SCUOLA DI METODOLOGIA CLINICA

IRCCS "Sacro Cuore - Don Calabria" Negrar di Valpolicella

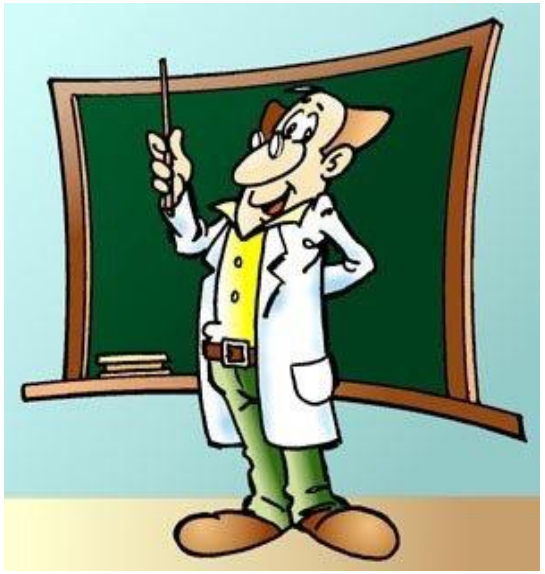
8ª EDIZIONE

1° MODULO

NEGRAR DI VALPOLICELLA
28-29 GENNAIO 2022

Centro Formazione IRCCS "Sacro Cuore-Don Calabria"


SCUOLA DI METODOLOGIA CLINICA
IRCCS "Sacro Cuore - Don Calabria" Negrar di Valpolicella
8ª EDIZIONE
1° MODULO
PRINCIPI DI SPERIMENTAZIONE CLINICA
NEGRAR DI VALPOLICELLA
28-29 GENNAIO 2022
Centro Formazione IRCCS "Sacro Cuore-Don Calabria"



28 Gennaio 2022

10.30-10.45 Saluti dell'amministratore delegato IRCCS
Sacro Cuore Don Calabria
Negrar di Valpolicella (VR)
Mario PICCININI

10.45-11.00 Presentazione ed obiettivi del Corso
Stefania GORI - Fabrizio NICOLIS
Giovanni L. PAPPAGALLO

11.00-12.00 Quesito clinico, plausibilità
e rilevanza del progetto, disegno di studio
Cristina MAZZI - Giovanni L. PAPPAGALLO

12.00-13.00 Lavoro di gruppo
Summing-Up (metodo: *What? So What? Now What?*)
(valido come prova ECM)

13.000-14.00 Colazione di lavoro

14.00-14.45 Scelta del braccio di controllo
e procedure di randomizzazione
Cristina MAZZI - Giovanni L. PAPPAGALLO

14.45-16.00 Variabili statistiche;
misure di effetto relative e assolute
Cristina MAZZI - Giovanni L. PAPPAGALLO

16.00-16.30 Coffee Break

16.30-17.30 Scelta dell'endpoint in base al quesito clinico
e al disegno dello studio
Cristina MAZZI - Giovanni L. PAPPAGALLO

17.30-18.30 Lavoro di gruppo
Summing-Up (metodo: *What? So What? Now What?*)
(valido come prova ECM)

29 Gennaio 2022

09.00-10.45 Principi di verifica di ipotesi statistica,
errori statistici, *target* di rilevanza clinica
Cristina MAZZI - Giovanni L. PAPPAGALLO

10.45-11.15 Coffee Break

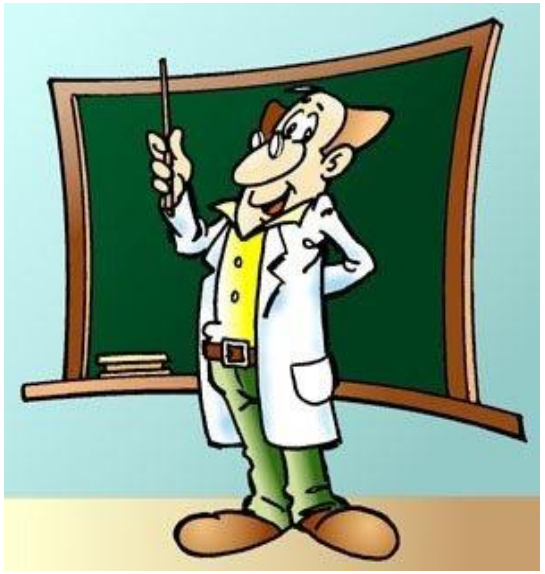
11.15-12.00 Calcolo della dimensione del campione
per diversi tipi di variabili statistiche
Cristina MAZZI - Giovanni L. PAPPAGALLO

12.00-13.00 Lavoro di gruppo
Summing-Up (metodo: *What? So What? Now What?*)
(valido come prova ECM)

13.00-13.15 Conclusione del Corso
Stefania GORI - Giovanni L. PAPPAGALLO

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi


SCUOLA DI METODOLOGIA CLINICA
IRCCS "Sacro Cuore - Don Calabria" Negrar di Valpolicella
8ª EDIZIONE
1° MODULO
PRINCIPI DI SPERIMENTAZIONE CLINICA
NEGRAR DI VALPOLICELLA
28-29 GENNAIO 2022
Centro Formazione IRCCS "Sacro Cuore-Don Calabria"




WHAT
SO WHAT
NOW WHAT

Liberating Structures® in sintesi

Favorire la creazione di reti tra Colleghi

Ricerca di innovazioni e soluzioni

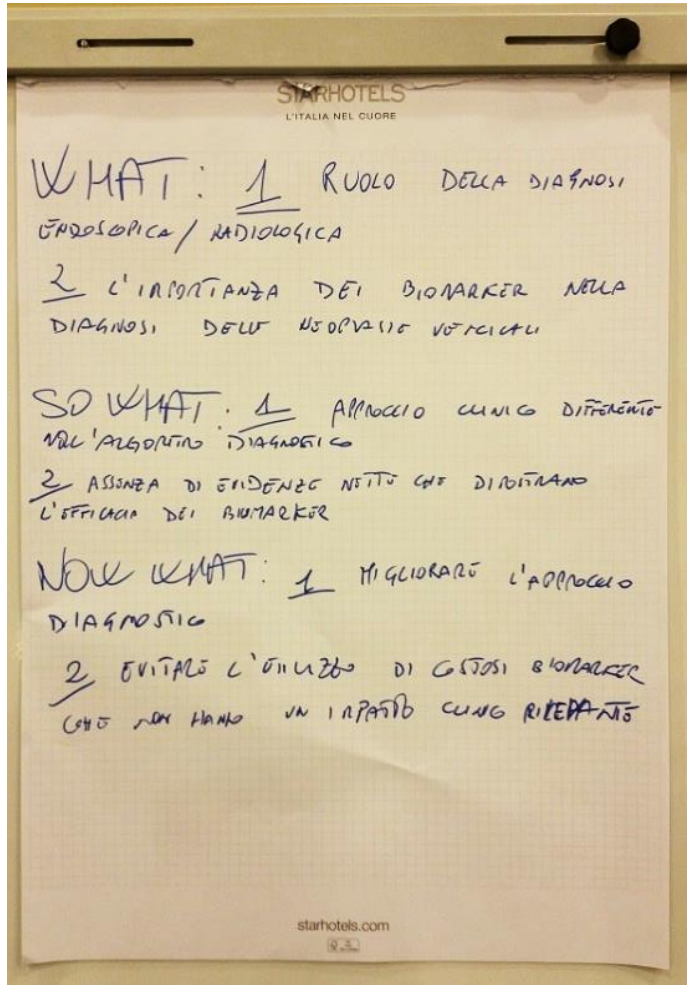
Analisi degli obiettivi

Condivisione di idee

Debriefing (riflessioni e sintesi)

Soluzioni a sfide

What?, So What?, Now What?



Partendo da quanto ascoltato, cosa è emerso di particolar-mente saliente / rilevante?

WHAT?

Per quale motivo le cose emerse sono così rilevanti?

SO WHAT?

Quali ricadute nell'immediato per la mia professione?

NOW WHAT?

28 Gennaio 2022

10.30-10.45 Saluti dell'amministratore delegato IRCCS
Sacro Cuore Don Calabria
Negrar di Valpolicella (VR)
Mario PICCININI

10.45-11.00 Presentazione ed obiettivi del Corso
Stefania GORI - Fabrizio NICOLIS
Giovanni L. PAPPAGALLO

11.00-12.00 Quesito clinico, plausibilità
e rilevanza del progetto, disegno di studio
Cristina MAZZI - Giovanni L. PAPPAGALLO

12.00-13.00 Lavoro di gruppo
Summing-Up (metodo: *What? So What? Now What?*)
(valido come prova ECM)

13.000-14.00 Colazione di lavoro

14.00-14.45 Scelta del braccio di controllo
e procedure di randomizzazione
Cristina MAZZI - Giovanni L. PAPPAGALLO

14.45-16.00 Variabili statistiche;
misure di effetto relative e assolute
Cristina MAZZI - Giovanni L. PAPPAGALLO

16.00-16.30 Coffee Break

16.30-17.30 Scelta dell'endpoint in base al quesito clinico
e al disegno dello studio
Cristina MAZZI - Giovanni L. PAPPAGALLO

17.30-18.30 Lavoro di gruppo
Summing-Up (metodo: *What? So What? Now What?*)
(valido come prova ECM)

29 Gennaio 2022

09.00-10.45 Principi di verifica di ipotesi statistica,
errori statistici, *target* di rilevanza clinica
Cristina MAZZI - Giovanni L. PAPPAGALLO

10.45-11.15 Coffee Break

11.15-12.00 Calcolo della dimensione del campione
per diversi tipi di variabili statistiche
Cristina MAZZI - Giovanni L. PAPPAGALLO

12.00-13.00 Lavoro di gruppo
Summing-Up (metodo: *What? So What? Now What?*)
(valido come prova ECM)

13.00-13.15 Conclusione del Corso
Stefania GORI - Giovanni L. PAPPAGALLO

- **L'approccio alla lettura**
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

Fattori da prendere in considerazione come primo approccio alla lettura di un articolo riguardante una sperimentazione clinica

① **Gravità dell'affezione.**

② **Efficacia/tollerabilità delle terapie disponibili.**

③ ***Unmet need* attuali (efficacia e/o tollerabilità).**

④ **Presumibile superiorità delle terapie sperimentali.**

Bruzzi & Pappagallo, 1987
mod. Pappagallo 2020

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial

Naomi B Haas, Judith Manola, Robert G Uzzo, Keith T Flaherty, Christopher G Wood, Christopher Kane, Michael Jewett, Janice P Dutcher, Michael B Atkins, Michael Pins, George Wilding, David Cella, Lynne Wagner, Surena Matin, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Toni K Choueiri, Roberto Pili, Igor Puzanov, Manish Kohli, Walter Stadler, Michael Carducci, Robert Coomes*, Robert S DiPaola

Lancet 2016; 387: 2008–16

Roughly 40% of the 338 000 people diagnosed worldwide with primary renal-cell carcinoma each year will ultimately die from metastatic disease. Risk of disease recurrence can be estimated based on algorithms ① incorporating clinical and histological features with TNM staging. Advances in genomic, epigenetic, and expression profiling, which can refine estimates of recurrence risk, are not available for clinical decision making.

The standard of care for primary resected renal-cell carcinoma at risk for recurrence has historically been observation. Previous adjuvant trials ② of interleukin 2, hormonal therapy, or chemotherapy in patients with high-③ risk resected disease have all been negative. Trials of adjuvant interferon, with potentially anti-angiogenic or immuno-stimulatory effects, were also negative.

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial

Naomi B Haas, Judith Manola, Robert G Uzzo, Keith T Flaherty, Christopher G Wood, Christopher Kane, Michael Jewett, Janice P Dutcher, Michael B Atkins, Michael Pins, George Wilding, David Cella, Lynne Wagner, Surena Matin, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Toni K Choueiri, Roberto Pili, Igor Puzanov, Manish Kohli, Walter Stadler, Michael Carducci, Robert Coomes*, Robert S DiPaola

Lancet 2016; 387: 2008–16

The vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitors (VEGFR inhibitors) sunitinib and sorafenib extend the progression-free survival of patients with advanced disease.

④ Extensive testing of VEGFR inhibitors in the adjuvant setting is underway, and these agents are being used in the community in this clinical setting. This study is the first randomised trial to compare disease-free survival with adjuvant sorafenib or sunitinib versus placebo in patients with resected primary renal-cell carcinoma at high risk for recurrence.

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

① Approximately 20% of all breast cancers have gene amplification or overexpression (or both) of human epidermal growth factor receptor 2 (HER2),[1] a tyrosine kinase transmembrane receptor, resulting in a more aggressive phenotype and a poor prognosis.

② Treatment with the anti-HER2 humanized monoclonal antibody trastuzumab in addition to chemotherapy, as compared with chemotherapy alone, significantly improves progression-free and overall survival among patients with HER2-positive metastatic breast cancer.

③ However, in most patients with HER2-positive metastatic breast cancer, the disease progresses,[8] highlighting the need for new targeted therapies for advanced disease.

Pertuzumab prevents HER2 from dimerizing with other ligand-activated HER receptors, most notably HER3.

④ Because pertuzumab and trastuzumab bind to different HER2 epitopes and have complementary mechanisms of action, these two agents, when given together, provide a more comprehensive blockade of HER2 signaling and result in greater antitumor activity than either agent alone in HER2-positive tumor models.

The Clinical Evaluation of Pertuzumab and Trastuzumab (CLEOPATRA) study assessed the efficacy and safety of pertuzumab plus trastuzumab plus docetaxel, as compared with placebo plus trastuzumab plus docetaxel, as first-line treatment for patients with HER2-positive metastatic breast cancer.

Oral Apixaban for the Treatment of Acute Venous Thromboembolism

Giancarlo Agnelli, M.D., Harry R. Buller, M.D., Ph.D., Alexander Cohen, M.D., Madelyn Curto, D.V.M., Alexander S. Gallus, M.D., Margot Johnson, M.D., Urszula Masiukiewicz, M.D., Raphael Pak, Ph.D., John Thompson, Ph.D., Gary E. Raskob, Ph.D., and Jeffrey I. Weitz, M.D., for the AMPLIFY Investigators*

N Engl J Med 2013;369:799-808.

① Venous thromboembolism, with an annual incidence of 1 to 2 cases per 1000 persons in the general population, is the third most common cause of vascular death after myocardial infarction and stroke.[1] Conventional treatment consists of a parenteral anticoagulant, ② such as enoxaparin, for at least 5 days, and warfarin begun during this time and continued for at least 3 months.[2] Although effective, this regimen ③ presents a challenge because enoxaparin requires daily subcutaneous injections, and warfarin therapy requires coagulation monitoring and dose adjustment.

Apixaban may simplify the treatment of venous thromboembolism by eliminating the need for initial parenteral anticoagulant therapy and laboratory monitoring, a concept supported by recent studies.

④ In the Apixaban for the Initial Management of Pulmonary Embolism and Deep-Vein Thrombosis as First-Line Therapy (AMPLIFY) trial, we compared apixaban with conventional anticoagulant therapy in patients with acute symptomatic venous thromboembolism.

Effectiveness and Safety of Tapentadol Prolonged Release (PR) Versus a Combination of Tapentadol PR and Pregabalin for the Management of Severe, Chronic Low Back Pain With a Neuropathic Component: A Randomized, Double-blind, Phase 3b Study

Ralf Baron, MD, PhD*; Emilio Martin-Mola, MD, PhD†; Matthias Müller, MD, MSc‡; Cecile Dubois, MS§; Dietmar Falke, PhD‡; Ilona Steigerwald, MD‡

Pain Practice, Volume ●●, Issue ●, 2014 ●●–●●

Chronic low back pain is a common chronic pain condition[1–4] that often has a neuropathic pain component, which may complicate its management.

Descending noradrenergic modulation mechanisms appear to play an important role in neuropathic pain modulation.[6] Therefore, the use of monotherapy directed at the ascending pathways or specific individual targets (eg, μ-opioid receptor [MOR] agonists) may not fully address the neuropathic component of low back pain.

To manage severe chronic pain with a neuropathic component, patients may be treated with a combination of strong opioids with co-analgesics (eg, anticonvulsants, antidepressants).[9] However, combination therapy may be associated with a higher incidence of side effects and related discontinuations.

Tapentadol represents a new class of centrally acting analgesic with both MOR agonist and noradrenaline reuptake inhibitor (NRI) activities.

Based on the MOR-NRI concept, which may address both nociceptive and neuropathic mechanisms of chronic pain, and previous results indicating that tapentadol PR is effective for managing chronic pain with a neuropathic component,[16,19] it was reasonable to examine whether tapentadol PR monotherapy could be as effective as a combination of tapentadol PR and pregabalin and to compare the tolerability profiles of both analgesic options.

Effect of a monoclonal antibody to PCSK9, REGN727/ SAR236553, to reduce low-density lipoprotein cholesterol in patients with heterozygous familial hypercholesterolaemia on stable statin dose with or without ezetimibe therapy: a phase 2 randomised controlled trial

Evan A Stein, Dan Gipe, Jean Bergeron, Daniel Gaudet, Robert Weiss, Robert Dufour, Richard Wu, Robert Pordy

Lancet 2012; 380: 29–36

① Heterozygous familial hypercholesterolaemia, a common genetic disorder in man with a minimum prevalence of one in 500 across all populations, is associated with substantial morbidity and mortality.[1]

② Substantial improvements in the treatment of familial hypercholesterolaemia have been achieved over the last 25 years with statins, ezetimibe, newer bile acid sequestrants, and extended-release niacin formulations.

③ An unmet medical need for additional, effective LDL-C lowering therapies remains for many patients with heterozygous familial hypercholesterolaemia.

The Treat-to-Target Trial

Randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients

MATTHEW C. RIDDLE, MD[1] JOHN GERICH, MD[2]
JULIO ROSENSTOCK, MD[2]
ON BEHALF OF THE INSULIN GLARGINE 4002 STUDY INVESTIGATORS*

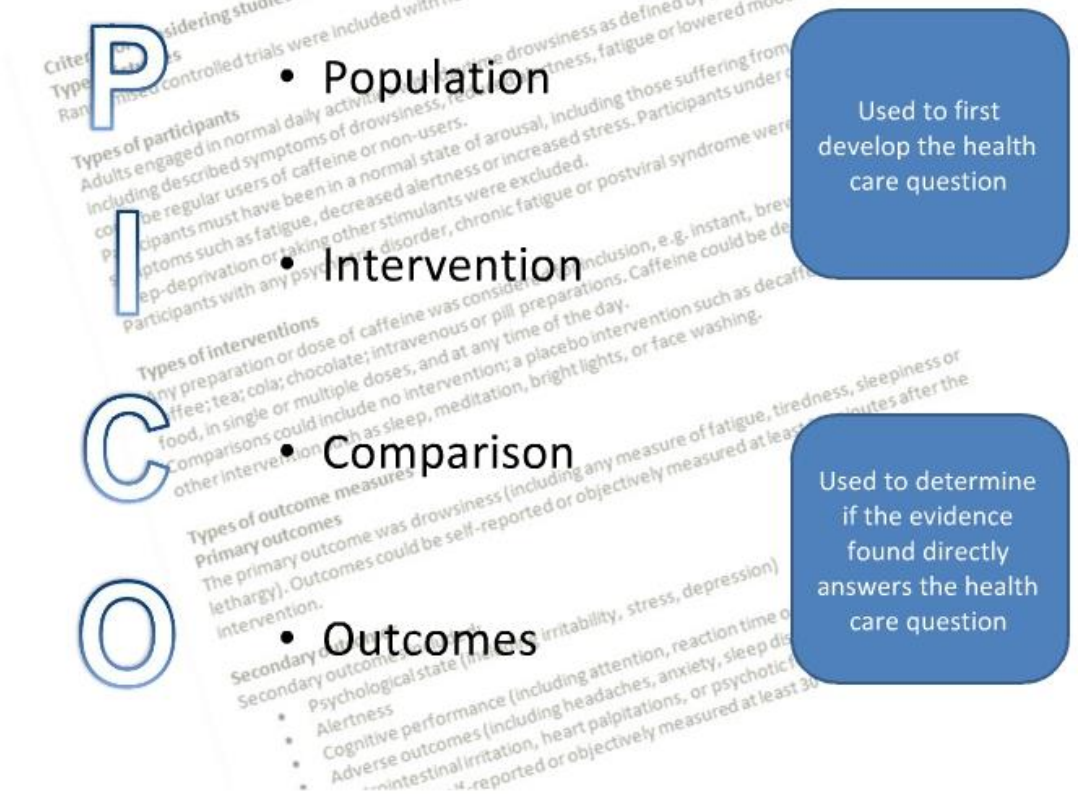
DIABETES CARE, VOLUME 26, NUMBER 11, NOVEMBER 2003

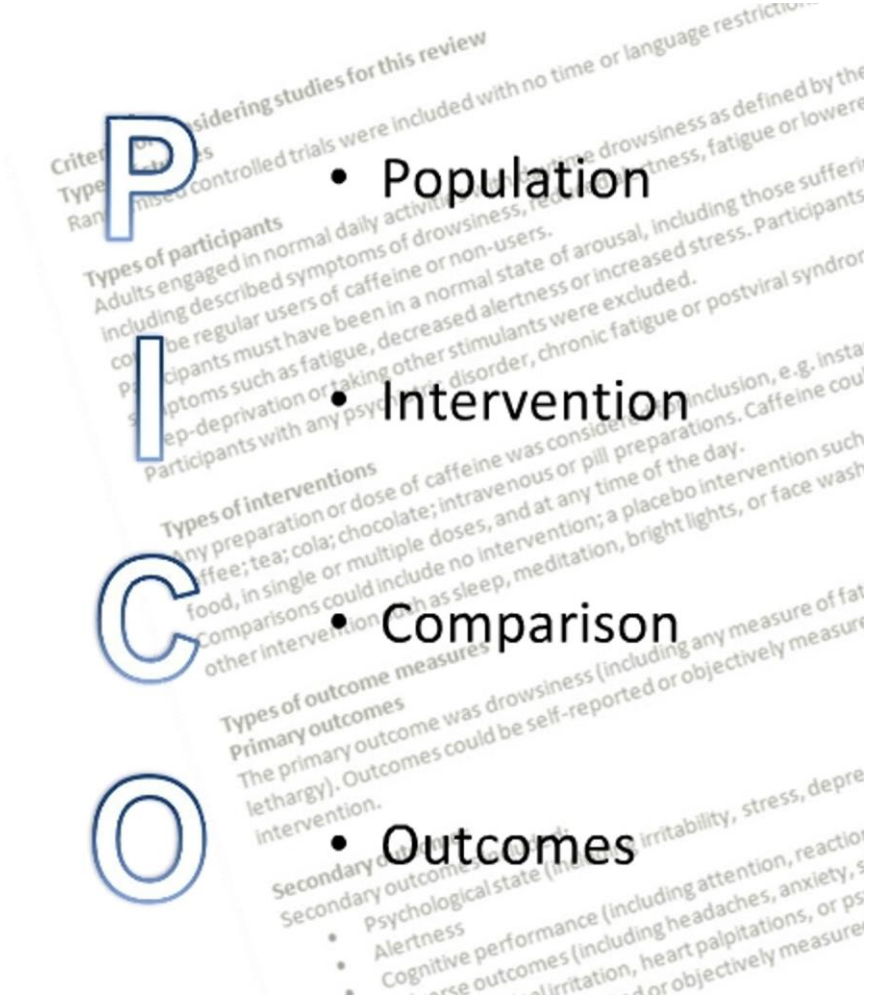
① The majority of patients with a longer duration of diabetes remain poorly controlled with oral agents, and use of insulin, which could improve glycemic control, is often long delayed and not aggressive enough. The reluctance to ② initiate insulin therapy seems partly due to its perceived complexity, the belief that ③ insulin is not effective for type 2 diabetes (8), and fear of hypoglycemia, which may be the greatest barrier (9).

④ Glargine, a new long-acting insulin analog with a more favorable 24-h time-action profile (no pronounced peak) than long- or intermediate-acting human insulin preparations (13,14), may be especially suited to this regimen.

- L'approccio alla lettura
- **Il quesito clinico**
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

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





In un paziente con neoplasia a cellule renali, sotto-posto a nefrectomia, ma ad alto rischio di recidiva,

… un trattamento medico adiuvante,

… è proponibile in alternativa al solo follow-up,

… con riguardo alla sopravvivenza libera da recidiva, sopravvivenza globale e tollerabilità?

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial

Naomi B Haas, Judith Manola, Robert G Uzzo, Keith T Flaherty, Christopher G Wood, Christopher Kane, Michael Jewett, Janice P Dutcher, Michael B Atkins, Michael Pins, George Wilding, David Cella, Lynne Wagner, Surena Matin, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Toni K Choueiri, Roberto Pili, Igor Puzanov, Manish Kohli, Walter Stadler, Michael Carducci, Robert Coomes*, Robert S DiPaola

Lancet 2016; 387: 2008–16

The vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitors (VEGFR inhibitors) sunitinib and sorafenib extend the progression-free survival of patients with advanced disease.

