[17] Burcin Simsek,[17] Andrea B. Apolo,[18] Robert J. Motzer[19]

Progression-free survival per BICR



It was estimated that ~350 progression or death events would result at least 95% power to detect a hazard ratio of 0.68

arativo

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

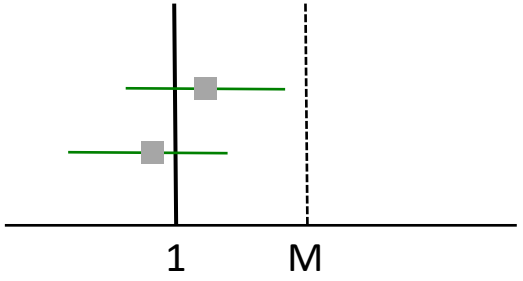
A > B di una quantità Δ di interesse clinico

studio di non inferiorità

A < B non oltre una quantità M di rilevanza clinica

Interpretazione clinica di uno Studio di Non-Inferiorità

(dato uno specifico M di interesse)



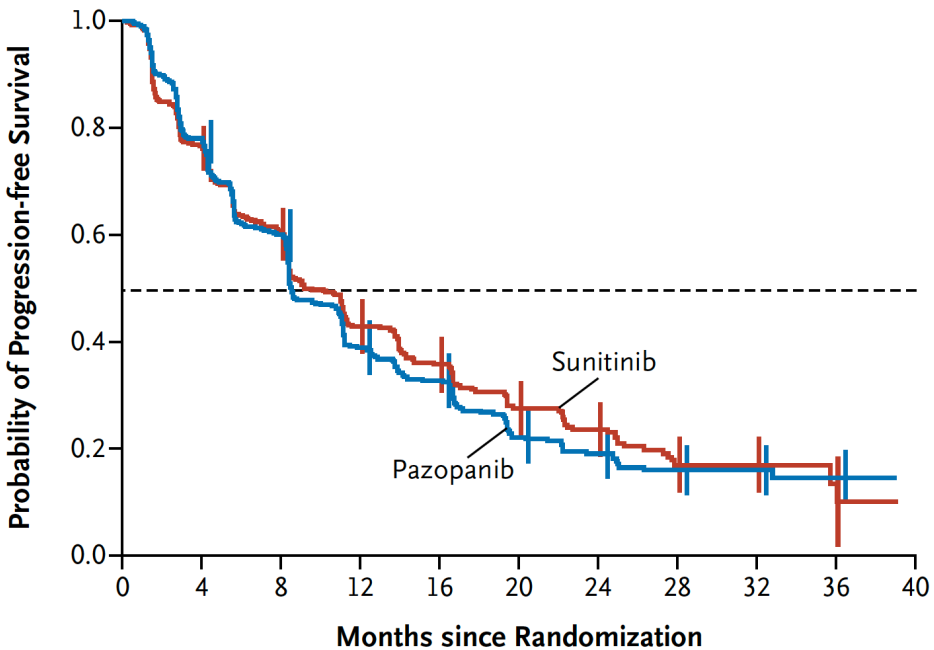
Dimostrazione di Non-Inferiorità

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

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



STATISTICAL ANALYSIS

We calculated that 631 disease-progression events were required for the study to have 80% power to reject the null hypothesis of an increased risk in the hazard of disease progression with pazopanib (hazard ratio, ≥1.25).

hazard ratio 1.05 (95% CI, 0.90 to 1.22)

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., 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	3			0.03	Pazopanib	0.14
Disease-related physical symptoms	3			0.03	Pazopanib	0.13
Disease-related emotional symptoms	3			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

Minimal (Clinical) Interesting Difference (MID / MCID)

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

- it’s easily understood by clinicians as a key concept in the interpretability of PRO scores;
- will inform judgments about the success-fulness of an intervention;

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.