Extensive testing of VEGFR inhibitors in the adjuvant setting is underway, and these agents are being used in the community in this clinical setting. This study is the first randomised trial to compare disease-free survival with adjuvant sorafenib or sunitinib versus placebo in patients with resected primary renal-cell carcinoma at high risk for recurrence.

Population

Intervention

Comparator

Outcome

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial

Naomi B Haas, Judith Manola, Robert G Uzzo, Keith T Flaherty, Christopher G Wood, Christopher Kane, Michael Jewett, Janice P Dutcher, Michael B Atkins, Michael Pins, George Wilding, David Cella, Lynne Wagner, Surena Matin, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Toni K Choueiri, Roberto Pili, Igor Puzanov, Manish Kohli, Walter Stadler, Michael Carducci, Robert Coomes*, Robert S DiPaola

Lancet 2016; 387: 2008–16

The vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitors (VEGFR inhibitors) sunitinib and sorafenib extend the progression-free survival of patients with advanced disease.
Extensive testing of VEGFR inhibitors in the adjuvant setting is underway, and these agents are being used in the community in this clinical setting. This study is the first randomised trial to compare disease-free survival with adjuvant sorafenib or sunitinib versus placebo in patients with resected primary renal-cell carcinoma at high risk for recurrence.

Population

Intervention

Comparator

Outcome

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial

Naomi B Haas, Judith Manola, Robert G Uzzo, Keith T Flaherty, Christopher G Wood, Christopher Kane, Michael Jewett, Janice P Dutcher, Michael B Atkins, Michael Pins, George Wilding, David Cella, Lynne Wagner, Surena Matin, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Toni K Choueiri, Roberto Pili, Igor Puzanov, Manish Kohli, Walter Stadler, Michael Carducci, Robert Coomes*, Robert S DiPaola

Lancet 2016; 387: 2008–16

The vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitors (VEGFR inhibitors) sunitinib and sorafenib extend the progression-free survival of patients with advanced disease.
Extensive testing of VEGFR inhibitors in the adjuvant setting is underway, and these agents are being used in the community in this clinical setting. This study is the first randomised trial to compare disease-free survival with adjuvant sorafenib or sunitinib versus placebo in patients with resected primary renal-cell carcinoma at high risk for recurrence.

Population
Intervention
Comparator
Outcome

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial

Naomi B Haas, Judith Manola, Robert G Uzzo, Keith T Flaherty, Christopher G Wood, Christopher Kane, Michael Jewett, Janice P Dutcher, Michael B Atkins, Michael Pins, George Wilding, David Cella, Lynne Wagner, Surena Matin, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Toni K Choueiri, Roberto Pili, Igor Puzanov, Manish Kohli, Walter Stadler, Michael Carducci, Robert Coomes*, Robert S DiPaola

Lancet 2016; 387: 2008–16

The vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitors (VEGFR inhibitors) sunitinib and sorafenib extend the progression-free survival of patients with advanced disease.

Extensive testing of VEGFR inhibitors in the adjuvant setting is underway, and these agents are being used in the community in this clinical setting. This study is the first randomised trial to compare disease-free survival with adjuvant sorafenib or sunitinib versus placebo in patients with resected primary renal-cell carcinoma at high risk for recurrence.

Population

Intervention

Comparator

Outcome

Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): ndomised, phase 3 trial

opher Kane, Michael Jewett, Janice P Dutcher, Timothy M Kuzel, Wade J Sexton, Yu-Ning Wong, Robert Coomes*, Robert S DiPaola

r receptor (VEGFR) tyrosine
itinib and sorafenib extend the
advanced disease.
rs in the adjuvant setting is
sed in the community in this
randomised trial to compare
nib or sunitinib versus placebo
cell carcinoma at high risk for

Outcomes

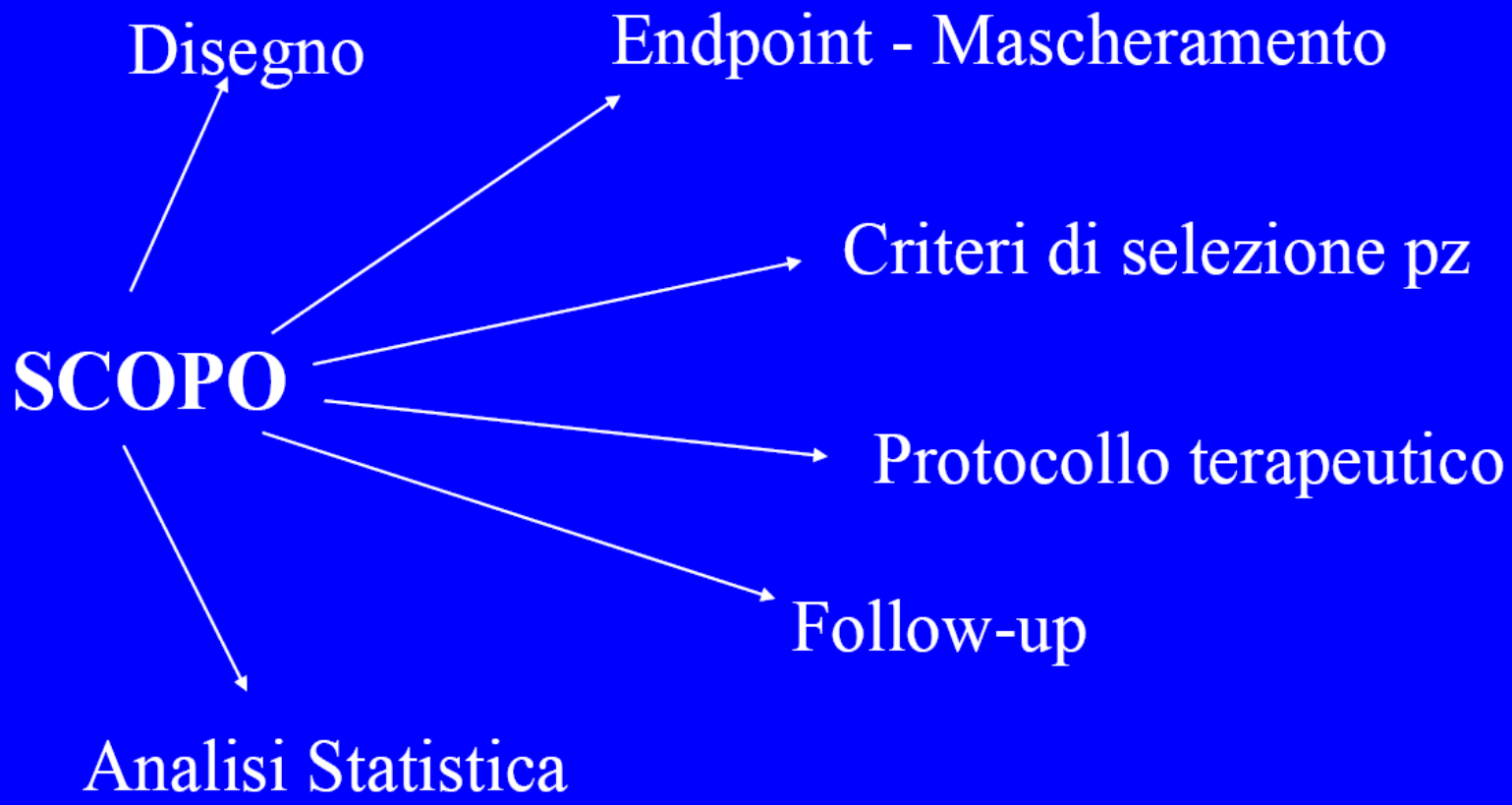
The primary endpoint was disease-free survival, defined as the time from randomisation to recurrence, development of second primary cancer, or death from any cause. Patients alive without disease recurrence at the time of analysis were censored on the date of last disease evaluation. Secondary endpoints included overall survival, disease-free survival for clear cell renal cancer, and toxic effects as assessed by NCI CTCAE version 3.0. The effect of treatment on cardiac function has been reported previously.[11] Data about a prospective quality-of-life assessment for fatigue were collected (patient-reported quality of life and assessment of a new measure of quality of life, PROMIS Fatigue-SF1) and will be published separately. A final secondary objective was to prospectively collect tumour and biological specimens to assess their characteristics and associations with various angiogenesis markers, genetic mutations and polymorphisms, DNA methylation profiles, drug metabolising, enzymes, and circulating endothelial cells, which will be reported in a later publication.

n
on
parator

Outcome

- L'approccio alla lettura
- Il quesito clinico
- **L'obiettivo dello studio**
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

L'aspetto piu' importante di uno studio e' il suo scopo primario



P. Bruzzi, qualche tempo fa…

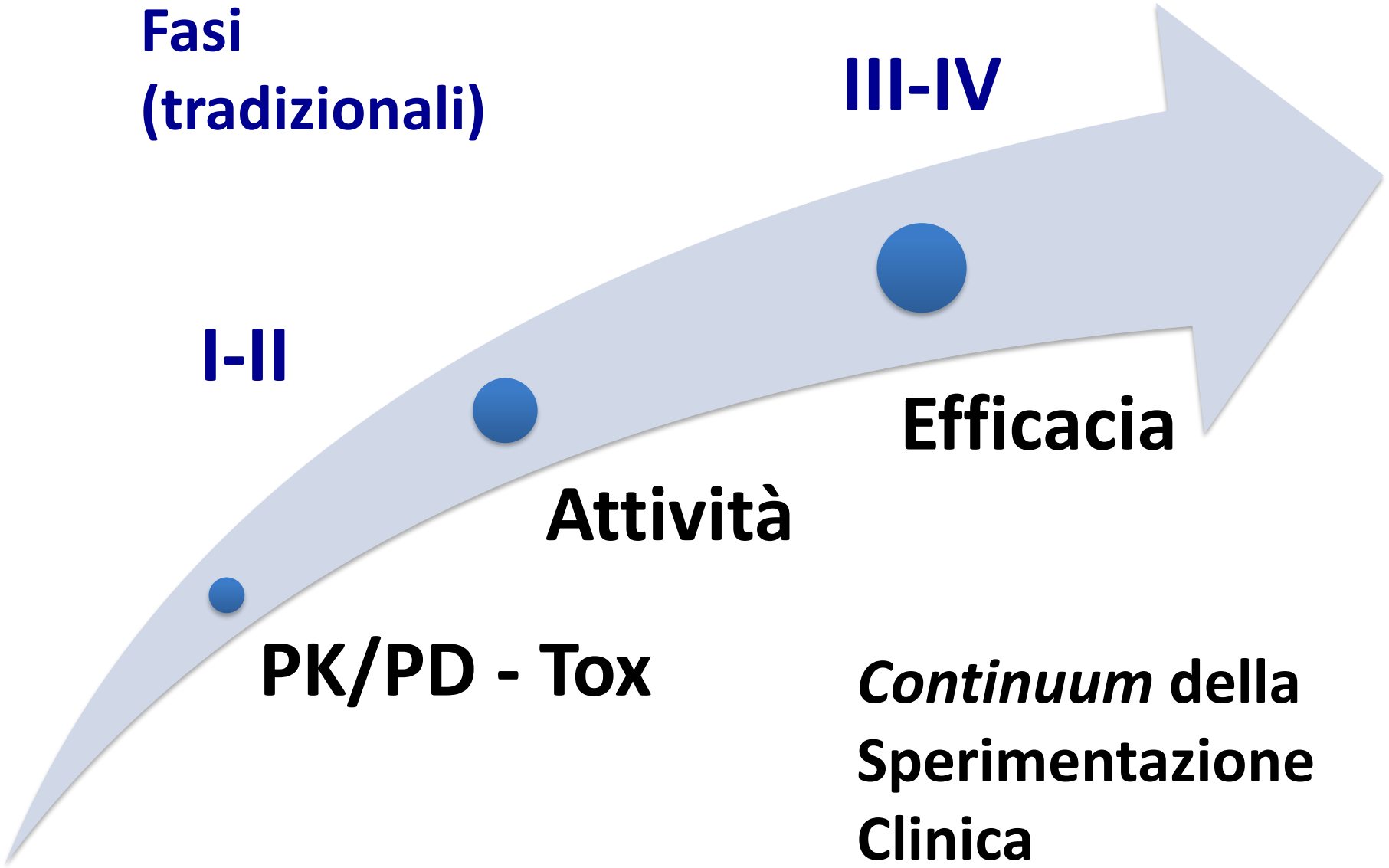
Hypotheses and Objectives

- KISS – keep it simple, stupid
- Too many objectives compromise a trial
  - A single hypothesis and a few secondary hypotheses
  - Can't study everything
- If you can't power an endpoint, it shouldn't be a primary or secondary objective

Primary and Secondary Questions/Objectives

- Common error – Sinking ship: Avoid overloading the study with too many objectives and too much data collection

- A single primary question around which to focus the development of the protocol and sample size estimates
- Secondary research questions: can be related to the primary question or to other hypotheses


Fasi
(tradizionali)
III-IV
I-II
Efficacia
Attività
PK/PD - Tox
Continuum della
Sperimentazione
Clinica

Attività *vs* Efficacia

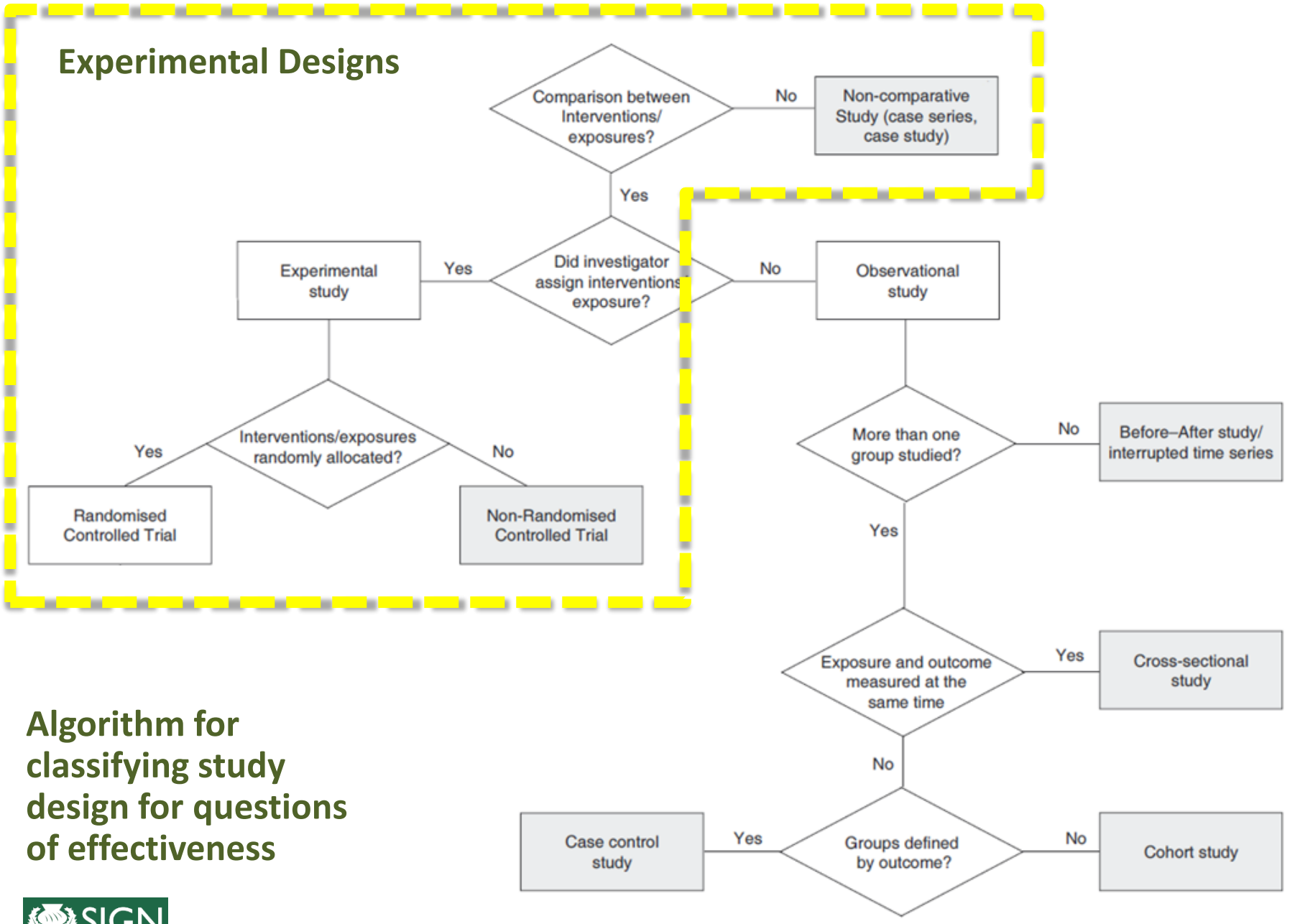
- **Attività**
  - **capacità di un trattamento di indurre le modificazioni attraverso le quali *si presume* di indurre dei benefici**
- **Efficacia**
  - **capacità di un trattamento di indurre i benefici per ottenere i quali esso viene somministrato**

Attività *vs* Efficacia

trattamento	*attività*	*efficacia*
diuretico	riduzione P.A.	riduzione malatt. C.V.
antidiab. orale	riduz. glicemia	riduz. mortalità
a.infiammat.	az. a.aggregante	riduzione malatt. C.V.
citotossico	riduz. tumorale	riduz. mortalità
citostatico	controllo malattia	riduz. mortalità
fatt. di crescita	stimolo crescita	riduz. complicanze

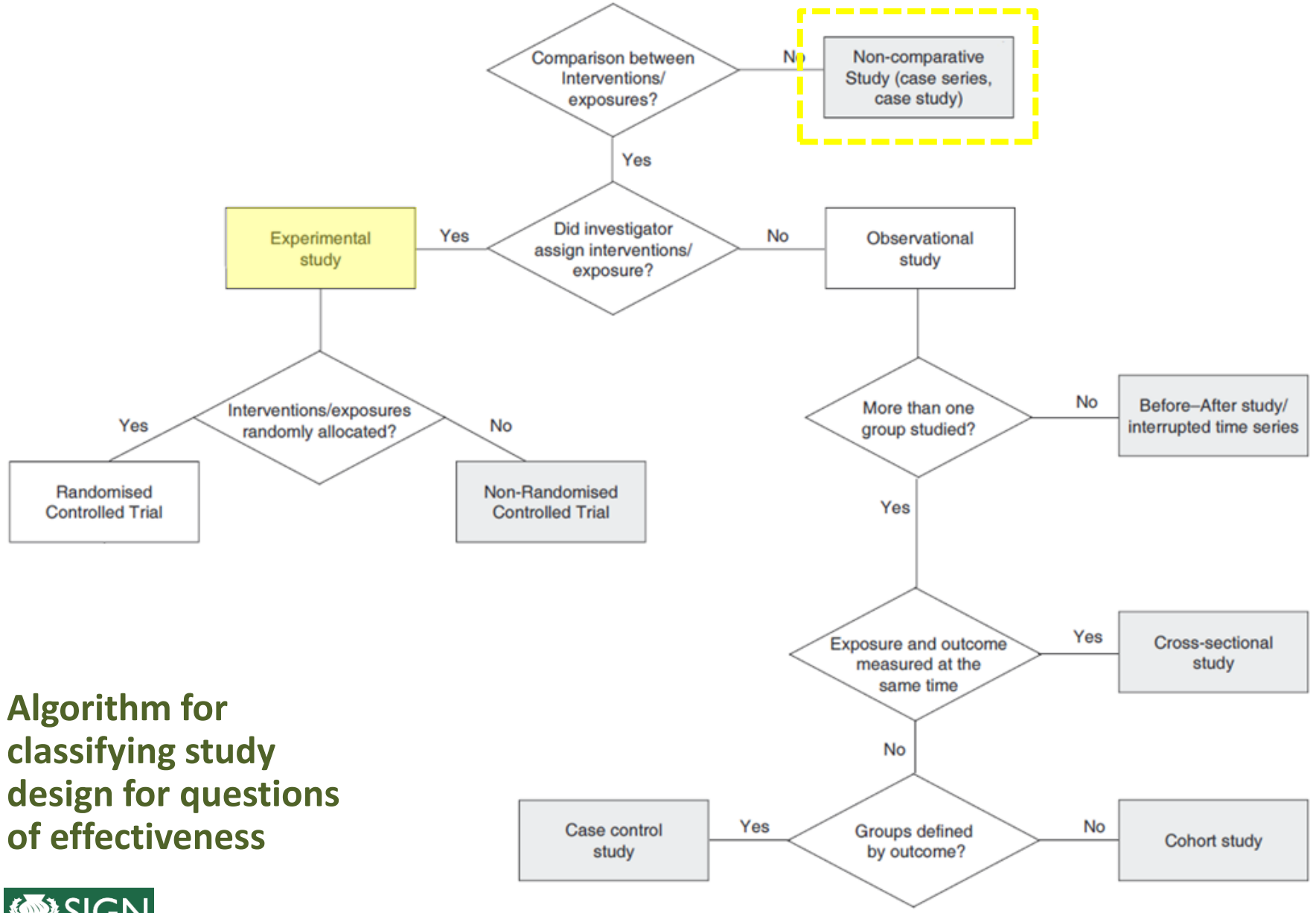
- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- **Il disegno dello studio: studi sperimentali**
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

Experimental Designs



Algorithm for classifying study design for questions of effectiveness

SIGN
Scottish Intercollegiate Guidelines Network



Algorithm for classifying study design for questions of effectiveness

SIGN
Scottish Intercollegiate Guidelines Network

Urologia Internationalis

Urol Int 1998;61:220–226

Received: March 18, 1998
Accepted after revision: November 11, 1998

R. Hurle[a]
A. Manzetti[a]
A. Losa[a]
E. Micheli[a]
A. Ranieri[a]
D. Chinaglia[b]
A. Lembo[a]

Intravesical Instillation of Mitomycin-C in 242 Patients with Superficial Bladder Cancer at High Risk of Recurrence: Long-Term Results

Intravesical chemotherapy has been used for almost three decades with the aim of reducing recurrence and progression rates. Several drugs have given encouraging results in the prevention of recurrences [3–6] but uncertainty persists about the effect of prophylactic chemotherapy on disease progression [7, 8]. Mitomycin C (MMC) significantly reduces the recurrence rate of superficial bladder cancer [6, 9].

The present study assessed the long-term results of intravesical MMC instillation after TUR in 242 consecutive patients with superficial bladder cancer at high risk of recurrence.

Noncomparative (old-style) efficacy study

Disegno di studi di fase II

Random o non-random?

- **Studi di fase II non randomizzati (a singolo braccio)** *(Fleming-Simon)*
- **Randomized, Discontinuation Design**
- **Randomized, Selection Design**
- **Randomized, Screening Design**

Salvage Therapy with Capecitabine Plus Weekly Paclitaxel in Heavily Pretreated Advanced Breast Cancer

A Multicenter Phase II Study

Mario Bari,[1] Mario Rosario D'Andrea,[1] Giuseppe Azzarello,[1] Giovanni L. Pappagallo,[1] Donata Sartori,[1] Aldo Iop,[2] Ferdinando Gaion,[3] Francesco Rosetti,[1] Barbara Silvestri,[1] Salvatore Bonura,[2] Antonietta D'Alessio[3] and Orazio Vinante[1]

Am J Cancer 2005; 4 (5): 307-313

...we planned to test the null hypothesis that the true response rate was <25% (i.e. no clinical interest) against the alternative hypothesis that the true response rate was at least 40% (level of clinical interest), with $\alpha = 0.05$ and $1\text{-}\beta = 90\%$. Thus, according to Simon's 'optimal design',[17] 20 patients had to be enrolled, with an upper limit for first stage rejection of the null hypothesis of four responses; the planned maximum sample size was 49 patients (first plus second stage rejection), with an upper limit for second stage rejection of 14 responses.