Valutazione delle evidenze disponibili

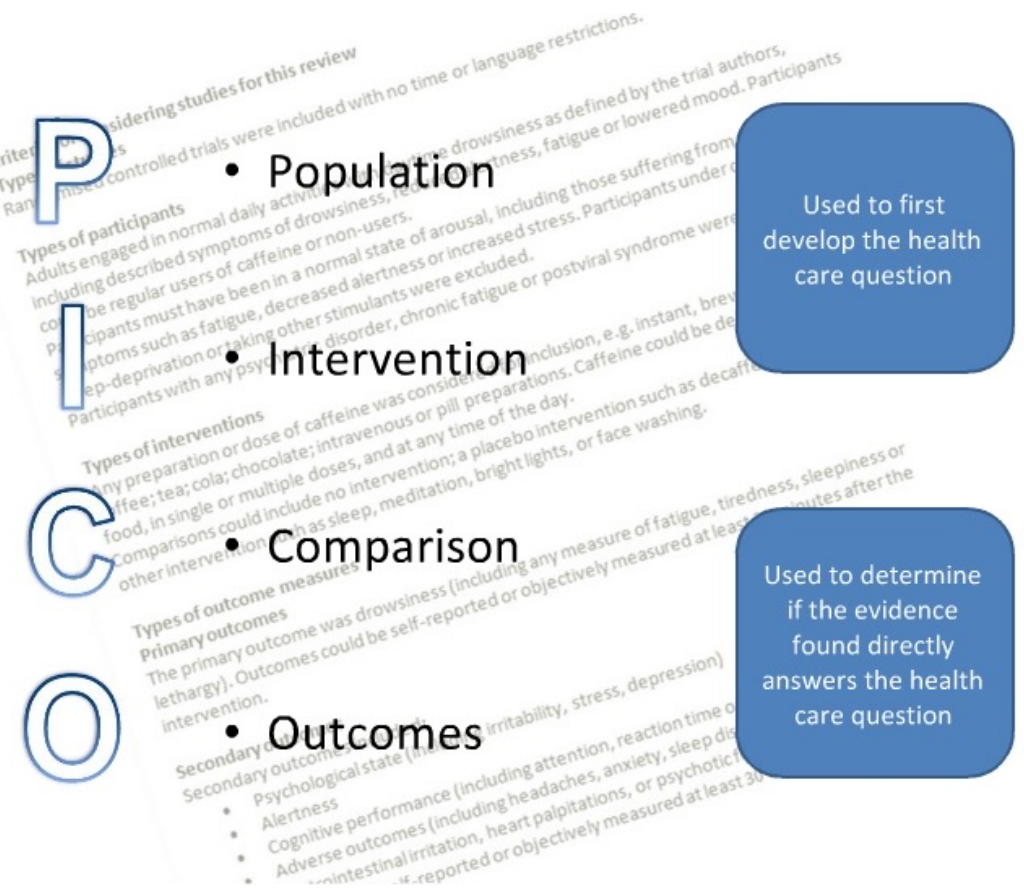
- Momento fondamentale (mai esclusivo!) nel processo di formazione della proposta terapeutica.
- Tre componenti fondamentali:
  - rilevanza degli effetti (desiderabili e non desiderabili) osservati
    - ✓ significatività statistica *Vs* rilevanza clinica
  - **trasferibilità delle evidenze disponibili alla situazione presente**
    - ✓ **aderenza al P.I.C.O.**
  - affidabilità (*confidence*) degli effetti osservati
    - ✓ imprecisione, rischio di bias, eterogeneità

Direct evidence…

…comes from research that:

- is conducted in the **Population** that we are providing answers for;
- includes the **Intervention** that we are interested in…
- …and compares these interventions with the appropriate **Alternatives**;
- measures the **Outcomes** in which we are interested

Un trial clinico non dovrebbe essere *letto così com'è*, ma avendo come riferimento uno specifico quesito clinico.