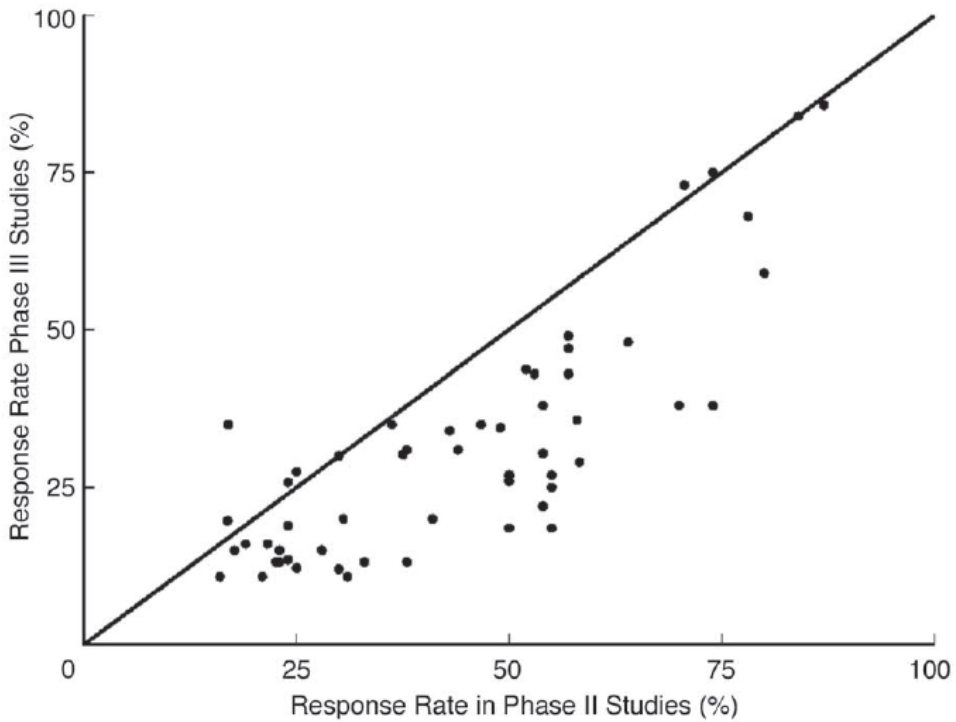
VOLUME 23 · NUMBER 28 · OCTOBER 1 2005

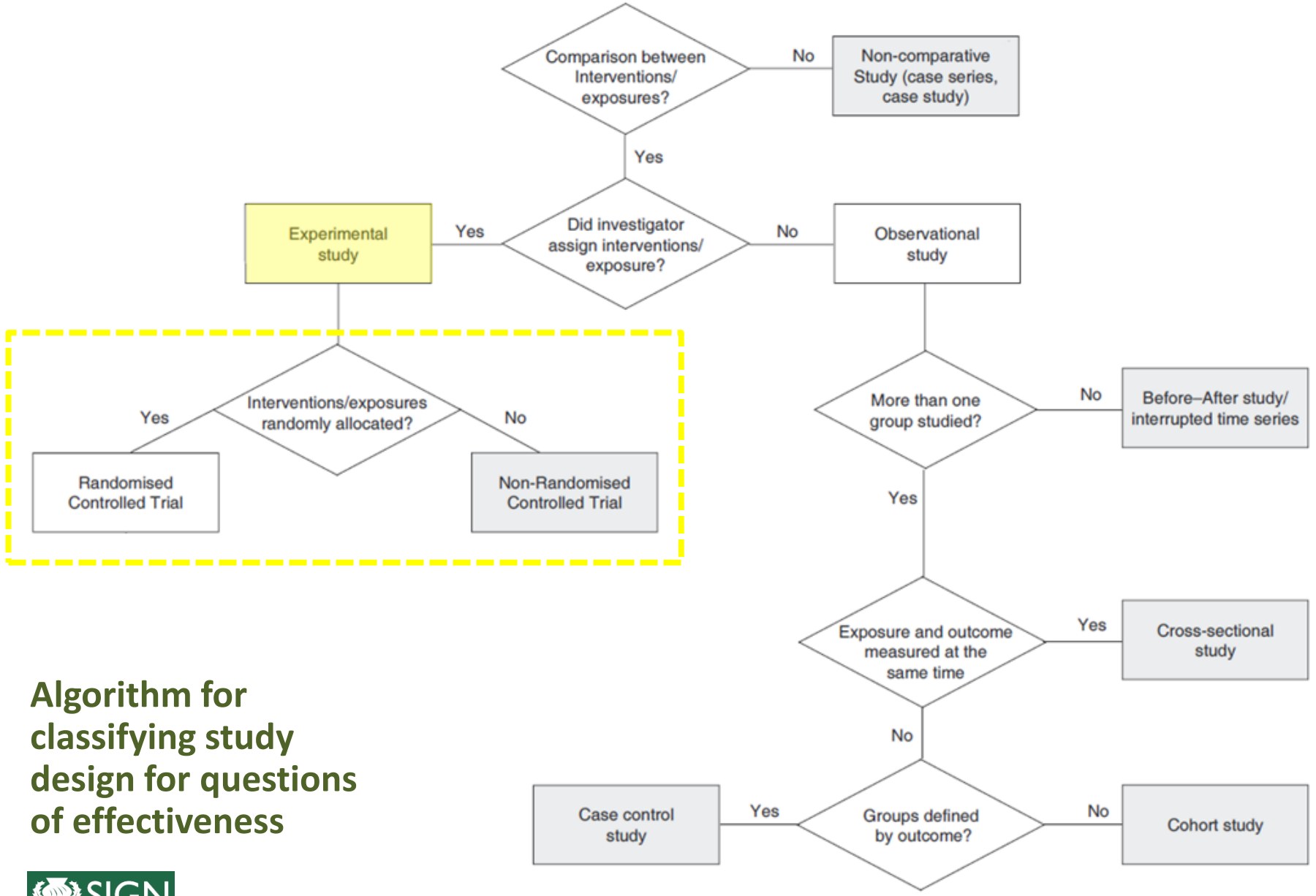
JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Comparison of Outcomes of Phase II Studies and Subsequent Randomized Control Studies Using Identical Chemotherapeutic Regimens

Mohammad I. Zia, Lillian L. Siu, Greg R. Pond, and Eric X. Chen





Algorithm for classifying study design for questions of effectiveness

SIGN
Scottish Intercollegiate Guidelines Network

Disegno di studi di fase II

Random o non-random?

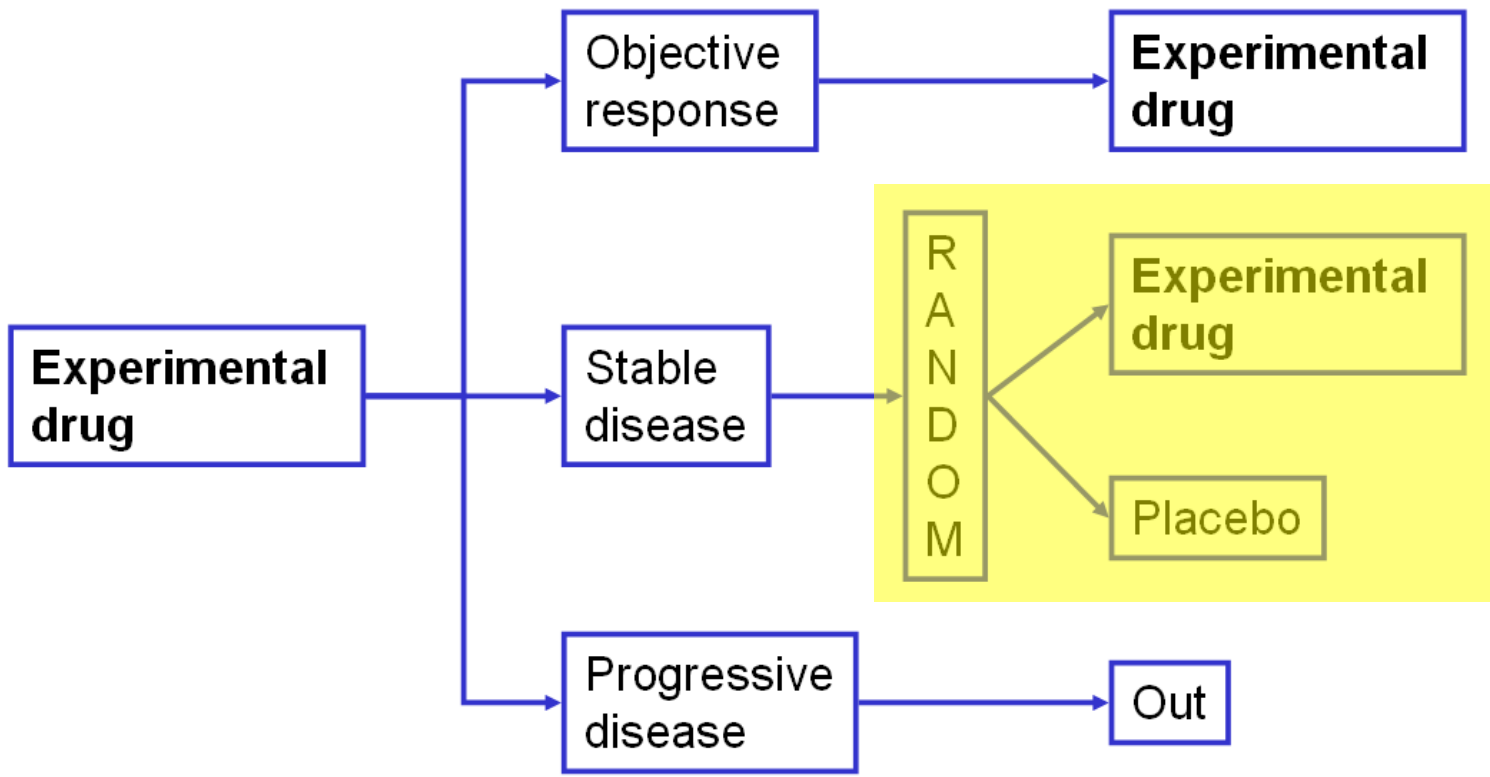
- **Studi di fase II non randomizzati (a singolo braccio)** *(Fleming-Simon)*
- **Randomized, Discontinuation Design**
- **Randomized, Selection Design**
- **Randomized, Screening Design**

Randomized Discontinuation Design: Application to Cytostatic Antineoplastic Agents

By Gary L. Rosner, Walter Stadler, and Mark J. Ratain

J Clin Oncol 20:4478-4484. © *2002*

Results: **By selecting a more homogeneous population, the randomized portion of the study requires fewer patients than would a study randomizing all patients at entry.**



Disegno di studi di fase II

Random o non-random?

- **Studi di fase II non randomizzati (a singolo braccio)** *(Fleming-Simon*
- **Randomized, Discontinuation Design**
- **Randomized, Selection Design**
- **Randomized, Screening Design**

Randomized Phase II Selection Designs

- K experimental arms, no control arm
- Select arm with highest response rate or disease control rate further development

- Simon, Wittes, Ellenberg; Cancer Treatment Reports 69:1375, 1985

R. Simon, ASCO 2010

Disegno di studi di fase II

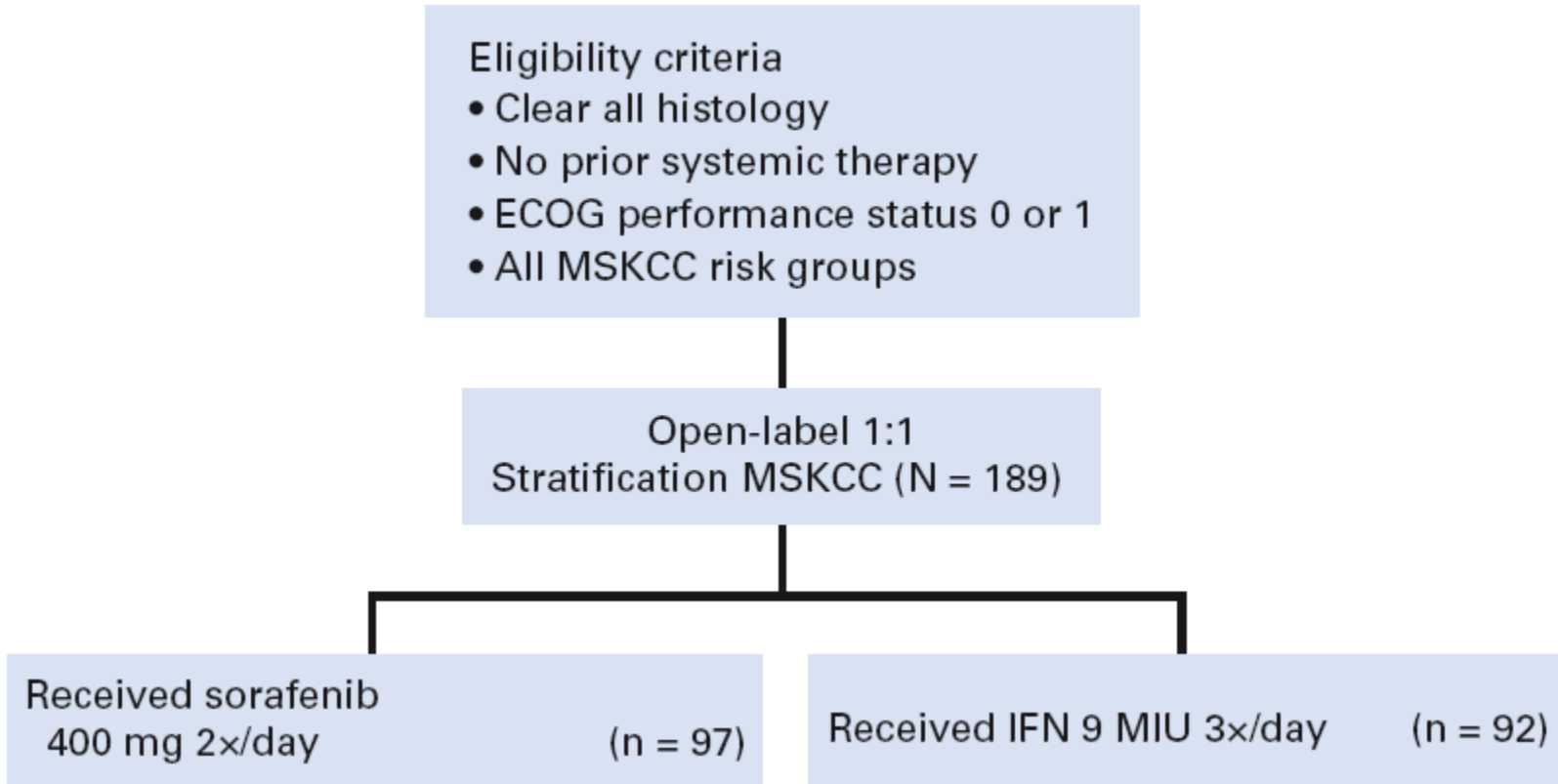
Random o non-random?

- Studi di fase II non randomizzati (a singolo braccio) *(Fleming-Simon*
- Randomized, Discontinuation Design
- Randomized, Selection Design
- **Randomized, Screening Design**

Randomized Phase II Trial of First-Line Treatment With Sorafenib Versus Interferon Alfa-2a in Patients With Metastatic Renal Cell Carcinoma

Bernard Escudier, Cezary Szczylik, Thomas E. Hutson, Tomasz Demkow, Michael Staehler, Frédéric Rolland, Sylvie Negrier, Nicole Laferriere, Urban J. Scheuring, David Cella, Sonalee Shah, and Ronald M. Bukowski

J Clin Oncol 27:1280-1289.



Design Issues of Randomized Phase II Trials and a Proposal for Phase II Screening Trials

Lawrence V. Rubinstein, Edward L. Korn, Boris Freidlin, Sally Hunsberger, S. Percy Ivy, and Malcolm A. Smith

RANDOMIZED PHASE II SCREENING DESIGNS

The most important caveat in using the phase II screening design is that it may compromise the ability to conduct definitive phase III trials. The screening design should not be applied unless investigators can be reasonably certain that a positive result in their small study will not be appreciated as definitive and will not preclude conduct of a definitive phase III test of the experimental regimen.

Superiorità Vs Non-inferiorità

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

eriorità

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

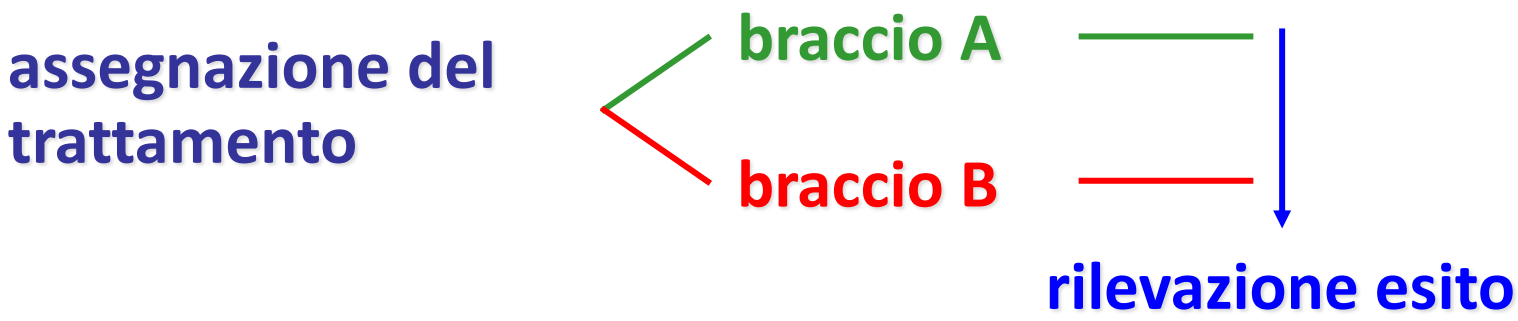
A > B di una quantità Δ di interesse clinico

studio di non inferiorità

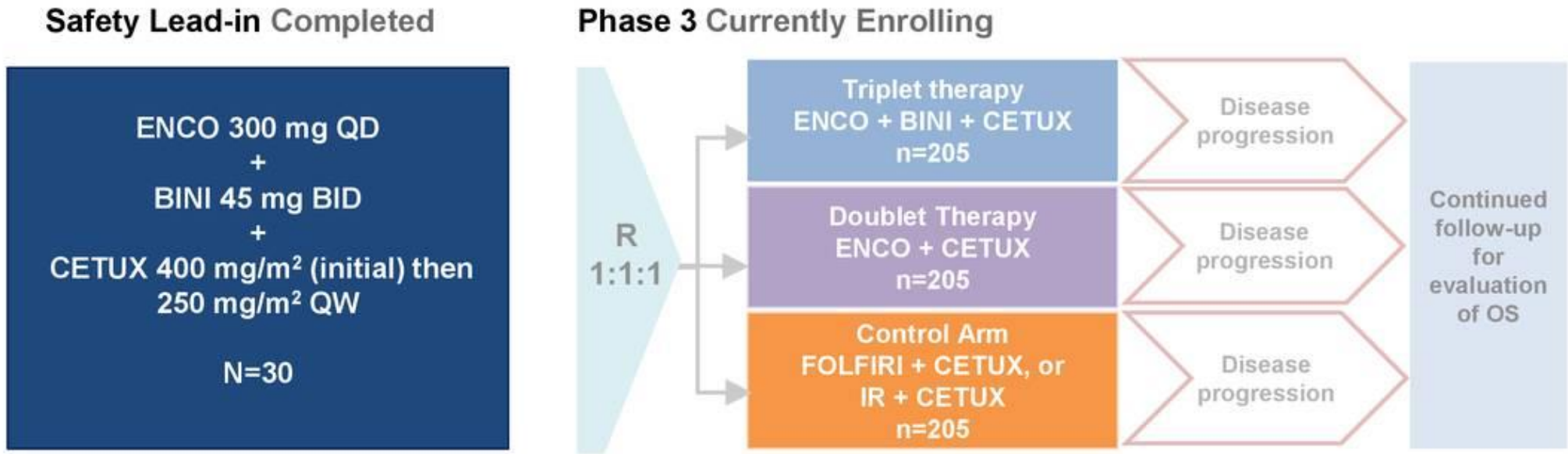
A < B non oltre una quantità M di rilevanza clinica

DISEGNO A BRACCI PARALLELI

Assegnazione del paziente a un gruppo di trattamento, al quale si appartiene per l'intera durata dello studio:



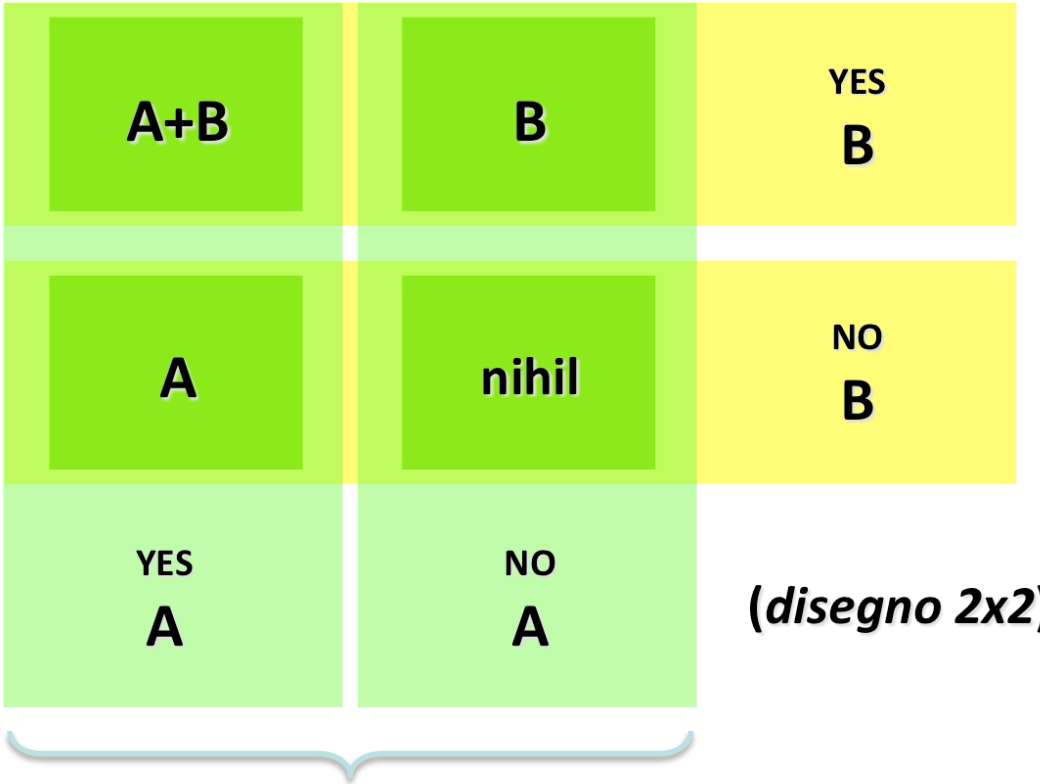
BEACON CRC Phase 3 Study Design[1]



1. Clinicaltrials.gov/ct2/show/NCT02928224; https://clinicaltrials.gov/ct2/show/NCT02928224 (February 2018).

FOLFIRI=folinic acid, fluorouracil, and irinotecan hydrochloride; IR=irinotecan

DISEGNO FATTORIALE



Valuta l'efficacia di B

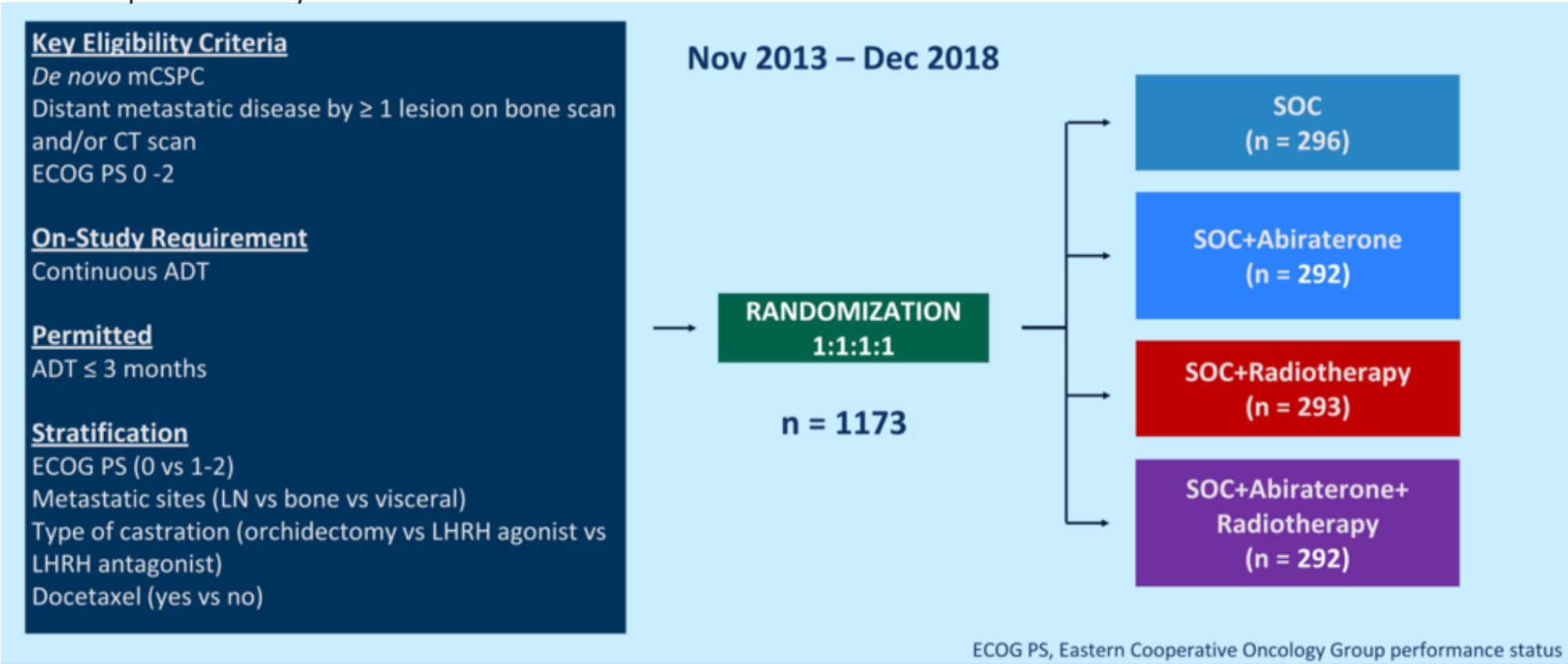
(disegno 2x2)

Valuta l'efficacia di A

Prerequisito: non interazione tra gli effetti degli interventi ("righe *Vs* colonne")

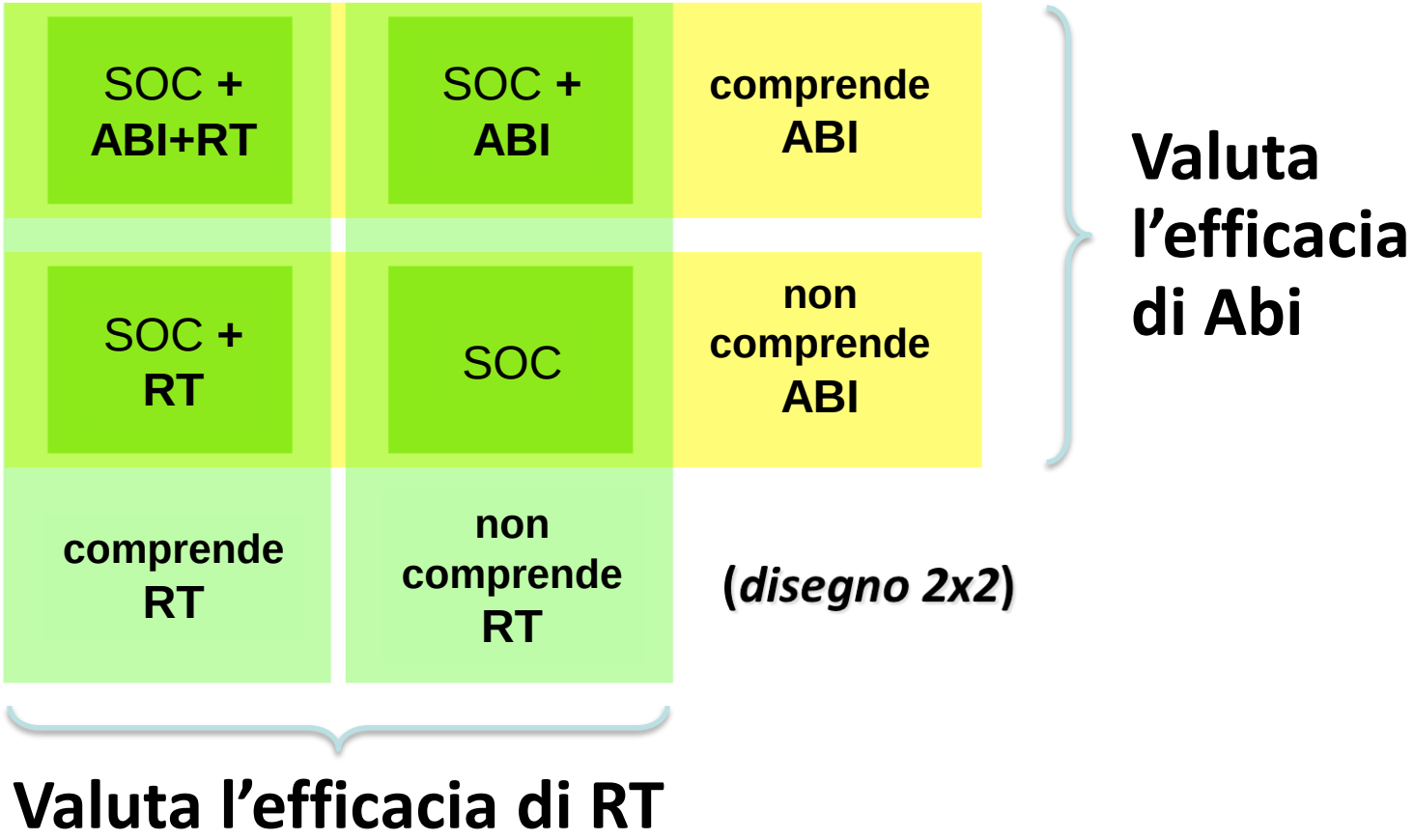
PEACE-1: A Phase 3 Trial with a 2x2 Factorial Design of Abiraterone Acetate plus Prednisone And/or Local Radiotherapy in Men with De Novo Metastatic Castration-Sensitive Prostate Cancer (mCSPC)

The trial used a 2 x 2 factorial design aimed at answering two questions, the role of abiraterone and the role of radiotherapy on top of the standard of care. There was no evidence of interaction between abiraterone and radiotherapy in the ADT + docetaxel population, and the two abiraterone arms (with or without radiotherapy) were thus pooled for subsequent analysis.



Fizazi K, et al. 2021 ASCO Annual Meeting, Virtual Annual Meeting #ASCO21, June, 4-8, 2021

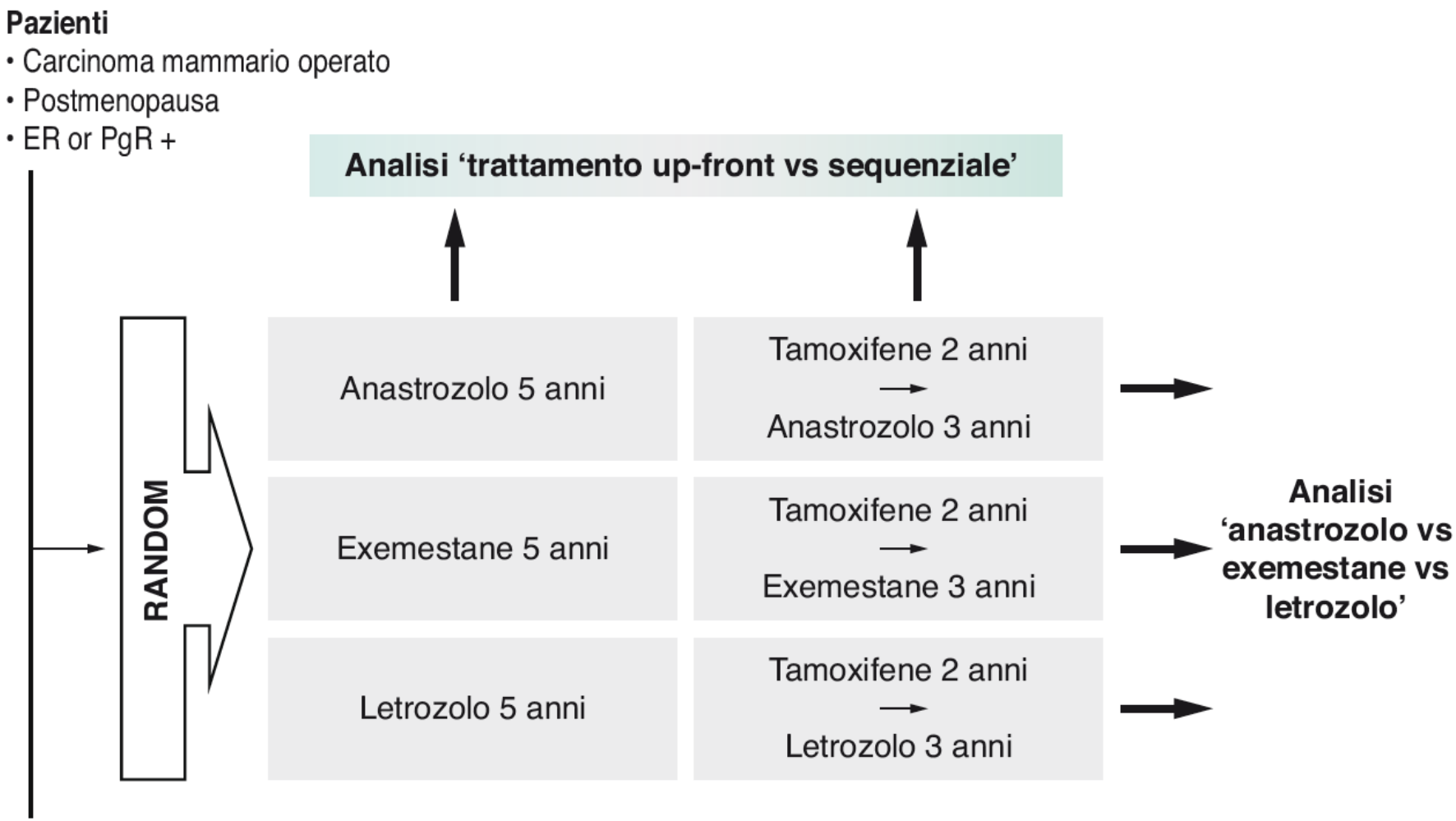
DISEGNO FATTORIALE



(*disegno 2x2*)

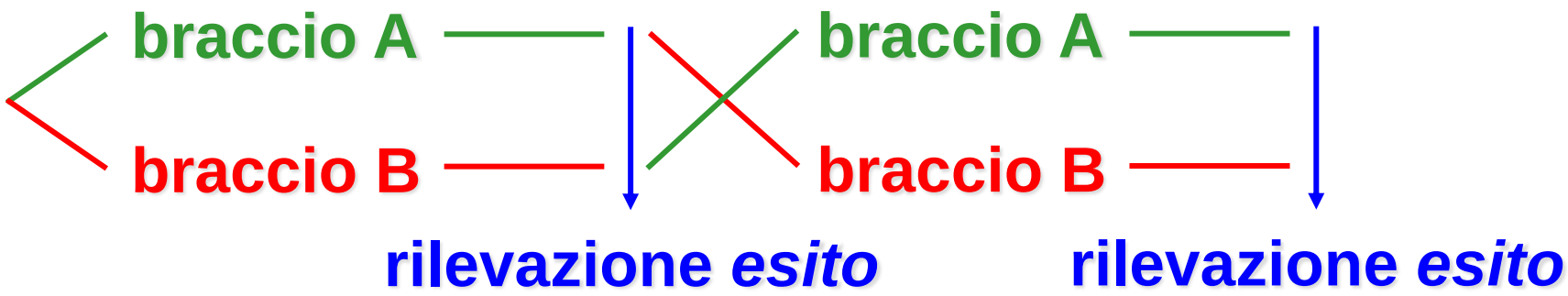
Gruppo Italiano Mammella (GIM) Studies

Source: Trial Sponsors > Index > G > Gruppo Italiano Mammella (GIM)



DISEGNO CROSSOVER

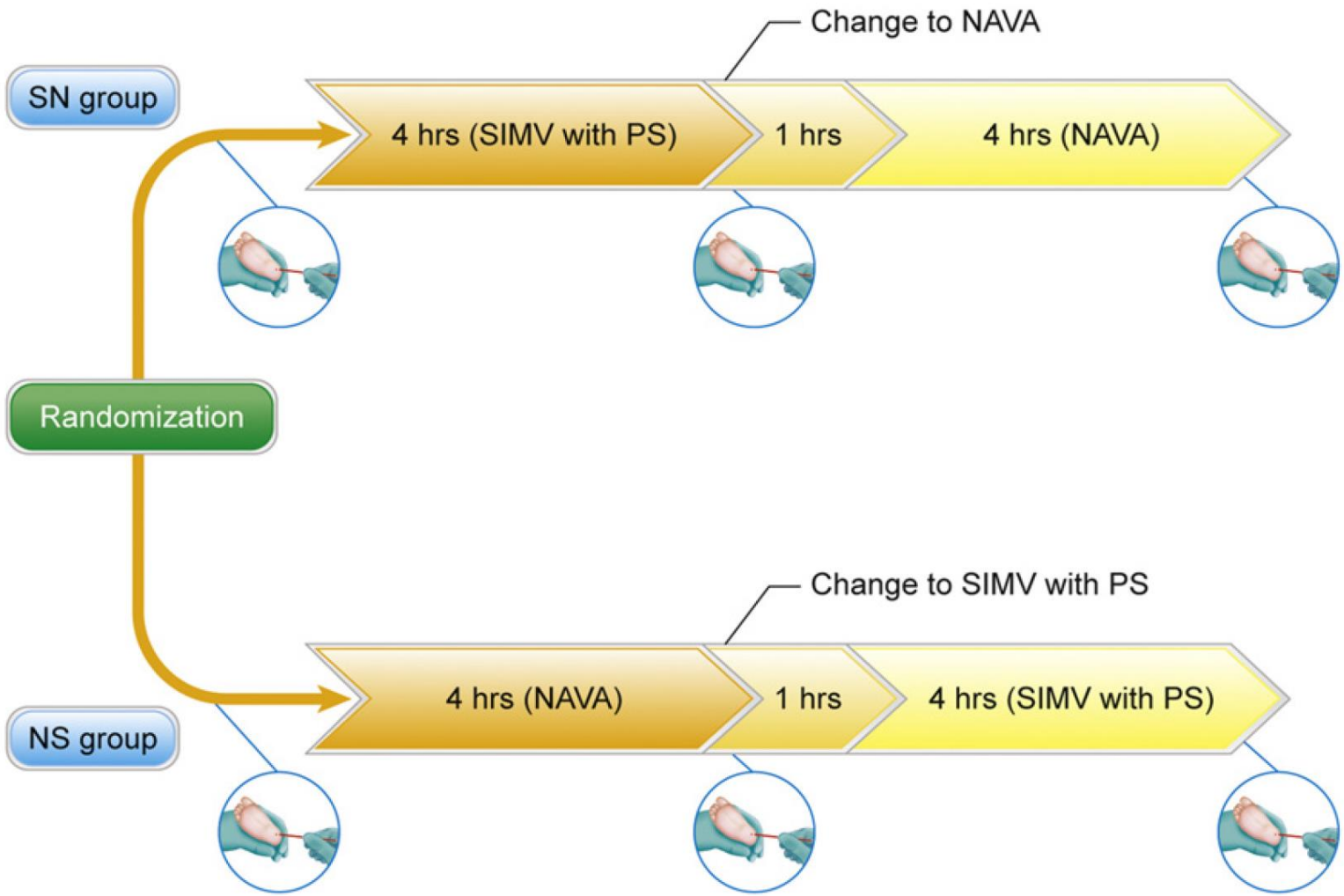
Ciascun paziente riceve entrambi i trattamenti oggetto di sperimentazione clinica (*within patient* vs *between patient*):



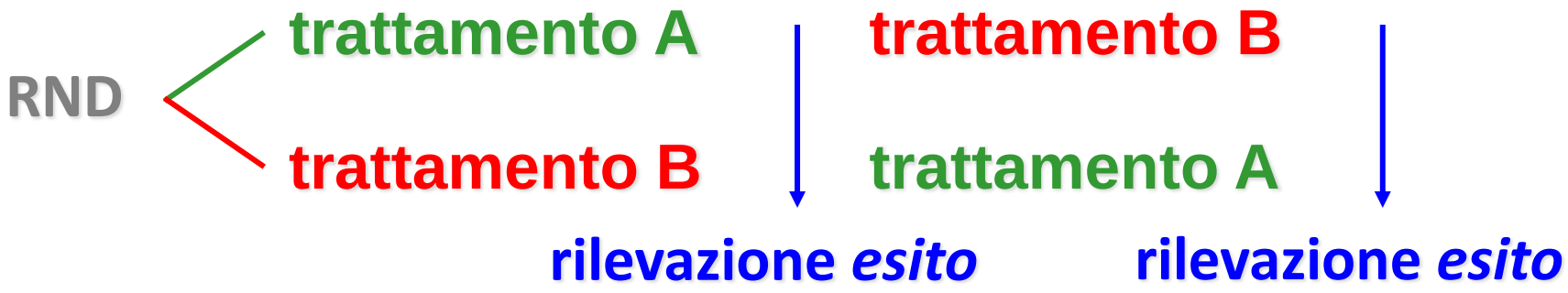
Viene "sottratta" dal confronto dei trattamenti l'influenza delle caratteristiche del paziente, le quali possono influire sull'indicatore di effetto → dimensione campionaria minore rispetto a uno studio a bracci paralleli.

Randomized Crossover Study of Neurally Adjusted Ventilatory Assist in Preterm Infants

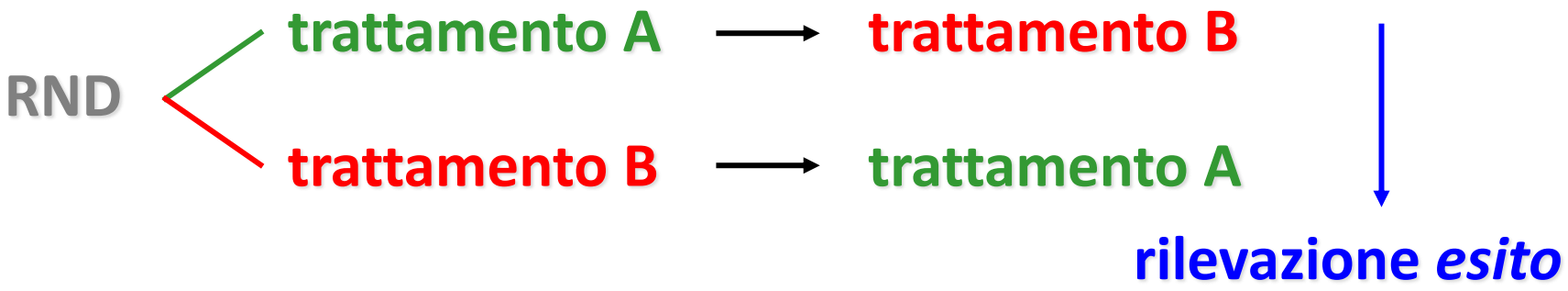
J Pediatr. 2012 Nov;161(5):808-13.



ATTENZIONE A NON CONFONDERE UN DISEGNO CROSSOVER…



…CON UN DISEGNO A BRACCI PARALLELI DI TIPO SEQUENZIALE



Choice of Control Group

- The selection of an appropriate control group is a critical decision which impacts on the scientific validity and ethical acceptability of a clinical investigation.
- The proper control group allows for discrimination between patient outcomes caused by the test treatment, and outcomes caused by other factors such as the natural progression of the disease, observer or patient expectations, or other treatments.

Health Authority Guidelines

- US Regulations require the establishment of safety and efficacy
  - Guideline for control group:
    - The standard regimen should have a well-characterized clinical benefit (survival benefit)
- EMEA
  - Reference therapy should be selected from the best available, evidence-based therapeutic options
    - Widely used, but not necessarily licensed regimen with a favourable benefit-risk convincingly documented through randomised trials and at least as good as alternative evidence-based treatment options

S.D. Averbuch, ASCO 2008

Health Authority Guidelines

- US Regulations require the establishment of

Considering Usual Medical Care in Clinical Trial Design

Liza Dawson[1]*, Deborah A. Zarin[2], Ezekiel J. Emanuel[3], Lawrence M. Friedman[4], Bimal Chaudhari[5], Steven N. Goodman[6]

PLoS Med 6(9): e1000111. doi:10.1371/journal.pmed.1000111

- EMEA
  - Reference therapy should be selected from the best available, evidence-based therapeutic options
    - Widely used, but not necessarily licensed regimen with a favourable benefit-risk convincingly documented through randomised trials and at least as good as alternative evidence-based treatment options

S.D. Averbuch, ASCO 2008

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

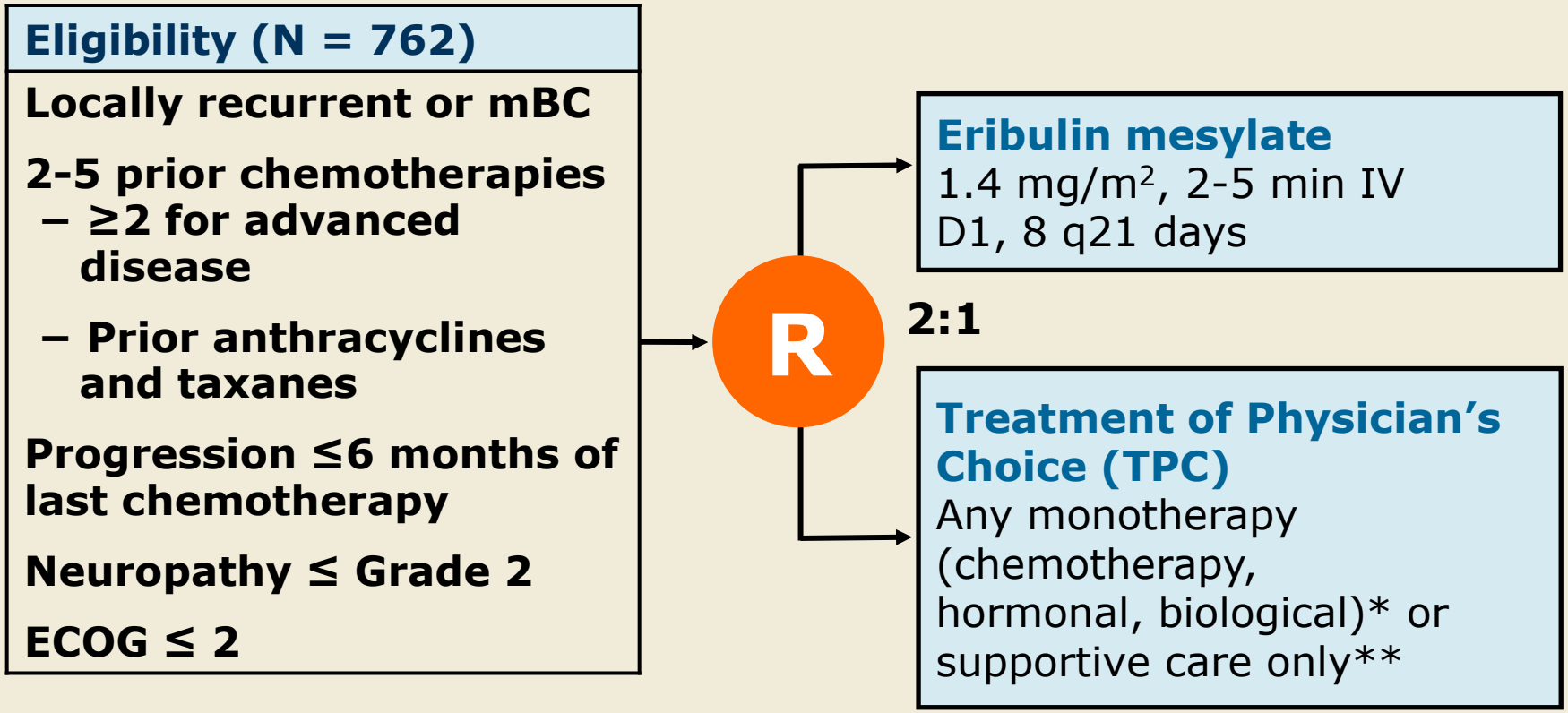
Eli M. Roth [a,*], Marja-Riitta Taskinen [b], Henry N. Ginsberg [c], John J.P. Kastelein [d], Helen M. Colhoun [e], Jennifer G. Robinson [f], Laurence Merlet [g], Robert Pordy [h], Marie T. Baccara-Dinet [i]

International Journal of Cardiology 176 (2014) 55–61

Ezetimibe was utilized as the comparator in this study as it is one of the options recommended for treating patients with statin intolerance [6].

The primary objective of this study was to evaluate the efficacy and safety of alirocumab monotherapy compared with ezetimibe in patients with hypercholesterolemia and at moderate cardiovascular (CV) risk (i.e. a 10-year risk of fatal CV events ≥1% and <5%) [5], who were not receiving statin or other lipid-lowering therapy.