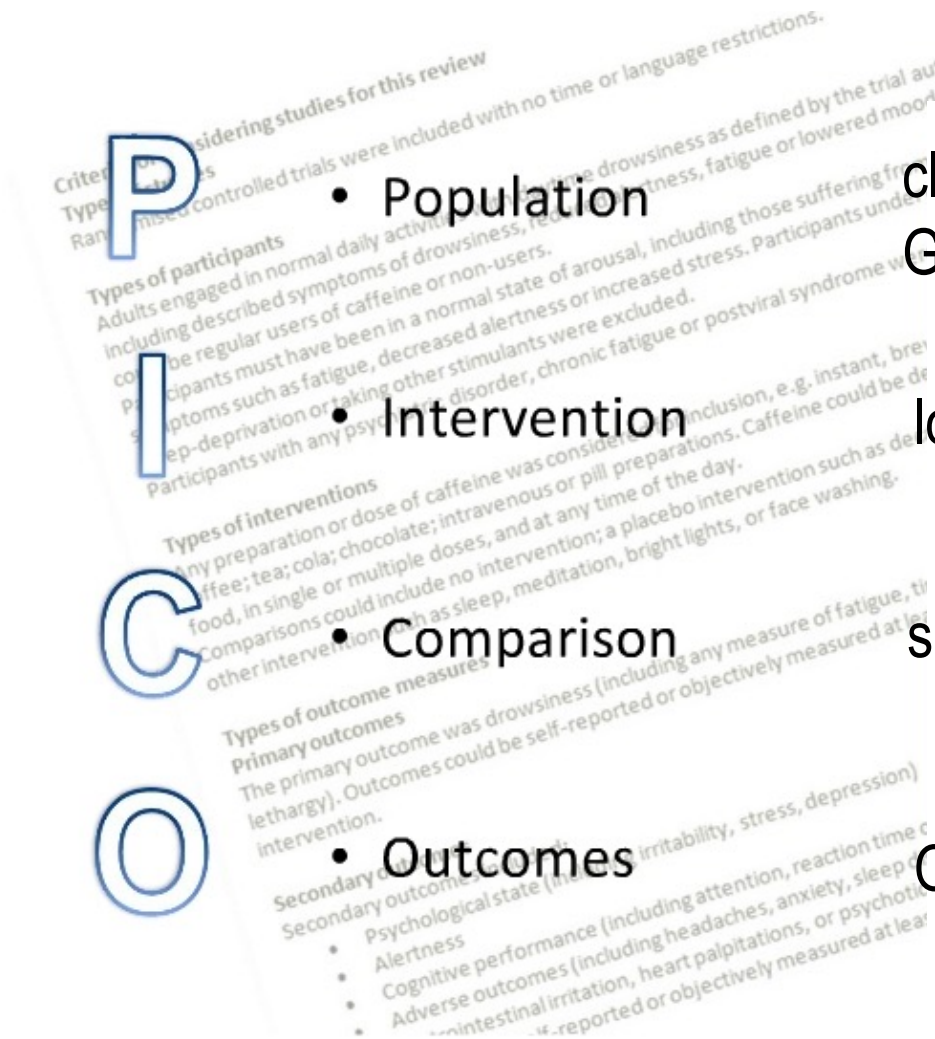


High-dose radiotherapy with short-term or long-term androgen deprivation in localised prostate cancer (DART01/05 GICOR): a randomised, controlled, phase 3 trial

Almudena Zapatero, Araceli Guerrero, Xavier Maldonado, Ana Alvarez, Carmen Gonzalez San Segundo, Maria Angeles Cabeza Rodríguez, Victor Macias, Agustí Pedro Olive, Francesc Casas, Ana Boladeras, Carmen Martín de Vidales, Maria Luisa Vazquez de la Torre, Salvador Villà, Aitor Perez de la Haza, Felipe A Calvo

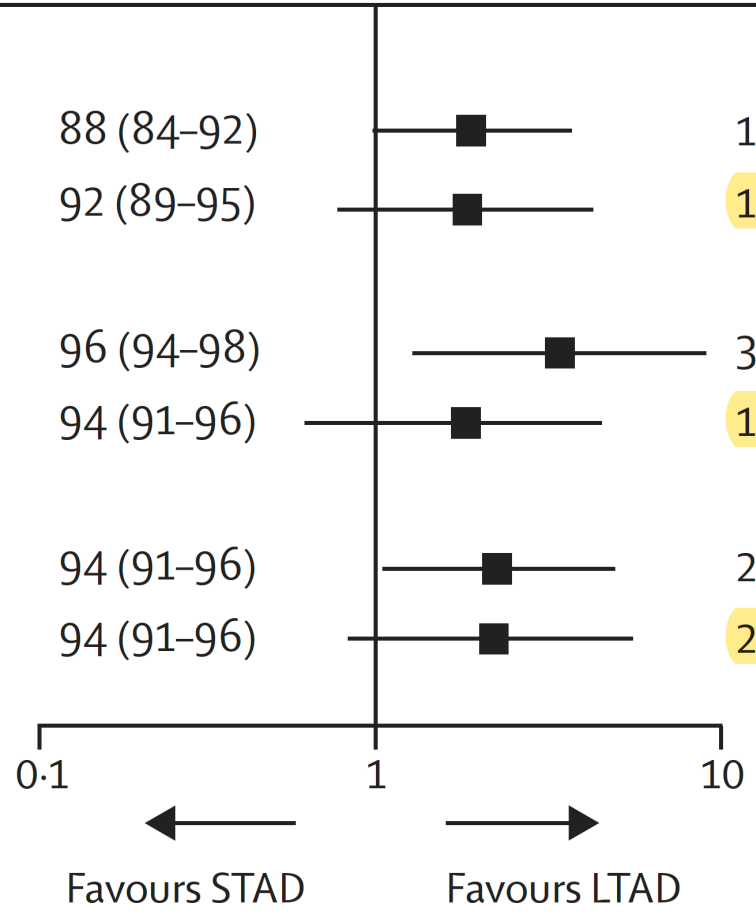
Lancet Oncol 2015; 16: 320–27

Findings Between Nov 7, 2005, and Dec 20, 2010, 178 patients were randomly assigned to receive short-term androgen deprivation and 177 to receive long-term androgen deprivation. After a median follow-up of 63 months (IQR 50–82), 5-year biochemical disease-free survival was significantly better among patients receiving long-term androgen deprivation than among those receiving short-term treatment (90% [95% CI 87–92] *vs* 81% [78–85]; hazard ratio [HR] 1·88 [95% CI 1·12–3·15]; p=0·01). 5-year overall survival (95% [95% CI 93–97] *vs* 86% [83–89]; HR 2·48 [95% CI 1·31–4·68]; p=0·009) and 5-year metastasis-free survival (94% [95% CI 92–96] *vs* 83% [80–86]; HR 2·31 [95% CI 1·23–3·85]; p=0·01) were also significantly better in the long-term androgen deprivation group than in the short-term androgen deprivation group.



- **P** Population: clinical stage T2N0M0 prostate adenocarcinoma, Gleason 7, PSA 20 ng/ml
- **I** Intervention: long-term androgen deprivation (LTAD)
- **C** Comparison: short-term androgen deprivation (STAD)
- **O** Outcomes: OS, MFS, bDFS

	Number of events			5-year rate (%, 95% CI)		Hazard ratio (95% CI)	p value
	N	STAD	LTAD	STAD	LTAD		
Biochemical disease-free survival							
High risk	189	23	13	76 (71–80)	88 (84–92)	1·91 (0·97–3·77)	0·054
Intermediate risk	166	14	8	88 (84–91)	92 (89–95)	1·82 (0·76–4·33)	0·174
Overall survival							
High risk	189	17	5	82 (77–86)	96 (94–98)	3·43 (1·26–9·32)	0·015
Intermediate risk	166	10	6	91 (88–95)	94 (91–96)	1·67 (0·61–4·60)	0·318
Metastasis-free survival							
High risk	189	20	9	79 (74–83)	94 (91–96)	2·27 (1·04–5·01)	0·041
Intermediate risk	166	13	6	89 (85–93)	94 (91–96)	2·14 (0·81–5·66)	0·124



Lancet Oncol 2015; 16: 320–27

Procedures

Radiotherapy was administered with three-dimensional conformal radiotherapy techniques done with a six-field isocentric beam setup based on a CT scan. The target volume included the prostate and the seminal vesicles. In view of the controversy regarding the role of prophylactic pelvic radiotherapy and the absence of definitive data, elective pelvic radiotherapy was left to the criteria of each participating centre.