EMBRACE Study Design

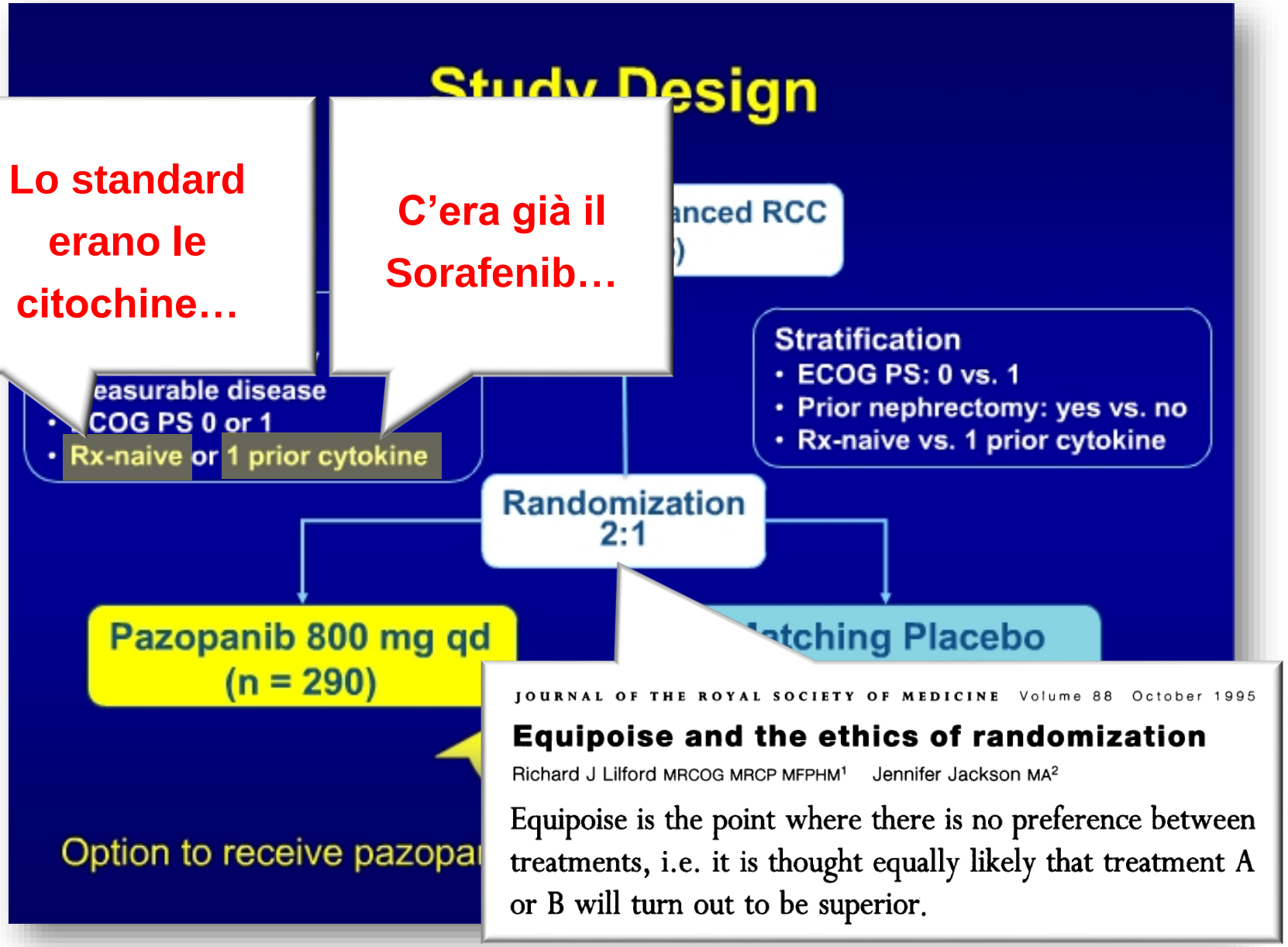


* Approved for cancer treatment
** Or palliative treatment or radiotherapy according to local practice

Twelves C et al. *Proc ASCO* 2010;Abstract CRA1004.

EFFICACIA E GRUPPO DI CONTROLLO

- **Studio di superiorità**
  - *L'uso del nuovo farmaco comporta un beneficio clinico per i pazienti? (controllo = placebo o nulla)*
  - *Il nuovo farmaco è più efficace di altri farmaci? (controllo = il miglior trattamento disponibile)*
- **Studio di non inferiorità**
  - *Cosa siamo disposti a "perdere" in cambio di minori effetti collaterali (disagi, costi)? (controllo = il miglior trattamento disponibile)*


Study Design
Lo standard erano le citochine…
C'era già il Sorafenib…
anced RCC
easurable disease
COG PS 0 or 1
Rx-naive or 1 prior cytokine
Stratification
• ECOG PS: 0 vs. 1
• Prior nephrectomy: yes vs. no
• Rx-naive vs. 1 prior cytokine
Randomization 2:1
Pazopanib 800 mg qd (n = 290)
atching Placebo
Option to receive pazopa
JOURNAL OF THE ROYAL SOCIETY OF MEDICINE Volume 88 October 1995
Equipoise and the ethics of randomization
Richard J Lilford MRCOG MRCP MFPHM[1] Jennifer Jackson MA[2]
Equipoise is the point where there is no preference between treatments, i.e. it is thought equally likely that treatment A or B will turn out to be superior.

Placebo/No Tx Arms In Absence Of Effective Therapy

- Control subjects typically not worse off than they would be outside the trial
- No-treatment controls acceptable when:
  - Alternative designs inadequate
  - Risks minimized and benefits maximized, *while ensuring answer to study question*
  - Fastidious attention to informed consent

JCO 26:1371, 2008

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- **Randomizzazione e mascheramento**
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

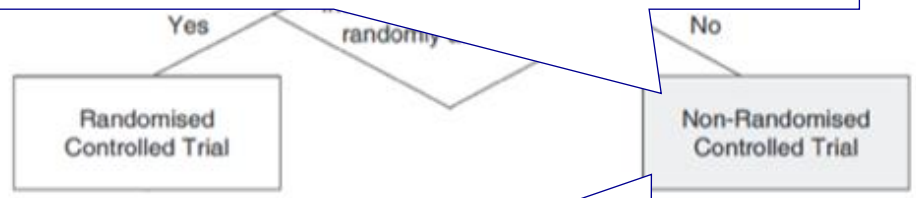
The 3 most important rules in drug development

- 1. Randomize
- 2. Randomize
- 3. Randomize

M Ratain, ASCO 2010

Nonrandomized Studies

A possible **bias** occurs because there is no random assignment of units in a target population to treatments.



Any observed difference between the outcomes of study arms may be attributable to baseline differences rather than to a true treatment effect.

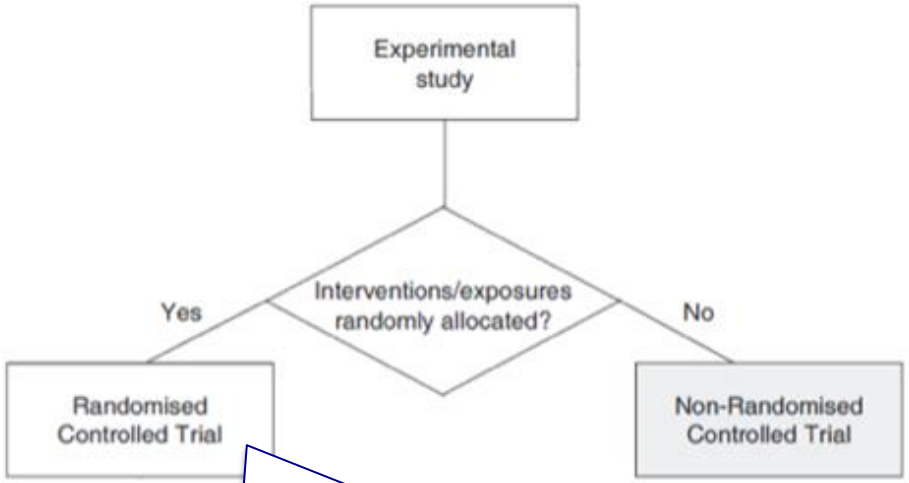
RANDOMIZZAZIONE

Assegnazione casuale dei pazienti al gruppo sperimentale o di controllo, al fine di assicurare che tutti i fattori prognostici - noti e sconosciuti - si distribuiscano omogeneamente nei due gruppi.

Tutti i requisiti della randomizzazione hanno lo scopo di assicurare che il **processo con cui vengono creati i due gruppi a confronto segua le leggi del caso**, e che **nessun fattore possa interferire** con la sua casualità.

Lachin, 2000

Randomization



Minimizes *allocation* ***bias***, balancing both known and unknown prognostic factors, in the assignment of treatments.

RANDOMIZATION COMPONENTS

Item	Descriptor
Sequence generation	Method used to generate the random allocation sequence, including details of any restriction (eg, blocking, stratification)
Allocation concealment	Method used to implement the random allocation sequence (eg, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned
Implementation	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups

RANDOMIZZAZIONE A BLOCCHI

- La sequenza totale delle assegnazioni previste viene divisa in un certo numero di blocchi successivi.
- Il blocco rappresenta un gruppo di assegnazioni all'interno del quale vi è bilanciamento nel numero di pazienti assegnati ai due trattamenti, in modo da rispettare il rapporto di assegnazione previsto.
- I blocchi dovrebbero essere di dimensione variabile, in dipendenza dalle dimensioni campionarie e dal numero di strati)
  - es. blocco di 4: ABAB
  - es. blocco di 6: ABABAB
  - es. blocco di 8: ABABABAB

RANDOMIZZAZIONE A BLOCCHI

- ✓ Se n = 4 (dimensione del blocco)
- ✓ Se x = 2 (numero dei trattamenti)
- ✓ Se A:B = 1:1 (rapporto di assegnazione)

Quante (e quali) sono le possibili permutazioni?

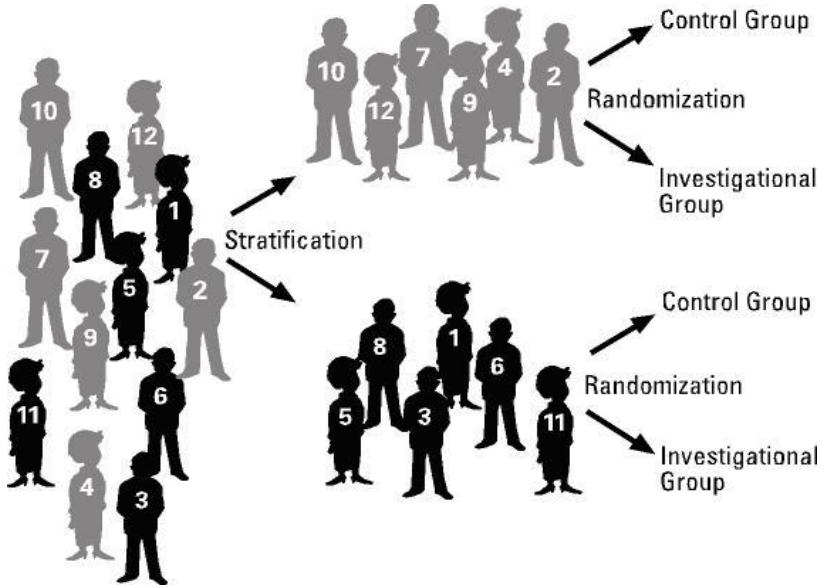
$$\frac{n!}{x!\,(n-x!)} = \frac{4!}{2!\,2!} = 6$$

AABB
ABAB
BABA
ABBA
BAAB
BBAA

RANDOMIZZAZIONE STRATIFICATA

Allestimento di liste di randomizzazione separate per una o più caratteristiche pre-trattamento:

- ✓ misura atta a evitare sbilanciamenti fra i trattamenti a confronto per specifici fattori prognostici;
- ✓ possibili vantaggi di tipo gestionale e organizzativo (es. stratificazione per Centro);
- ✓ considerare solo fattori di stratificazione oggettivamente definibili.



RANDOMIZZAZIONE STRATIFICATA

Il numero di liste random che si viene a formare con la stratificazione è uguale al prodotto del numero degli strati di ogni fattore di stratificazione:

RAPID axSpA

- ✓ Site * 104
- ✓ mNY status * 2
- ✓ Prior TNF inhibition * 2

416 Liste di Randomizzazione

(325 pazienti)

RAPID PsA

- ✓ Site * 92
- ✓ Prior TNF inhibitor use * 2

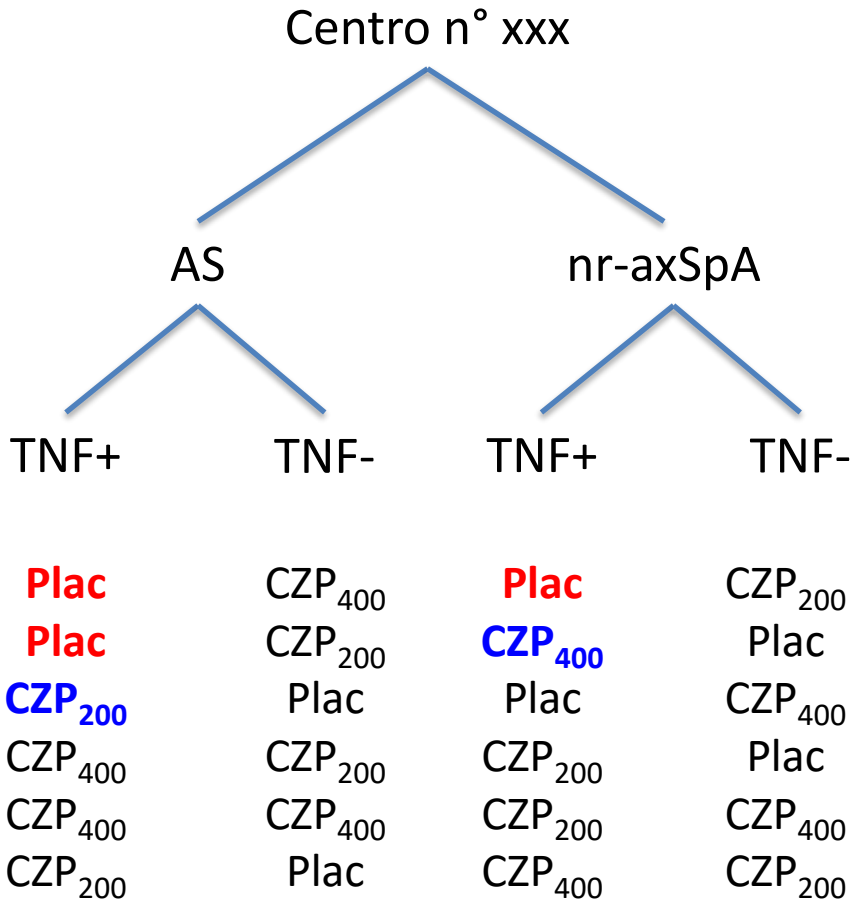
184 Liste di Randomizzazione

(409 pazienti)

Attenzione alla *overstratification* !

OVERSTRATIFICATION IN RAPID axSpA?

Possibile Scenario (potrebbe valere anche l'ipotesi opposta)



(alcune delle possibili permutazioni del *blocco* di 6)

Nel Centro n° xxx sono stati arruolati 5 pazienti con precedente esposizione a TNFi:

- ✓ di 3 pazienti con AS, 2 sono stati assegnati a Plac e 1 a CZP_{200}
- ✓ di 2 pazienti con nr-axSpA, 1 è stato assegnato a Plac e 1 a CZP_{400}

Gli strati TNF+ hanno un arruolamento non sufficiente a garantire il riempimento del *blocco* di 6 pazienti

Placebo-Controlled Trial

Experimental VS. Placebo

Open-Label With Active Arm*

Active VS. Experimental

Double-Dummy Double-Blind*

Experimental + Placebo VS. Placebo + Active

Experimental + Placebo VS. Placebo + Active

● Placebo ● Experimental product ● Safe and effective product * Trials with active control arm

MASCHERAMENTO (Blinding)

- Insieme delle procedure atte a prevenire distorsioni dovute al fatto che il Paziente, il Medico o il Valutatore sono a conoscenza del trattamento ricevuto dal Paziente:
  - *Paziente = singolo cieco*
  - *Paziente + Medico = doppio cieco*
  - *Paziente + Medico + Valutatore = triplo cieco*
- Necessità connessa agli obiettivi dello studio e al tipo di variabile di effetto utilizzata

MASCHERAMENTO (Blinding)

- Insieme del
  distorsioni d
  Medico o il V
  trattamento
  - *Paziente =*
  - *Paziente +*
  - *Paziente + Medico* *ore = triplo cieco*
- Necessità connessa agli obiettivi dello studio e al tipo di variabile di effetto utilizzata

	paziente	*medico*	*valutatore*
Decesso (per ogni causa)	**no**	**no**	**no**
Decesso per causa specifica	**no**	**no**	**si**
Recidiva, progressione	**no**	**no(?)**	**si**
Risposta clinica	**no (?)**	**no(?)**	**si**
Risposta soggettiva	**si**	**si**	**si**
Dolore	**si**	**si**	**si**
Stato psichico	**si**	**si**	**si**

Less Hypoglycemia With Insulin Glargine in Intensive Insulin Therapy for Type 1 Diabetes

ROBERT E. RATNER, MD
IRL B. HIRSCH, MD
JAMES L. NEIFING, MD
THOMAS E. MECCA, PHD
CRAIG A. WILSON, PHD
SATISH K. GARG, MD

FOR THE U.S. STUDY GROUP OF INSULIN GLARGINE IN TYPE 1 DIABETES

Diabetes Care 23:639–643, 2000

A double-blind design was not feasible because insulin glargine is a clear solution and is distinguishable from cloudy NPH insulin.

Placebo-Controlled Trials in Surgery

A Systematic Review and Meta-Analysis

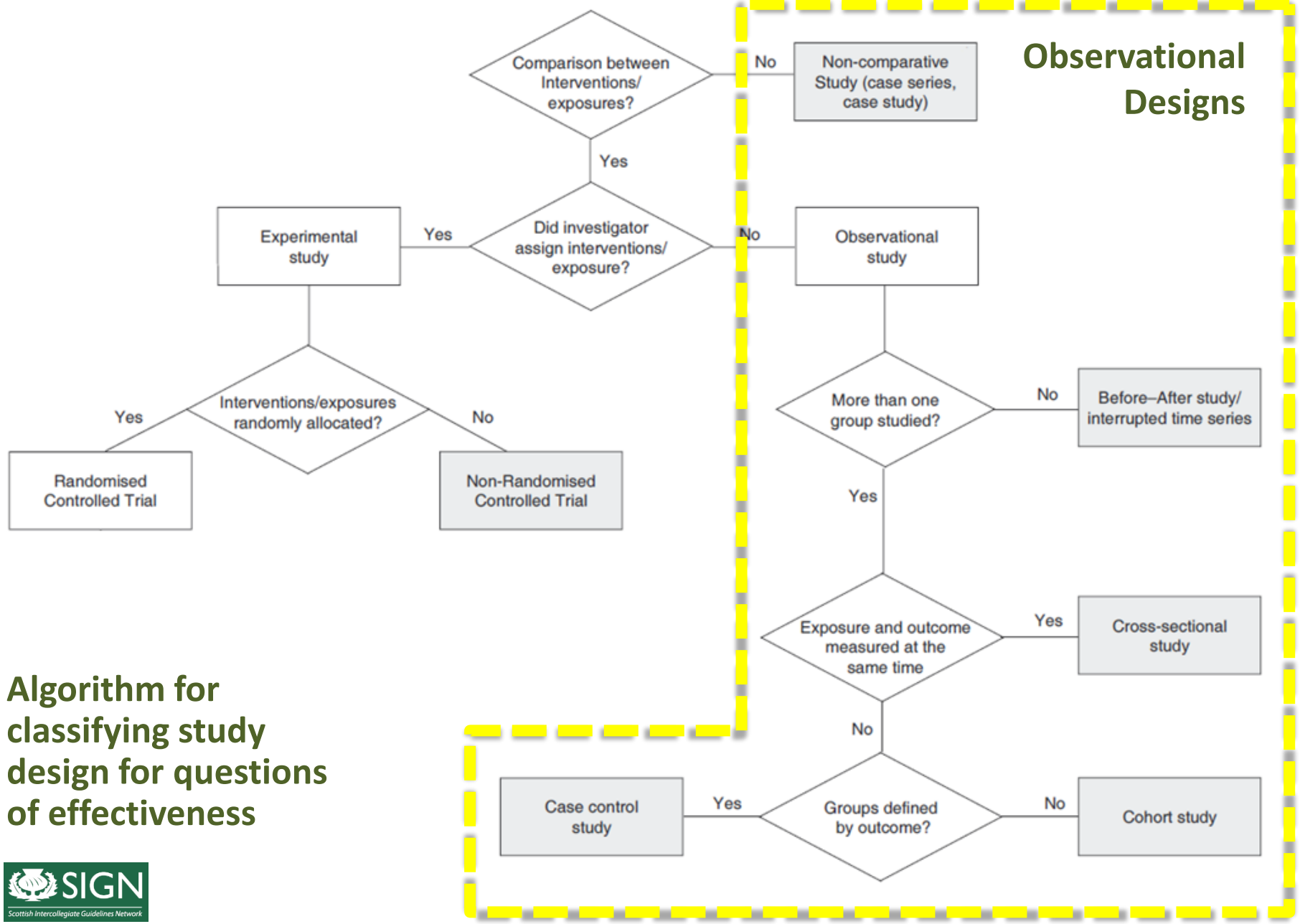
Pascal Probst, MD, Kathrin Grummich, HD, Julian C. Harnoss, MD, Felix J. Hüttner, MD, Katrin Jensen, PhD, Silvia Braun, MD, Meinhard Kieser, PhD, Alexis Ulrich, MD, Markus W. Büchler, MD, and Markus K. Diener, MD

Medicine • Volume 95, Number 17, April 2016

Properly conducted placebo-controlled trials can be feasible and are valid methodological instruments for the evaluation of efficacy of surgical interventions. Placebo controlled trials have a scientific value, and are mostly published in high-impact journals. However, surgical placebos entail a considerable risk for study participants and should be used only if justified by the clinical question and methodological necessity.

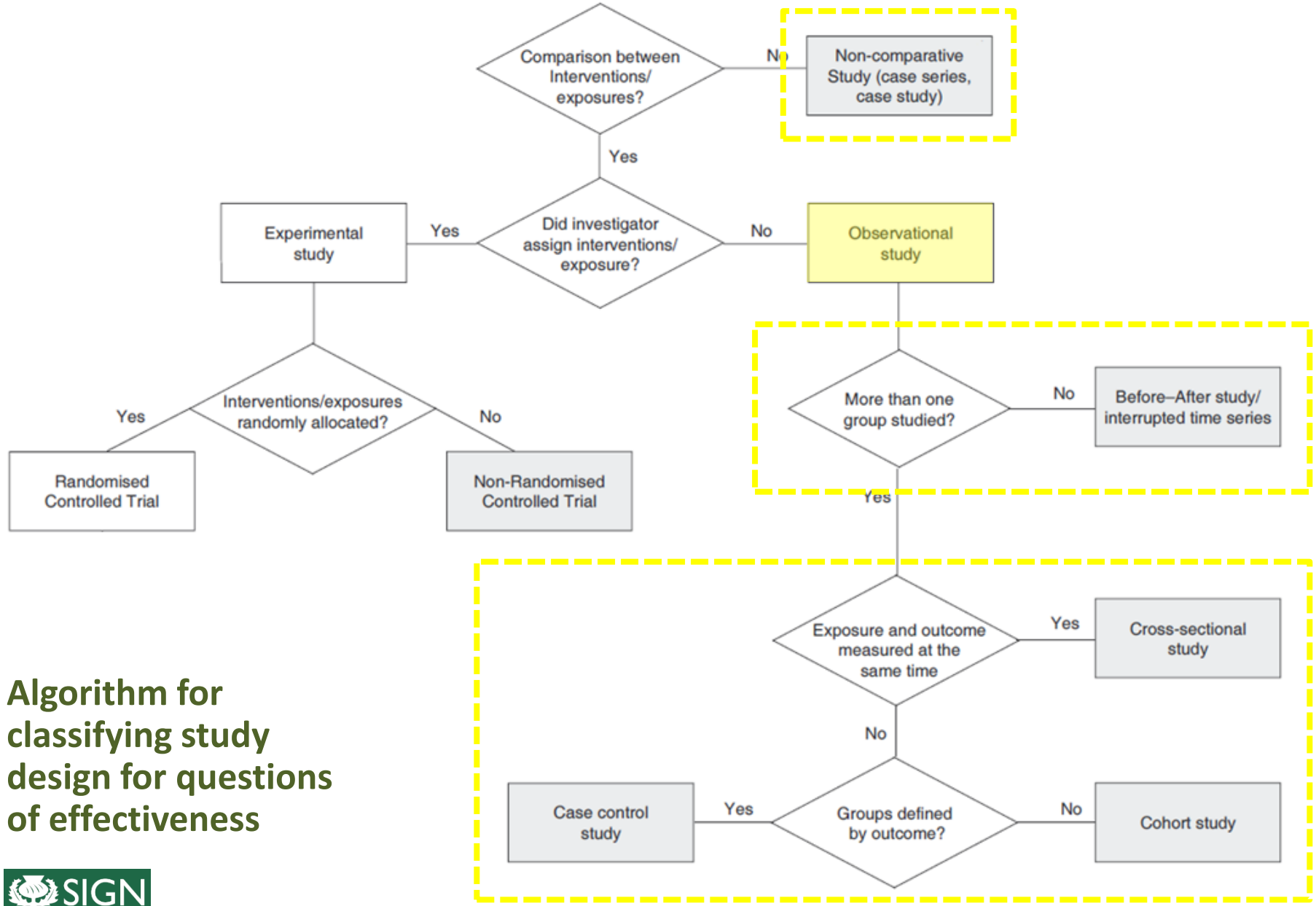
- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- **Il disegno dello studio: studi osservazionali**
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

Observational Designs



Algorithm for classifying study design for questions of effectiveness

SIGN
Scottish Intercollegiate Guidelines Network



Algorithm for classifying study design for questions of effectiveness

SIGN
Scottish Intercollegiate Guidelines Network

Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

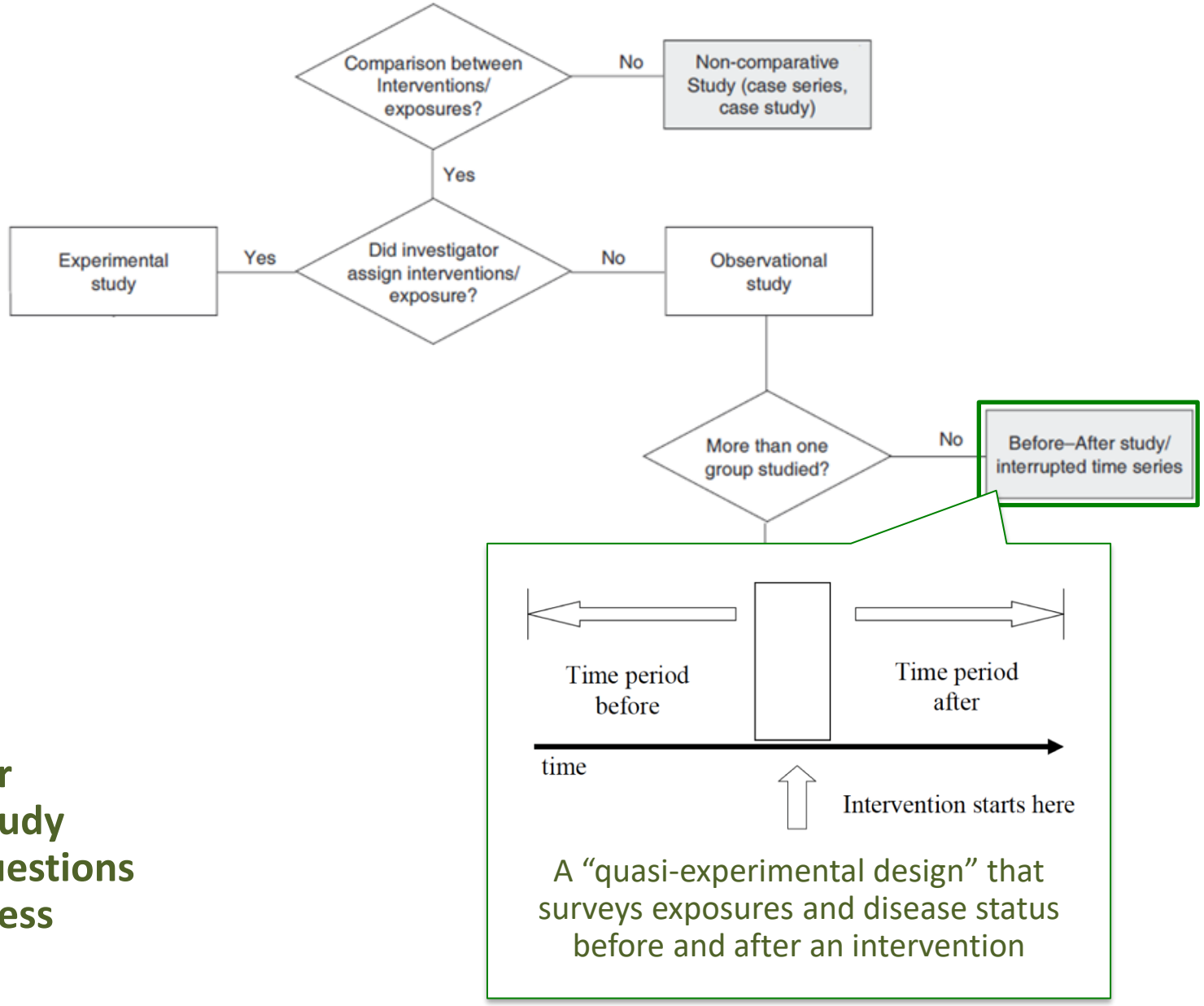
Identify the incidence of a particular outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions



Algorithm for classifying study design for questions of effectiveness

SIGN
Scottish Intercollegiate Guidelines Network

Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

Identify the incidence of a particular outcome over time

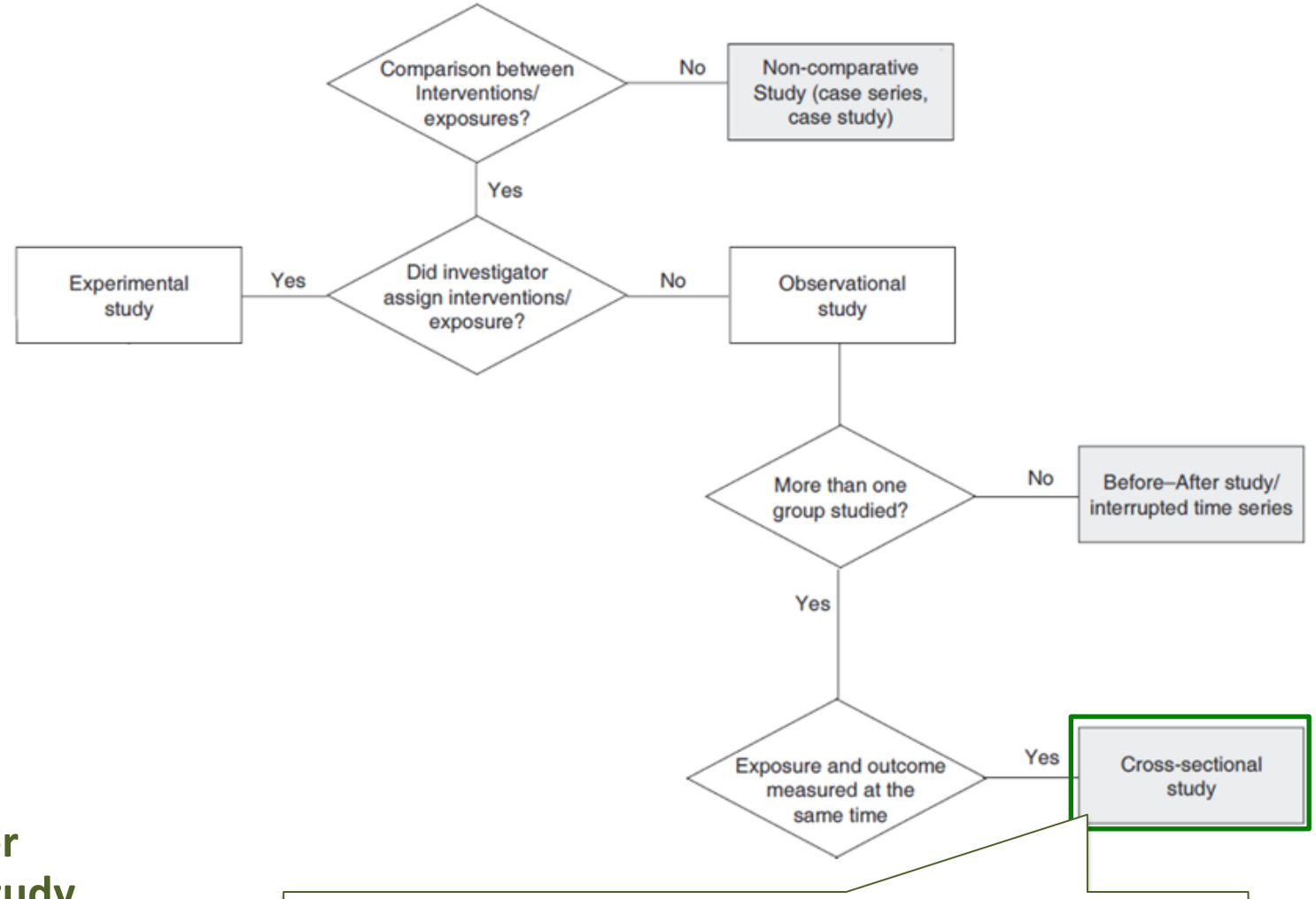
Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

Algorithm for classifying study design for questions of effectiveness



Subjects selected irrespective of the presence or absence of the characteristics of interest. Similar to a case series, except that the purpose of the analysis is to record associations between variables, rather than merely to report frequencies of their occurrence

SIGN
Scottish Intercollegiate Guidelines Network

Observational Studies

"EPIDEMIOLOGIC" Vs "THERAPEUTIC"

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

Identify the incidence of a particular outcome over time

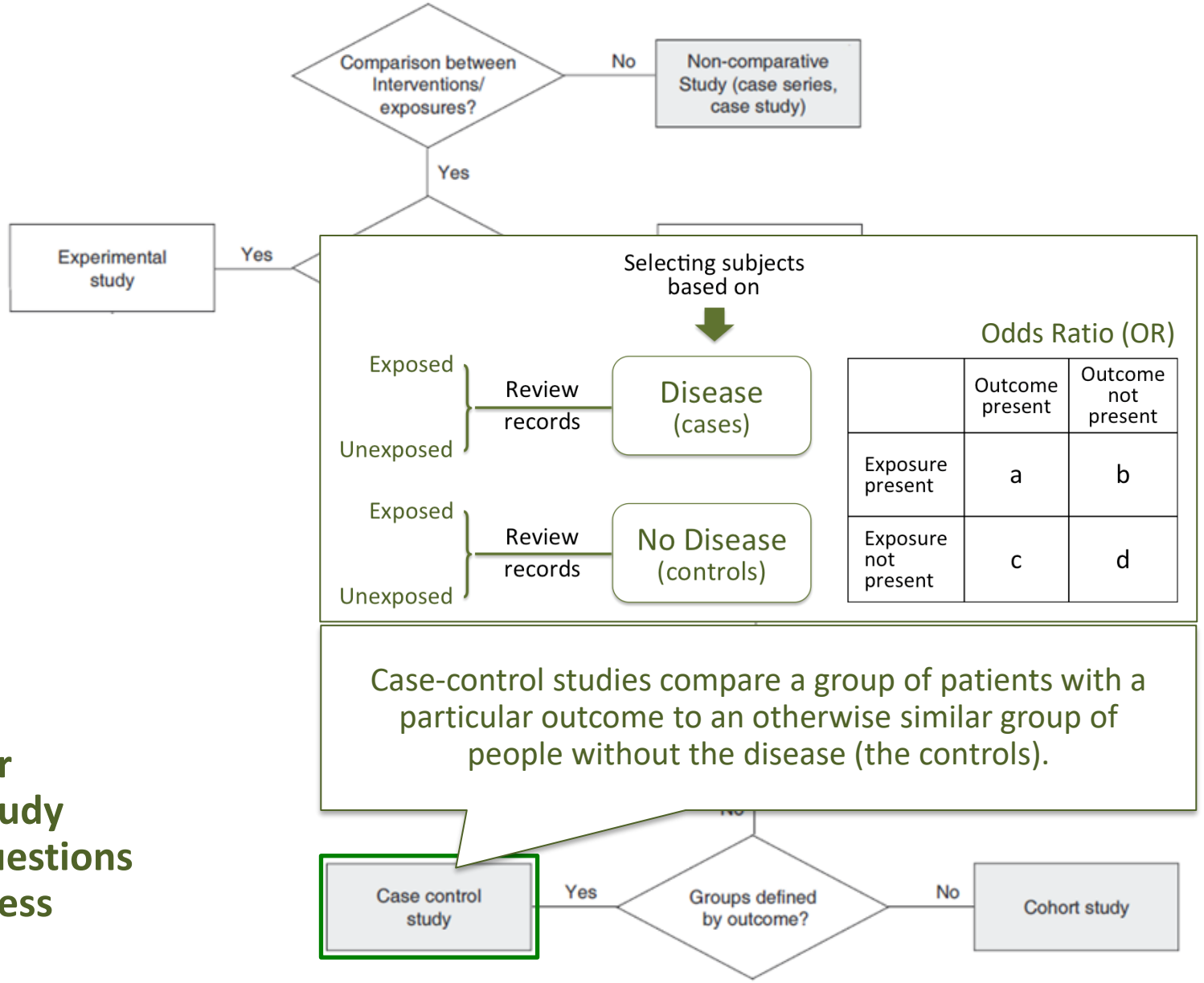
Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

Algorithm for classifying study design for questions of effectiveness



Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

Identify the incidence of a particular outcome over time

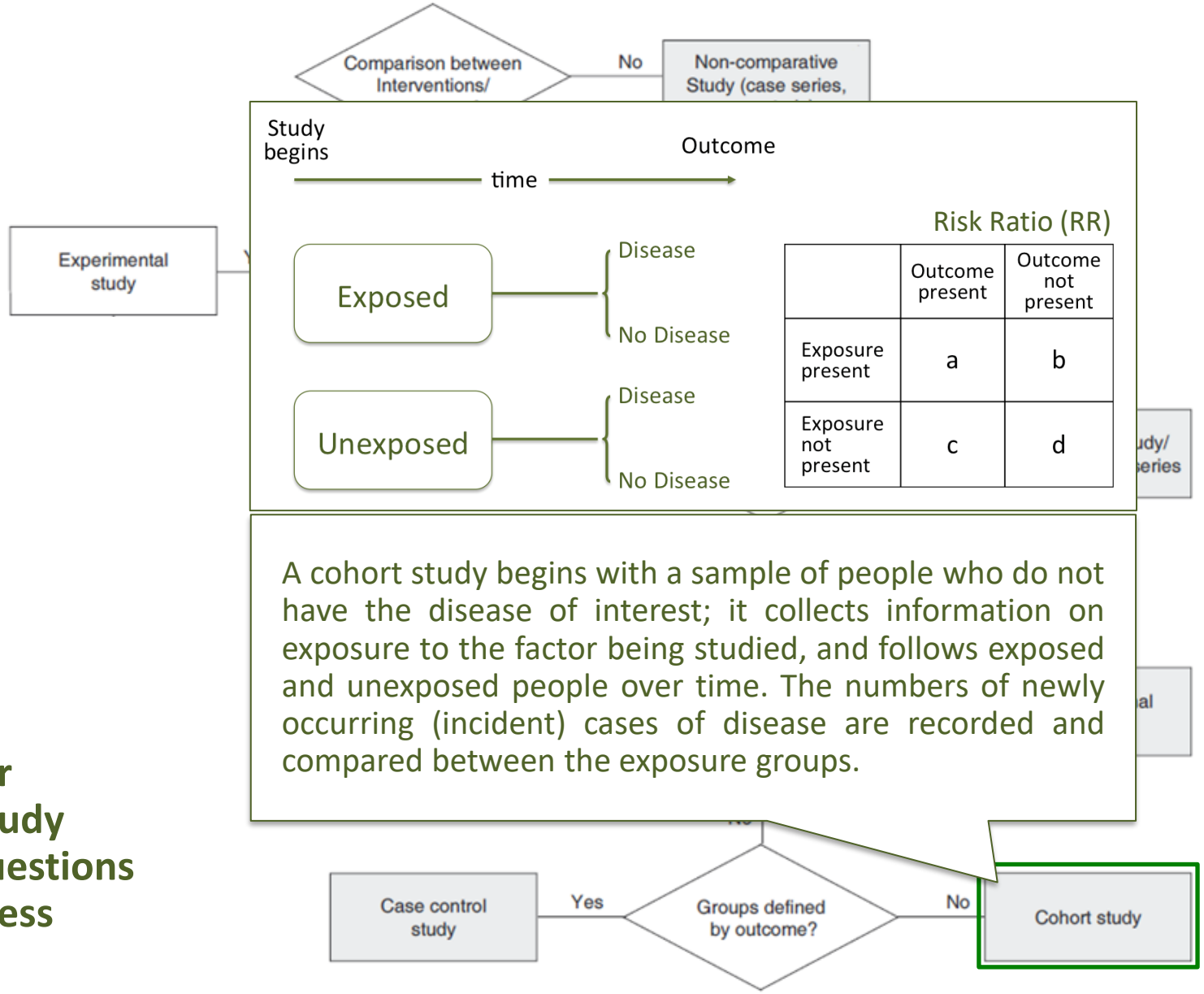
Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

Algorithm for classifying study design for questions of effectiveness



Time matters…

Cross-Sectional Studies
(exposure and outcome measured at the same time)

Case-Control Studies
(groups defined by the outcome)

Cohort Studies
(groups not defined by the outcome)

OBSERVATIONAL STUDY: A DEFINITION

An observational study draws inferences from a sample to a population where the independent variable is **not under the control** of the researcher.

The term observational study covers a wide range of study designs, a common feature of which is that they are noninterventional, in the sense that the **study protocol does not determine the** precise **features of any therapy** given to the participants in the study.

Observational Studies

“EPIDEMIOLOGIC” Vs “THERAPEUTIC”

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particular outcome

Cohort

Identify the incidence of a particular outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

Retrospective, non comparative efficacy analysis

Mature results on metastatic breast cancer patients with prolonged (≥1 year) exposure to first-line bevacizumab combined with paclitaxel from a large observation study

P2 – 16 - 03

M Schmidt[1], A Schneeweiss[2], F Foerster[3], M Geberth[4], C Schumacher[5], W Hollburg[6], U Söling[7], B Aktas[8], S Kümmel[9]

[1]*University Hospital Mainz, Mainz, Germany;* [2]*University of Heidelberg, National Center for Tumor Diseases, Heidelberg;* [3]*University of Applied Sciences Zwickau, Zwickau;* [4]*SPGO-Mannheim, Mannheim;* [5]*St Elisabeth-Krankenhaus, Köln;* [6]*Hämatologisch-Onkologische Praxis Altona im Struenseehaus, Hamburg;* [7]*Clinical Practice, Kassel;* [8]*University Clinic, Essen;* [9]*Kliniken Essen-Mitte, Essen, Germany*

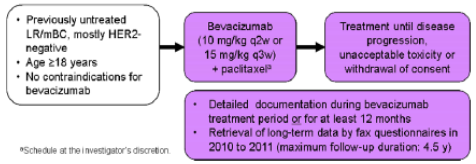
Background

- Combining first-line bevacizumab with standard chemotherapy significantly improved PFS and response rate compared with chemotherapy alone in three randomised phase III trials in HER2-negative locally recurrent or metastatic breast cancer (LR/mBC).[1–4]
- The efficacy and tolerability observed in these randomised phase III trials was supported by results of large cohort studies in routine oncology practice, with 'real-life' patient selection and treatment procedures:
  - The multinational ATHENA study, evaluating first-line bevacizumab-containing therapy in 2264 patients from overall 37 countries.[5,6]
  - A German non-interventional study evaluating first-line bevacizumab–paclitaxel, according to the European label, in 865 patients.[7]
- In both of the 'real-life' studies and in two smaller prospective/ambispective studies conducted in Japan[8] and France[9] a considerable proportion of patients received bevacizumab for prolonged periods.
- We report the final analysis of efficacy and safety, including mature long-term follow-up, in the subset of patients who received bevacizumab for ≥1 year in the German non-interventional study.

Methods – Study design

- The design of this non-interventional study is shown in Figure 1.
  - Endpoints were efficacy (overall response rate, PFS, OS) and safety (adverse events [AEs], AEs of special interest, serious AEs).
- Paclitaxel schedule, diagnostics and frequency of follow-up visits were at the discretion of the physician.
  - Detailed therapy data were collected during bevacizumab therapy, at least for up to 1 year, with further follow-up for efficacy at several time points after the end of intensely documented observation or discontinuation of bevacizumab, up to a maximum of 4.5 years (y).
- For this analysis, data from patients treated with bevacizumab for ≥1 year were extracted and analyzed retrospectively.
- Data base closure was in July 2012.

Figure 1. Study design



[a]Schedule at the investigator's discretion.

Results

Patients

Between May 2007 and September 2009, a total of 1123 patients were documented in this observational study
- Of these, 865 patients received a bevacizumab–paclitaxel combination as 1st line therapy.
- Of these, 167 (20%) had received bevacizumab for ≥1 year.

- Baseline characteristics of bevacizumab-paclitaxel patients with antibody treatment < or ≥1 year are shown in Table 1.
- Patients with long-term bevacizumab treatment typically show characteristics which are generally associated with less aggressive disease: fewer patients with G3 tumors, disease-free interval ≤ 12 months, ≥3 metastatic sites, liver lesions, hormone receptor or triple-negative status, previous (neo)adjuvant chemotherapy, and ECOG performance ≥1.

Treatment exposure

- In 78% of those treated for ≥1 year, bevacizumab was continued as a single agent after discontinuation of chemotherapy. Other changes of the cytotoxic regimen during 1st line were rare.
- 9% of patients treated for ≥1 year, and 12% in the complementary group received bevacizumab beyond progression (ie in combination with second- or later lines of chemotherapy).
- In the first cycle, 80% of patients in this subgroup received bevacizumab at a dose of 10 mg/kg, administered every two weeks. Dose reductions of the antibody were rare in the total study population (2% of patients) and not distinctly more common in the long-term subgroup (3%). Respective numbers for chemotherapy dose reductions were 21% and 28%.
- The corresponding numbers for treatment delays were 11% and 12% for bevacizumab, based on cycles, and 61% and 50%, based on patients (at least once). For chemotherapy these rates were generally higher, with 17% and 18% of cycles, 81% and 67% of patients, respectively.

Table 1. Baseline characteristics

Characteristic	Bevacizumab for <1 year (n=698)	Bevacizumab for ≥1 year (n=167)
Median age, years (range)	58 (26–87)	57 (28–79)
Metastatic at initial diagnosis, %	20	18
Grading G3 at diagnosis, %	42	32
Disease-free interval <12 months, %	26	14
Metastatic sites, %		
≥3	34	29
Liver	44	35
Lung	33	43
Bone	54	53
Hormone receptor-positive disease, %	70	75
Triple-negative disease (TNBC), %	20	12
Prior endocrine therapy for mBC, %	25	33
Prior (neo)adjuvant chemotherapy, %	67	61
Prior trastuzumab therapy, %	4	4
ECOG status		
0	41	49
1	50	43
2	8	8
3	1	1

Efficacy

- The overall response rate (ORR) in patients treated for ≥1 year was 80%, including complete responses in 19% of patients. The same ORR was observed in the TNBC subgroup.
  - In the population with bevacizumab for < 1 year, the overall response rate was 57%, including complete responses in 9%.
- Long-term median progression-free survival (PFS, Fig. 2) was:
  - 9.6 months in the overall population (715/865 events observed, 83%).
  - 18.4 months in patients treated for ≥1 year (125/167 events observed, 75%)
  - 8.0 months in patients treated for <1 year (590/698 events observed, 85%)
- Median overall survival (OS, Fig. 3) was:
  - 21.6 months in the overall population (524/865 events observed, 61%).
  - 35.7 months in patients treated for ≥1 year (72/167 events observed, 43%)
  - 18.0 months in patients treated for <1 year (452/698 events observed, 65%)

Figure 2. Progression-free survival by duration of bevacizumab treatment (<1 vs. ≥ 1 year)

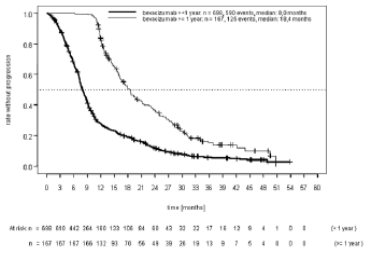
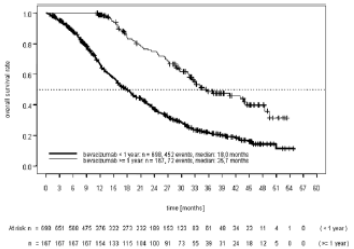


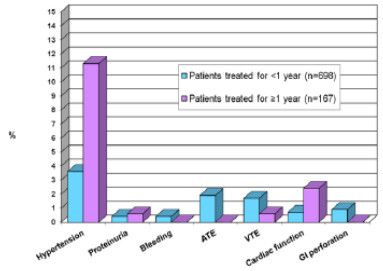
Figure 3. Overall survival by duration of bevacizumab treatment (<1 vs. ≥ 1 year)



Safety

- Grade 3/4 adverse events of special interest for bevacizumab (those reported in previous clinical trials) are shown in Figure 4.
- The most common grade 3/4 adverse events in patients treated with bevacizumab for ≥1 year were:
  - Hypertension (11% of patients)
  - Pain (9%)
  - Leukopenia (8%).
- There were no cases of gastrointestinal perforation or arterial thromboembolic events in those treated for ≥1 year. No confirmed reversible posterior leukoencephalopathy syndrome was reported in the total observation study population.
- In relation to the subgroup with bevacizumab for < 1 year, long-term treatment patients showed distinctly higher frequencies (all grades) with respect to hypertension (35% vs. 25%), proteinurea (17% vs. 9%), and sensory neuropathy (43% vs. 26%). However, the latter is obviously due to the higher amount of paclitaxel chemotherapy administered in the ≥1 year subgroup.