Lancet Oncol 2015; 16: 320–27

Outcomes

The primary endpoint was biochemical disease-free survival, defined as the time from randomisation to progression of biochemical disease, or death from any cause, or censoring at the date of the last contact. The RTOG-ASTRO Phoenix Consensus Conference definition[20] (an increase in the PSA concentration of ≥2 ng/mL above the nadir) was used to define biochemical failure. Secondary endpoints included overall survival, distant metastasis-free survival, and cause-specific survival. Overall survival was defined as the time from

Lancet Oncol 2015; 16: 320–27

Valutazione delle evidenze disponibili

- Momento fondamentale (mai esclusivo!) nel processo di formazione della proposta terapeutica.
- Tre componenti fondamentali:
  - rilevanza degli effetti (desiderabili e non desiderabili) osservati
    - ✓ significatività statistica *Vs* rilevanza clinica
  - trasferibilità delle evidenze disponibili alla situazione presente
    - ✓ aderenza al P.I.C.O.
  - **affidabilità (*confidence*) degli effetti osservati**
    - ✓ **imprecisione**, rischio di bias, eterogeneità

Caso Vs Bias

CASO

Errore Random

Risultati Imprecisi

Errore in diminuzione con l'aumentare delle dimensioni del campione

BIAS

Errore Sistematico

Risultati Inesatti

Errore non influenzato dalle dimensioni del campione

Journal of Clinical Epidemiology 64 (2011) 1283–1293

GRADE guidelines 6. Rating the quality of evidence—imprecision

Gordon H. Guyatt[a,b,*], Andrew D. Oxman[c], Regina Kunz[d,e], Jan Brozek[a], Pablo Alonso-Coello[f], David Rind[g], PJ Devereaux[a], Victor M. Montori[h], Bo Freyschuss[i], Gunn Vist[c], Roman Jaeschke[b], John W. Williams Jr.[j], Mohammad Hassan Murad[h], David Sinclair[k], Yngve Falck-Ytter[l], Joerg Meerpohl[m,n], Craig Whittington[o], Kristian Thorlund[a], Jeff Andrews[p], Holger J. Schünemann[a,b]

Key Points

- GRADE's primary criterion for judging precision is to focus on the 95% confidence interval (CI) around the difference in effect between intervention and control for each outcome.
- In general, the CIs to consider are those around the absolute, rather than the relative effect.
- If a recommendation or clinical course of action would differ if the upper versus the lower boundary of the CI represented the truth, consider the rating down for imprecision.

PRESENTED AT: 2020 ASCO ANNUAL MEETING #ASCO20 #TheraP

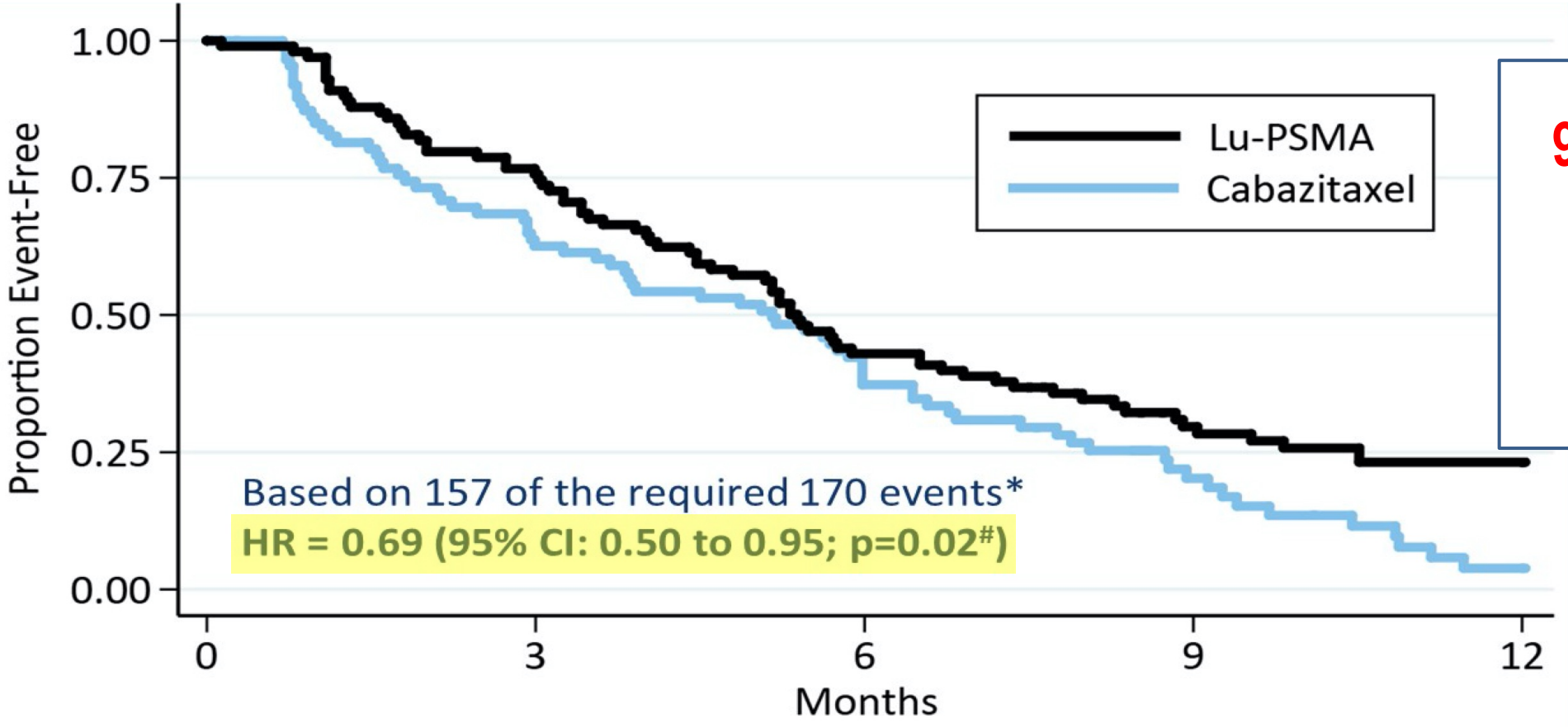
Slides are the property of the author, permission required for reuse.

A randomised phase II trial of ^{177}Lu-PSMA-617 (Lu-PSMA) theranostic versus cabazitaxel in metastatic castration resistant prostate cancer (mCRPC) progressing after docetaxel: Initial results

TheraP (ANZUP 1603)

Michael Hofman, Louise Emmett, Shahneen Sandhu, Amir Iravani, Anthony Joshua, Jeffrey Goh, David Pattison, Hsiang Tan, Ian Kirkwood, Siobhan Ng, Roslyn Francis, Craig Gedye, Natalie Rutherford, Alison Zhang, Margaret McJannett, Martin Stockler, John Violet, Scott Williams, Andrew Martin, Ian Davis

- median follow-up: 13.3 months
- baseline risk: 95.1%
- absolute risk: 8 events fewer (95%CI: 1 fewer to 17 fewer)



95%CI of absolute effect consistent with opposite interpretations

Caso Vs Bias

CASO

↓

Errore Random

↓

Risultati Imprecisi

Errore in diminuzione con l'aumentare delle dimensioni del campione

BIAS

↓

Errore Sistematico

↓

Risultati Inesatti

Errore non influenzato dalle dimensioni del campione

I *bias* fondamentali… in pillole

- selezione dei soggetti da assegnare a uno specifico trattamento, rimozione di alcuni pazienti dall'analisi, analisi non preordinata su sottogruppi di pazienti (*selection bias*)

  rimedio: randomizzazione (e stratificazione)