Figure 4. Grade ≥3 adverse events of special interest for bevacizumab



Comparison to ATHENA results

	This study	ATHENA[5,6]
Patient no. (recruitment)	865 (2007 – 2009)	2264 (2006 – 2009)
Country	Germany	multinational (37 countries)
Concurrent chemotherapy	paclitaxel only	paclitaxel (35%), docetaxel (33%) others/combinations (32%)
Median age / rate ≥ 70 y	58 years / 16%	53 years / 8%
Total population efficacy (ORR/ median PFS/ median OS/ 1 y OS rate)	57% / 9.6 mo / 21.6 mo / 73%	52% / 9.7 mo / 25.2 mo / 73%
Rate with treatment ≥ 1 y	20%	21%
Characteristics associated with <u>no</u> prolonged bevacizumab treatment (≥ 1 y)	G3, disease-free interval ≤ 12mo, ≥3 metastatic sites, liver lesions, HR- or triple-neg. status, previous adjuvant chemotherapy, ECOG ≥1	identical findings, except for missing impact of adjuvant chemotherapy (grading not reported)
Maturity of OS data	events observed in 43%	events observed in 27%[5]
≥ 1 y bevacizumab treatment population efficacy (ORR/ median PFS/ OS)	80% / 18.4 mo / 35.7 mo	68% / 19.9 mo / 29.6 mo

Discussion and Conclusions

- A notable proportion of patients seems to derive benefit from prolonged exposure to first-line bevacizumab-containing therapy.
- In the present analysis, baseline characteristics appeared more favorable in the subset of patients treated for ≥1 year than in the overall population. However, this might be an underlying association without a specific causal relation to antibody treatment duration.Typically, only these conventional parameters are collected in non-interventional studies, preventing the analysis and identification of new characteristics. Therefore, further prospective clinical research is needed to detect predictive biomarkers for bevacizumab.
- Efficacy data obviously have a bias towards improved outcome in those able to continue bevacizumab for ≥1 year, as these patients had sustained disease control for ≥1 year. This is a common limitation in single-cohort studies, when comparing time-related outcomes in patients treated for different durations.
- Nevertheless, the favorable survival outcomes of prolonged bevacizumab-containing therapy are of interest and suggest that some patients achieve sustained disease control with continued first-line bevacizumab–paclitaxel with limited side effects.
- Although our study is clearly different from the ATHENA project with respect to geographical, demographic and therapeutic homogeneity, the results reveal some striking similarities in "long-term" patient numbers and characterization, efficacy and safety endpoints. (The moderate dissimilarity in median OS is easily explained by the 5 year difference in median age.)
- Moreover, we confirmed and extended the positive findings from ATHENA, both in a population distinctly shifted to elderly patients, and based on a more mature overall survival data coverage.

References and acknowledgments

1. Miller K, et al. N Engl J Med 2007;357:2666-76.
2. Miles D, et al. J Clin Oncol 2010;28:3239–47.
3. Robert N, et al. J Clin Oncol 2011;29:1252–60.
4. Miles DW, et al. Ann Oncol 2013; published online: July 25, 2013.
5. Smith IE, et al. Ann Oncol 2011;22:595-602.
6. Smith I, et al. Breast Cancer Res Treat 2011;130:133–43.
7. Klare P, et al. J Clin Oncol 2011;29(suppl): abstract 1079.
8. Aogi K, et al. Breast Cancer Res Treat 2011;129:829–38.
9. Saghatchian M et al. ESMO Congress 2012, abstr. 937

The ML21165 study was sponsored by Roche Germany. Support for third-party writing assistance for this poster was provided by Roche Germany.
Presented at the 36th San Antonio Breast Cancer Symposium, December 10 to 14, 2013.

UNIVERSITY OF LIVERPOOL

VICTOR: Vinflunine (Vin) in advanced metastatic Transitional Cell Carcinoma of the Urothelium (TCCU): A retrospective analysis of the use of Vinfluinine in multi-centre real life setting as second line chemotherapy through the free of charge programme (FOCP) for patients in the UK and Ireland.

S.A.Hussain[1,2], J. Ansari[3], R. Huddart[4], Alison Hassall[2], E. Harrold[5], D.G. Power[5], J. Lyons[6], J. Wylie[6], M. Vilarino-Varela[7], D.Wilkinson[8], R. McMenemin[8], L.M. Pickering[9], G. Faust[10], S. Chauhan[1], R. J. Jackson[1]

LCTU Liverpool Cancer Trials Unit

Real-World Vinflunine Outcomes in Bladder Cancer in a Single-Institution Study: Moving Beyond Clinical Trials

Guillaume Moriceau,[1] Alexis Vallard,[2] Romain Rivoirard,[1] Benoîte Méry,[1] Sophie Espenel,[2] Julien Langrand-Escure,[2] Majed Ben Mrad,[2] Guoping Wang,[2] Peng Diao,[2] Cécile Pacaut,[1] Aline Guillot,[1] Olivier Collard,[1] Pierre Fournel,[1] Nicolas Magné[2]

Clinical Genitourinary Cancer, Vol. 13, No. 6, 588-92 © 2015 Elsevier Inc.

Effectiveness, toxicity, and economic evaluation of vinflunine for the treatment of patients with transitional cell carcinoma in the Spanish outpatient setting

Beatriz Guglieri-López[a], Alejandro Pérez-Pitarch[b], Begoña Porta-Oltra[a], Francisco Ferriols-Lisart[b], Mónica Climente-Martí[a] and Manuel Alós-Almiñana[b]

Anti-Cancer Drugs 2015, **26**:860–865

Treatment of relapsed urothelial bladder cancer with vinflunine: real-world evidence by the Hellenic Genitourinary Cancer Group

Nikolaos Pistamaltzian[a], Kimon Tzannis[b], Vassiliki Pissanidou[e], Stavros Peroukidis[g], Georgia Milaki[h], Vasilis Karavasilis[e], Iraklis Mitsogiannis[c], Ioannis Varkarakis[c], Athanasios Papatsoris[c], Athanasios Dellis[d], Ioannis Adamakis[d], Konstantinos Stravodimos[d], Dimitra Molyva[f], Ilias Athanasiadis[a], Nikos Androulakis[h], Charalambos Andreadis[f], Charalambos Kalofonos[g], Dionisios Mitropoulos[d], Charalambos Deliveliotis[c], Constantinos Constantinides[d], Meletios A. Dimopoulos[b] and Aristotelis Bamias[b]

Anti-Cancer Drugs 2015, **00**:000–000

Safety and effectiveness of vinflunine in patients with metastatic transitional cell carcinoma of the urothelial tract after failure of one platinum-based systemic therapy in clinical practice

Daniel Castellano[1], Javier Puente[2], Guillermo de Velasco[3], Isabel Chirivella[4], Pilar López-Criado[5], Nicolás Mohedano[6], Ovidio Fernández[7], Icíar García-Carbonero[8], María Belén González[9] and Enrique Grande[10*]

BMC Cancer 2014, **14**:779

POSTER P268 (ECCO 2013)

HISTORICAL DATA IN REAL LIFE FROM PATIENTS TREATED BY VINFLUNINE FOR ADVANCED OR METASTATIC UROTHELIAL CARCINOMA (UC): RESULTS OF THE CURVE STUDY.

J.Médioni[1], A.Guillot[2], D.Spaeth[3], M.Di Palma[4], C.Théodore[5]

1 - Georges Pompidou Hospital, Paris, France; 2 - Lucien Neuwirth Cancerology Institute, Saint Priez en Jarez, France; 3 - Gentilly Oncology Centre, Nancy, France; 4 - Gustave Roussy Institute, Villejuif, France; 5 - Foch Hospital, Suresnes France

POSTER P130 (ECC 2015)

EFFECTIVENESS AND POSSIBLE MOLECULAR FACTORS PREDICTIVE OF CLINICAL OUTCOMES IN PATIENTS WITH TRANSITIONAL CELL CARCINOMA OF THE UROTHELIAL TRACT (TCCU) TREATED WITH VINFLUNINE: A MULTICENTER RETROSPECTIVE STUDY (MOVIE) OF THE GRUPPO ONCOLOGICO ITALIANO DI RICERCA CLINICA (GOIRC)

R. Passalacqua[1], R. Montironi[2], S. Lazzarelli[1], M. Donini[1], B. Perrucci[1], F. Nolè[3], G.L. Ceresoli[4], S. Pignata[5], F. Torricelli[6], P. Giannatempo[7], L. Doni[6], M. Ungari[1], S. Panni[1], A. Necchi[7], E. Betri[1], U. De Giorgi[8], R. Sabbatini[9], E. Rondini[10], M. Sequino[11], C. Caminiti[11].

1-Istituti Ospitalieri, Cremona, Italy. 2-University of Ancona, Italy. 3-IEO, Milano, Italy. 4-Humanitas Gavazzeni, Bergamo, Italy. 5-Istituto Tumori, Napoli, Italy. 6-Ospedale Careggi, Firenze, Italy. 7-INT, Milano, Italy. 8-IRST, Meldola, Italy. 9-Ospedale Universitario, Modena, Italy. 10-Arcispedale Santa Maria Nuova, Reggio Emilia, Italy. 11-University Hospital, Research, Parma, Italy.

The Value of Observational Cohort Studies for Cancer Drugs

Randomized controlled trials — the gold standard for clinical drug evaluation — can't always predict adverse events in real-world settings. For the new cancer therapies, observational cohort studies (OCSs) can help evaluate their effects in broader populations and provide valuable information for future clinical trials.

BY DAVID R. SPIGEL, MD **BIOTECHNOLOGY HEALTHCARE** · SUMMER 2010

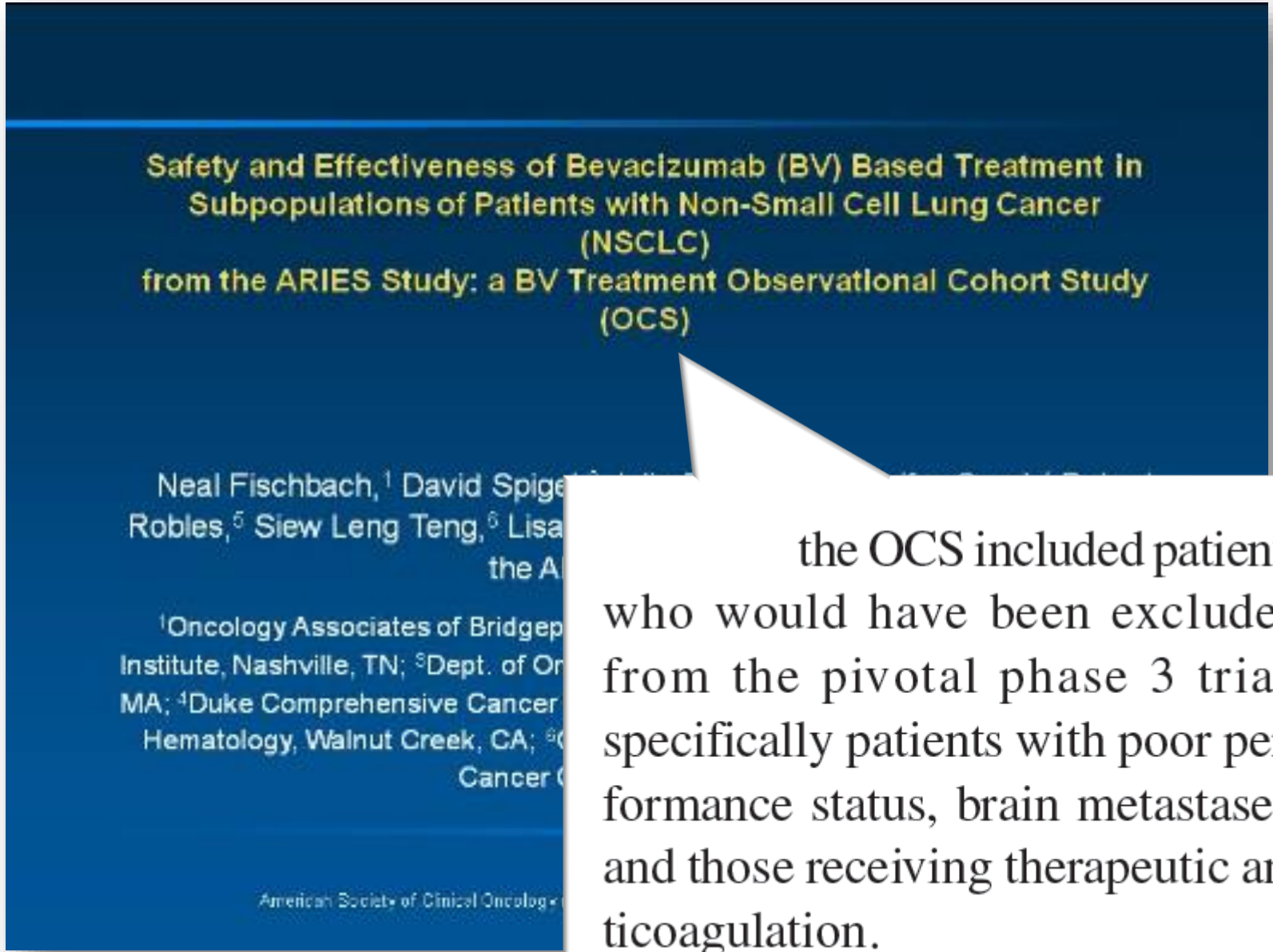
WHAT IS AN OCS?

An OCS is an analysis of a group of individuals who have specific features in common and who are followed over a defined period of time.

Prospective OCSs are designed to examine predefined primary outcomes.

Post-approval OCSs generally follow a single cohort, although patient subgroups may be analyzed separately.

To represent a broad and diverse patient base and to detect rare adverse events, large community-based, multicenter OCSs are useful in the post-approval setting for new therapeutics.



the OCS included patients who would have been excluded from the pivotal phase 3 trial, specifically patients with poor performance status, brain metastases, and those receiving therapeutic anticoagulation.

First-line bevacizumab plus taxane-based chemotherapy for locally recurrent or metastatic breast cancer: safety and efficacy in an open-label study in 2251 patients

I. E. Smith[1*], J.-Y. Pierga[2], L. Biganzoli[3], H. Cortés-Funes[4], C. Thomssen[5], X. Pivot[6], A. Fabi[7], B. Xu[8], D. Stroyakovskiy[9], F. A. Franke[10], B. Kaufman[11], P. Mainwaring[12], T. Pienkowski[13], B. De Valk[14], A. Kwong[15], J. L. González-Trujillo[16], I. Koza[17], K. Petrakova[18], D. Pereira[19] & K. I. Pritchard[20], on behalf of the ATHENA Study Group

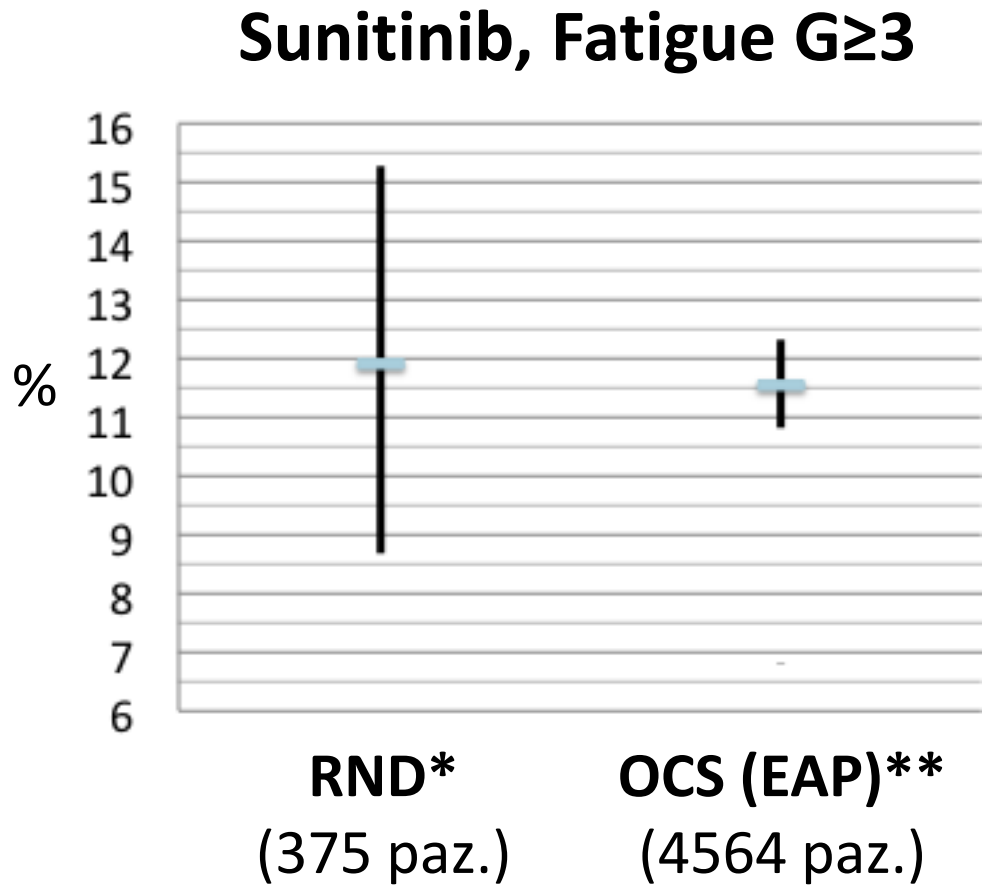
Annals of Oncology 22: 595–602, 2011

First-line Bevacizumab–Paclitaxel in 220 Patients with Metastatic Breast Cancer: Results from the AVAREG Study

MAGDOLNA DANK[1], LASZLO BUDI[2], BELA PIKO[3], LASZLO MANGEL[4], JOZSEF ERFAN[5], JOZSEF CSEH[6], AGNES RUZSA[7] and LASZLO LANDHERR[8]

ANTICANCER RESEARCH *34*: 1275-1280 (2014)

Studio RND registrativo vs OCS (EAP)



Quale dei due studi è più UTILE per la Clinica?

* Motzer, NEJM 2007; ** Gore, Lancet Oncol 2009

Observational Studies

"EPIDEMIOLOGIC" Vs "THERAPEUTIC"

Before-After

Surveys disease status before and after an intervention

Cross-Sectional

Provide information on prevalence of a particular condition at a single time point (time window)

Case-Control

Identify predictors of a particu outcome

Cohort

Identify the incidence of a particular outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions

Any observed difference between the outcomes of study arms may be attributable to baseline differences rather than to a true treatment effect.

Effectiveness Versus Efficacy: More Than a Debate Over Language

Julie M. Fritz, PT, PhD, ATC[1] *Joshua Cleland, PT, DPT, OCS*[2]

To some, the best evidence may be viewed as research that minimizes bias to the greatest extent possible, while others may prioritize research that is deemed most pertinent to clinical practice.

Qualità = eliminare (ridurre) le fonti di bias?

From Randomized Controlled Trials to Observational Studies

Stuart L. Silverman, MD

The American Journal of Medicine (2009) 122, 114-120

Table 1 Randomized Controlled Trial Methodology

Strengths	
Well-defined study population	Excludes many patients requiring clinical treatment
Design maximizes internal validity	Outcomes are difficult to extrapolate to a more general patient population
Tightly controlled treatment conditions	Short duration and modest sample sizes limit ability to identify rare or long-term adverse events
Compliance maximized through strict protocols	

Effectiveness Versus Efficacy: More Than a Debate Over Language

Julie M. Fritz, PT, PhD, ATC[1] *Joshua Cleland, PT, DPT, OCS*[2]

To some, the best evidence may be viewed as research that minimizes bias to the greatest extent possible, while others may prioritize research that is deemed most pertinent to clinical practice.

Qualità = trasferibilità alla pratica clinica?

From Randomized Controlled Trials to Observational Studies

Stuart L. Silverman, MD

The American Journal of Medicine (2009) 122, 114-120

Table 1 Randomized Controlled Trial Methodology

Strengths	Limitations
Well-defined study population	Excludes many patients requiring clinical treatment
Design maximizes internal validity	Outcomes are difficult to extrapolate to a more general patient population
Tightly controlled treatment conditions	Short duration and modest sample sizes limit ability to identify rare or long-term adverse events
Compliance maximized through strict protocols	

Integrating real-life studies in the global therapeutic research framework

*Nicolas Roche, Helen K Reddel, Alvar Agusti, Eric D Bateman, Jerry A Krishnan, Richard J Martin, Alberto Papi, Dirkje Postma, Mike Thomas, Guy Brusselle, Elliot Israel, Cynthia Rand, Alison Chisholm, David Price, on behalf of the Respiratory Effectiveness Group

www.thelancet.com/respiratory **Vol 1 December 2013**

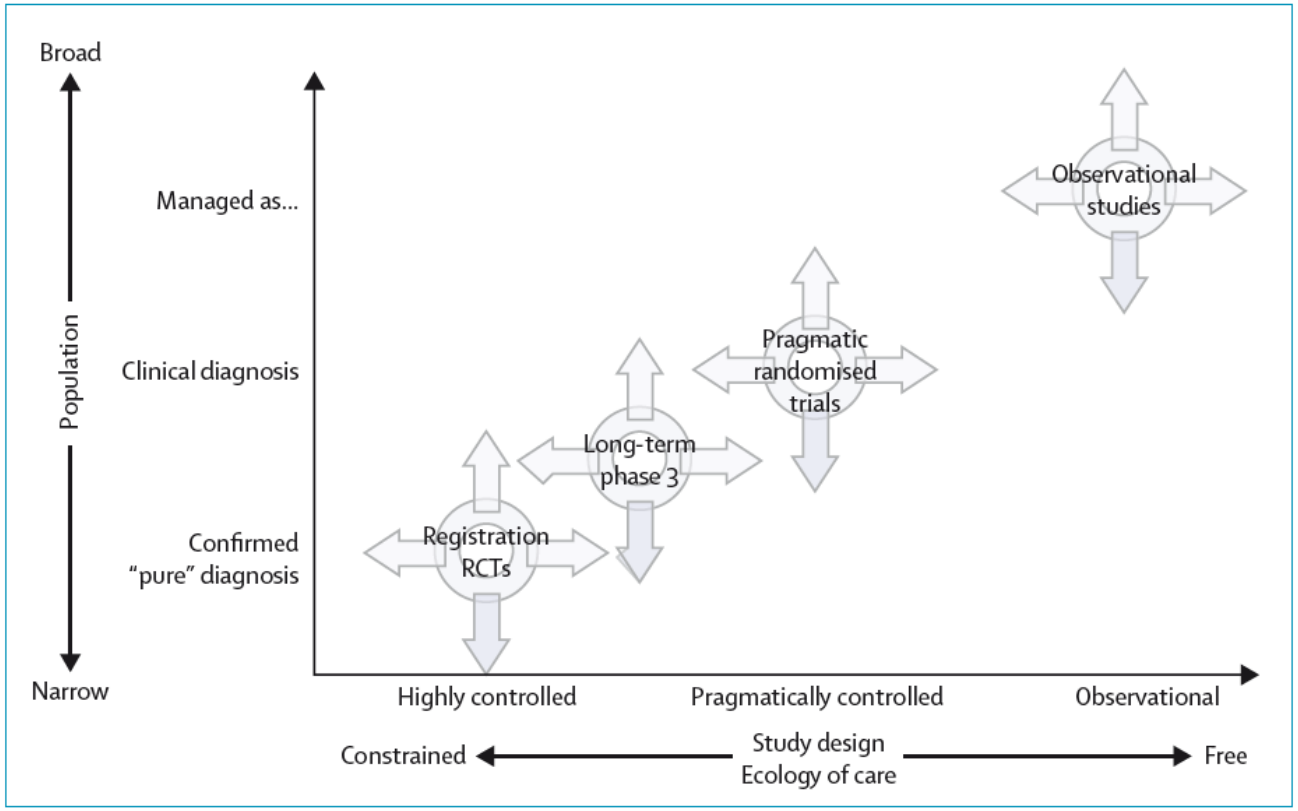


Figure 1: **A conceptual framework for therapeutic research**

Opportunities and challenges in using real-world data for health care

Vivek A. Rudrapatna[1,2] and Atul J. Butte[1,3,4]

J Clin Invest. 2020;130(2):565–574.