- conoscenza da parte del paziente/medico/valutatore del trattamento in atto (*performance/detection bias*)

  rimedio: mascheramento

- perdita dei pazienti alla valutazione di uno specifico outcome di interesse (*attrition bias*)

  rimedio: disegno di studio e procedure di follow-up

- selezione dei risultati da comunicare a mezzo presentazione/pubblicazione (*reporting bias*)

  rimedio: database degli studi clinici in corso

Valutazione delle evidenze disponibili

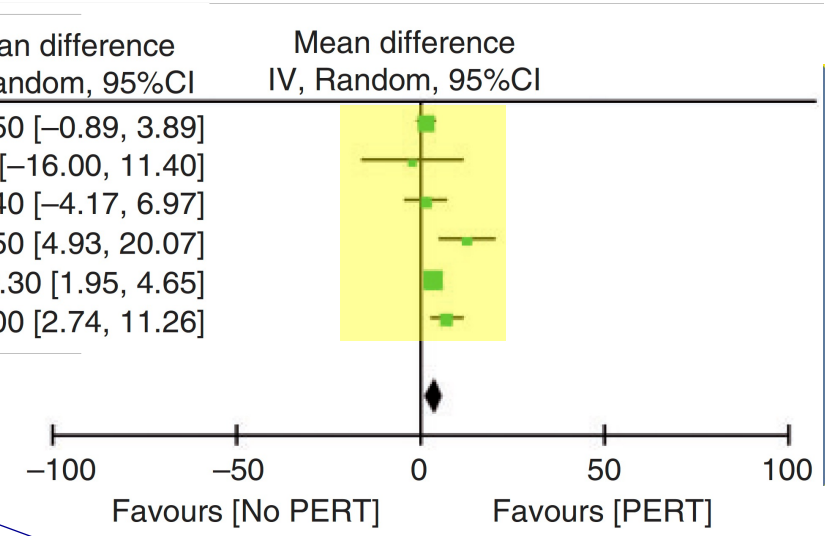
- Momento fondamentale (mai esclusivo!) nel processo di formazione della proposta terapeutica.
- Tre componenti fondamentali:
  - rilevanza degli effetti (desiderabili e non desiderabili) osservati
    - ✓ significatività statistica *Vs* rilevanza clinica
  - trasferibilità delle evidenze disponibili alla situazione presente
    - ✓ aderenza al P.I.C.O.
  - **affidabilità (*confidence*) degli effetti osservati**
    - ✓ imprecisione, rischio di bias, **eterogeneità**

Heterogeneity

Study or subgroup	PERT Mean	PERT SD	PERT Total	No pert Mean	No pert SD	No pert Total	Mean difference IV, random, 95%CI
Partelli 2012	12.5	5.5	97	11	8.5	97	1.50 [–0.89, 3.89]
Woo 2016	5.8	28.6	34	8.1	28.6	33.2	–2.30 [–16.00, 11.40]
Saito 2017	15	17.1	46	13.6	8.8	45	1.40 [–4.17, 6.97]
Zdenkowski 2017	16.9	10.2	9	4.4	5.5	9	12.50 [4.93, 20.07]
Dominguez-Munoz 2018	6.6	3.8	40	3.3	2.5	55	3.30 [1.95, 4.65]
Saito 2018	19	9.3	43	12	10.4	40	7.00 [2.74, 11.26]
Total (95% CI)							

Heterogeneity: Tau^2 = 4.13; Chi = 11.83, df = 5 (P = 0.04); I^2 = 58%

Test for overall effect: Z = 3.08 (P = 0.002)



- ✓ describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance)
- ✓ 50% threshold (50%-75% moderate heterogeneity, 75%-100% considerable heterogeneity)