The what and why of real-world data

The United States Food and Drug Administration (FDA) has been using the phrases "real-world data" (RWD) to mean "data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources". They define "real-world evidence" (RWE) as "clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD".

…

This definition actually covers a broader variety of data sources… RWE covers a broader range of analytic designs for causal inference.

Tutte le prove sono real world evidence

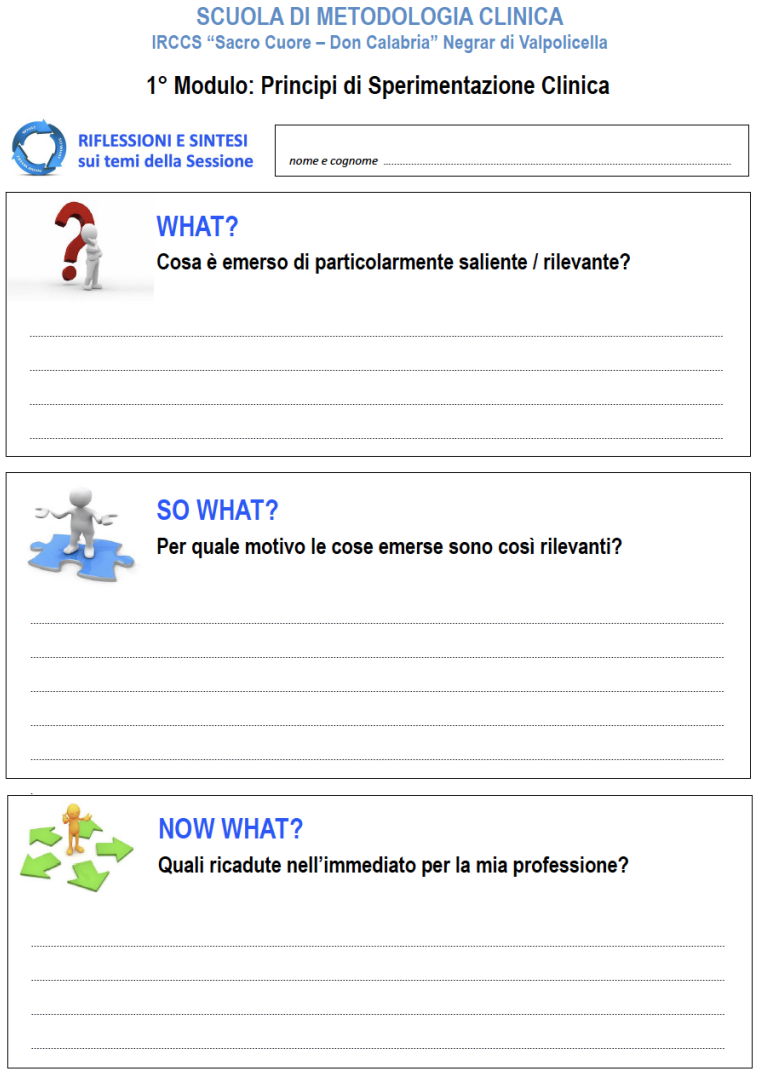
HOLGER SCHÜNEMANN

Recenti Prog Med 2019; 110: 165-167

Tutte le prove – se di buona qualità – sono “real world evidence”, sia che provengano da RCT, sia che giungano da studi osservazionali o da altri disegni di studio
Chi lavora alle linee guida dovrebbe valutare se i big data, i dati generati dai pazienti o i dati dei registri o delle cartelle cliniche elettroniche possono mitigare le principali preoccupazioni sollevate dalle prove che risultano da RCT per informare le decisioni: principalmente *indirectness*, *imprecision* e uso di prove appropriate per informare i fattori che influenzano la fiducia nelle raccomandazioni.

Dovrebbero quindi valutare la certezza delle prove, indipendentemente dalla loro etichetta. L’approccio GRADE fornisce già una guida dettagliata per ottenere questi risultati.

1. Riflettete da soli per 10 min. e compilate il form →
2. Confrontatevi con i Colleghi del Vostro tavolo per 15 min., declinate un W^3 condiviso e delegate un portavoce
3. Riportate sulla lavagna il Vostro W^3 condiviso su almeno due aspetti ritenuti rilevanti e impattanti sulla professione (in 5 min.)
4. Presentate ai Colleghi degli altri tavoli il Vostro W^3 condiviso



SCUOLA DI METODOLOGIA CLINICA

IRCCS "Sacro Cuore – Don Calabria" Negrar di Valpolicella

1° Modulo: Principi di Sperimentazione Clinica

RIFLESSIONI E SINTESI sui temi della Sessione

nome e cognome .....

WHAT?

Cosa è emerso di particolarmente saliente / rilevante?

SO WHAT?

Per quale motivo le cose emerse sono così rilevanti?

NOW WHAT?

Quali ricadute nell'immediato per la mia professione?

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- **Indicatori di effetto**
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

VARIABILE DI RISPOSTA

- di tipo **qualitativo (nominale)**
  - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- di tipo quantitativo (intervallare)
  - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- del tipo “tempo a evento”
  - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

Encorafenib, Binimetinib, and Cetuximab in *BRAF* V600E–Mutated Colorectal Cancer

S. Kopetz, A. Grothey, R. Yaeger, E. Van Cutsem, J. Desai, T. Yoshino, H. Wasan, F. Ciardiello, F. Loupakis, Y.S. Hong, N. Steeghs, T.K. Guren, H.-T. Arkenau, P. Garcia-Alfonso, P. Pfeiffer, S. Orlov, S. Lonardi, E. Elez, T.-W. Kim, J.H.M. Schellens, C. Guo, A. Krishnan, J. Dekervel, V. Morris, A. Calvo Ferrandiz, L.S. Tarpgaard, M. Braun, A. Gollerkeri, C. Keir, K. Maharry, M. Pickard, J. Christy-Bittel, L. Anderson, V. Sandor, and J. Tabernero

N Engl J Med 2019;381:1632-43.

Table 2. Tumor Response in Patients with Metastatic Colorectal Cancer with the *BRAF* V600E Mutation.*

Variable	Triplet Regimen (N=111)	Doublet Regimen (N=113)	Control (N=107)
Best overall response — no. (%)			
Complete response	4 (4)	6 (5)	0
Partial response	25 (23)	17 (15)	2 (2)
Stable disease†	47 (42)	61 (54)	31 (29)
Progressive disease	11 (10)	8 (7)	36 (34)

Risks and Odds

- ***Risk*** (proportion of persons with disease = *cumulative incidence*)
  - *Risk Ratio* = ratio of 2 cumulative incidence estimates = *Relative Risk*
- ***Odds*** (the number of events divided by the number of non events)
  - ***Odds Ratio*** = ratio of 2 odds

Incidenza Cumulativa

Probabilità (rischio) di sviluppare la malattia in uno specifico periodo di tempo *t*

- assume follow-up completo
- è una proporzione perciò può assumere valori da 0 ad 1
- deve riferirsi ad uno specifico periodo di tempo

$$IC = \frac{\text{N° di persone che ammalano tra il } t_0 \text{ e } t_1}{\text{N° di persone non malate all'inizio del periodo } t_0}$$

Es. 5 si ammalano / 10 inizialmente non malati = 0.5

Risks and Odds

- ***Risk*** (proportion of persons with disease = *cumulative incidence*)
  - ***Risk Ratio*** = ratio of 2 cumulative incidence estimates = *Relative Risk*
- ***Odds*** (the number of events divided by the number of non events)
  - ***Odds Ratio*** = ratio of 2 odds

Risks and Odds

	Events	Non-event	Total
Experimental	a: 29	b: 82	N1: 111
Control	c: 2	d: 105	N2: 107
Total	N+: 31	N−: 187	N: 218

$$\frac{(29/111)}{(2/107)}$$

Risk Ratio	CI Start	CI End
13.9775	3.4191	57.1412

$$\frac{(29/82)}{(2/105)}$$

Odds Ratio	CI Start	CI End
18.5671	4.3045	80.0867

Risks and Odds

- ***Risk*** (proportion of persons with disease = *cumulative incidence*)
  - *Risk Ratio* = ratio of 2 cumulative incidence estimates = *Relative Risk*
- ***Odds*** (the number of events divided by the number of non events)
  - ***Odds Ratio*** = ratio of 2 odds

Risks and Odds

- lative
- tes =
- Odds (th events divided by the number of non events)
  - ***Odds Ratio*** = ratio of 2 odds

Odds Ratios are used to compare the occurrence of the outcome of interest (e.g. disease or unfavourable event), given exposure to the variable of interest (e.g. health characteristic, or intervention).

Most commonly used in **case-control studies**

Risks and Odds

	Events	Non-event	Total
Experimental	a: 29	b: 82	N1: 111
Control	c: 2	d: 105	N2: 107
Total	N+: 31	N−: 187	N: 218

$$\frac{(29/111)}{(2/107)}$$

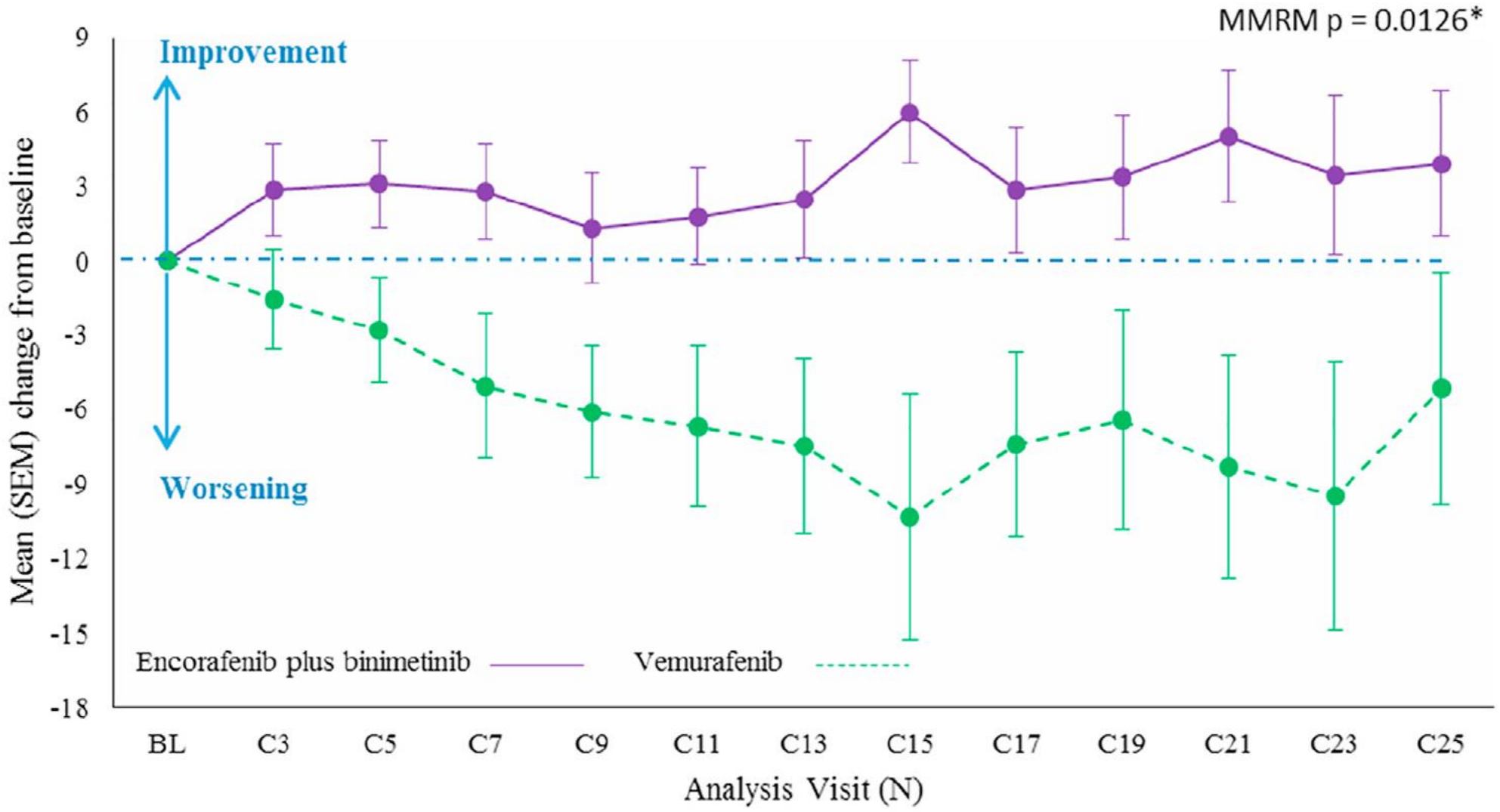
Risk Ratio	CI Start	CI End
13.9775	3.4191	57.1412

$$\frac{(29/82)}{(2/105)}$$

Odds Ratio	CI Start	CI End
18.5671	4.3045	80.0867

VARIABILE DI RISPOSTA

- di tipo qualitativo (nominale)
  - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- di tipo **quantitativo (intervallare)**
  - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- del tipo “tempo a evento”
  - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.



Mean change from baseline over time in the EORTC QLQ-C30 global health status

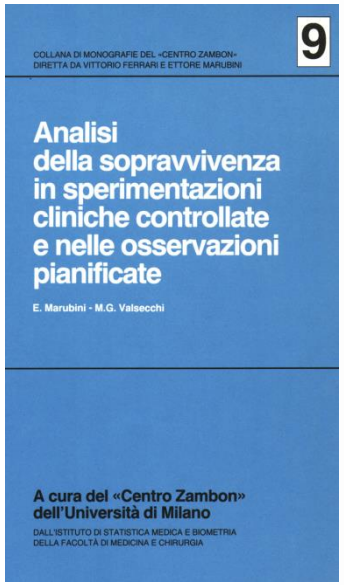
H. Gogas et al. / European Journal of Cancer 152 (2021) 116–128

VARIABILE DI RISPOSTA

- di tipo qualitativo (nominale)
  - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- di tipo quantitativo (intervallare)
  - assume uno spettro continuo di valori e viene misurata in riferimento a una scala a intervalli costanti.
- del tipo “**tempo a evento**”
  - rappresenta il tempo trascorso fino al verificarsi (o meno) di un evento.

VARIABILE “TEMPO A EVENTO”

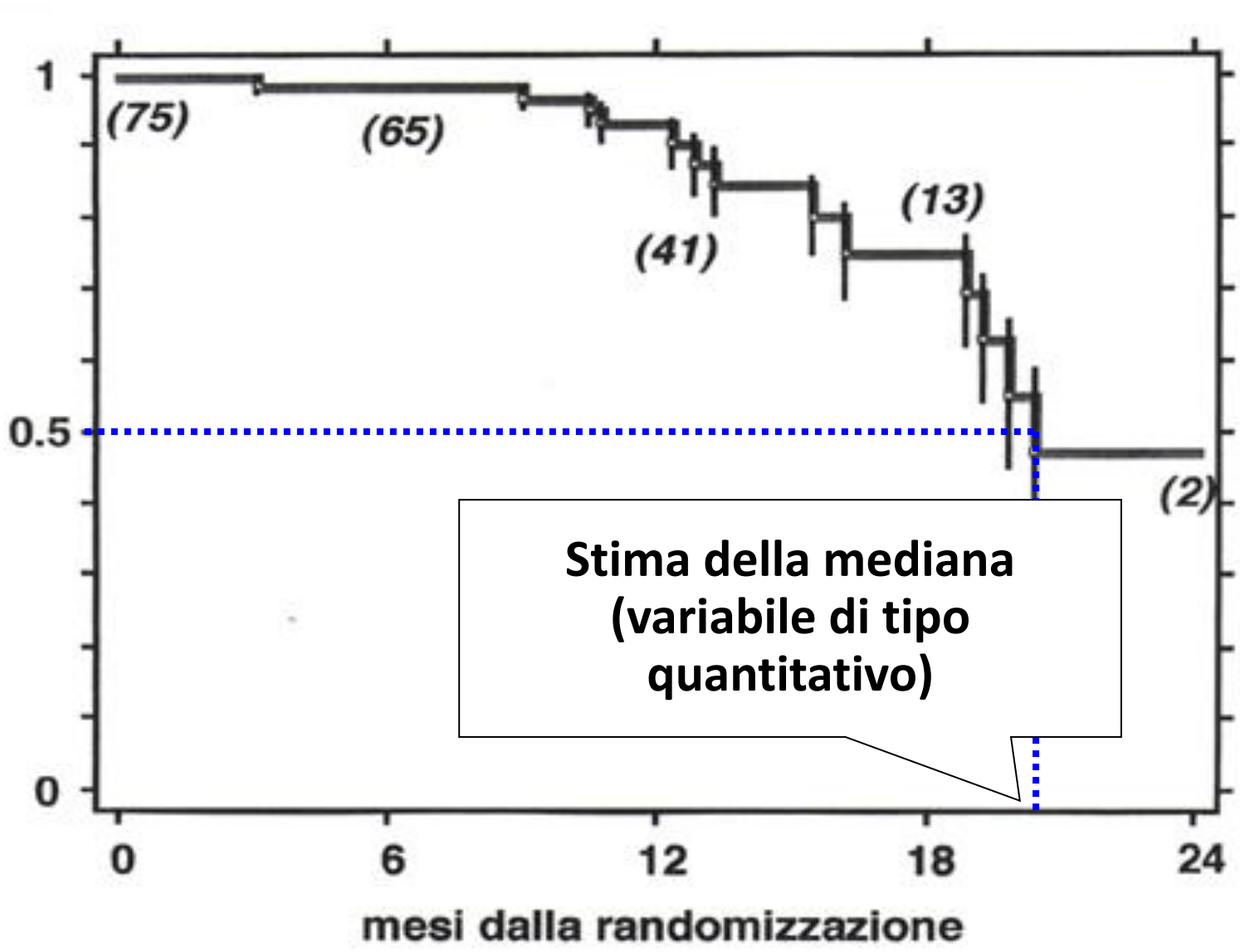
- Apparentemente assimilabile a una variabile di tipo quantitativo (intervallare).
- Ma il verificarsi o meno di un evento la rende assimilabile a una variabile di tipo qualitativo (nominale)
- In alcuni soggetti inoltre l’evento di interesse potrebbe non essersi ancora verificato al momento della analisi
- Tali risultati vengono quindi meglio rappresentati come stima della funzione di sopravvivenza:
  - probabilità di sopravvivere oltre un determinato tempo, misurato dalla data di inizio dell'osservazione.



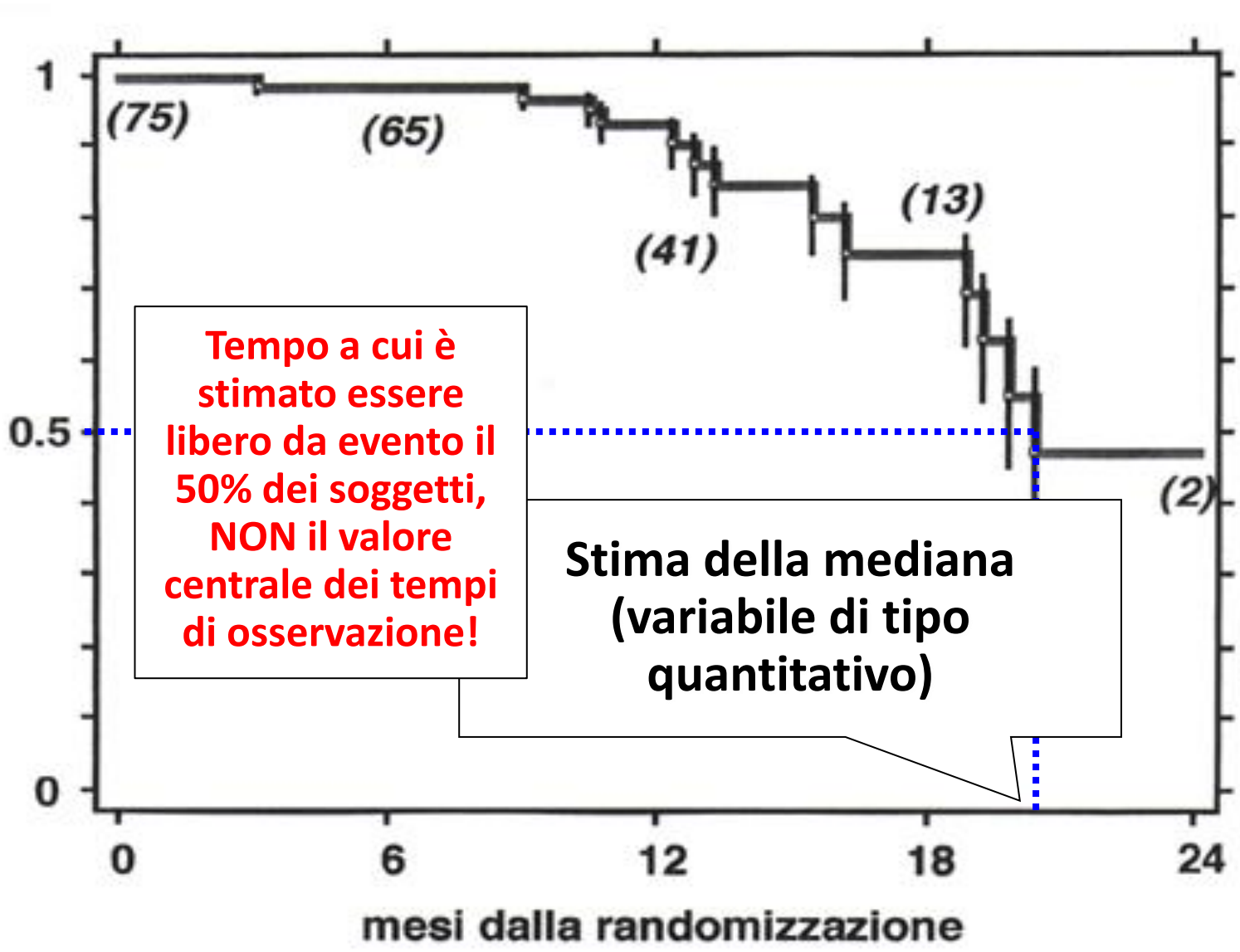
Tempi di risposta $t_{(j)}$	Tempi troncati* t^*	N° soggetti esposti a rischio n_j	N° eventi terminali d_j	Rischio istantaneo di "morte" $\hat{\lambda}(t_{(j)})$	Probabilità cumulativa di sopravvivere $t_{(j)}$ $\hat{P}_j$
9		20	1	1/20 = .050	(1 – 1/20) × 1 = .9500
13		19	1	1/19 = .053	(1 – 1/19) × .9500 = .8996
20		18	1	1/18 = .055	(1 – 1/18) × .8996 = .8501
26		17	1	1/17 = .059	(1 – 1/17) × .8501 = .7999
27		16	1	1/16 = .062	(1 – 1/16) × .7999 = .7503
28		15	1	1/15 = .067	(1 – 1/15) × .7503 = .7000
30		14	1	1/14 = .071	(1 – 1/14) × .7000 = .6503
32		13	2	2/13 = .154	(1 – 2/13) × .6503 = .5502
75		11	1	1/11 = .091	(1 – 1/11) × .5502 = .5001
79		10	1	1/10 = .100	(1 – 1/10) × .5001 = .4501
91		9	1	1/9 = .111	(1 – 1/9) × .4501 = .4001
	177*	8	0	0/8 = .0	(1 – 0/8) × .4001 = .4001
193		7	1	1/7 = .143	(1 – 1/7) × .4001 = .3429
541		6	1	1/6 = .167	(1 – 1/6) × .3429 = .2856
1129		5	1	1/5 = .200	(1 – 1/5) × .2856 = .2285
	1499*	4	0	0/4 = .0	(1 – 0/4) × .2285 = .2285
1585		3	1	1/3 = .333	(1 – 1/3) × .2285 = .1524

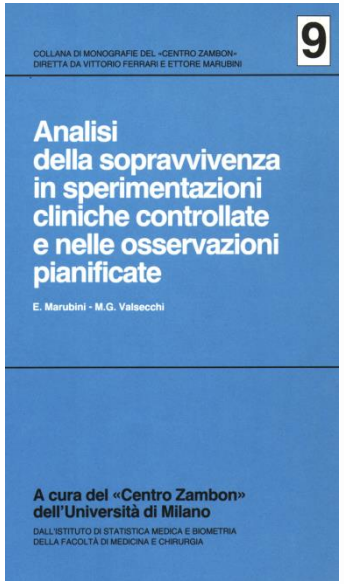
TABELLA 10.
Calcolo secondo Kaplan e Meier della curva di sopravvivenza del rene trapiantato nei pazienti di tabella 6.

CURVA DI SOPRAVVIVENZA



CURVA DI SOPRAVVIVENZA





Tempi di risposta $t_{(j)}$	Tempi troncati* t^*	N° soggetti esposti a rischio n_j	N° eventi terminali d_j	Rischio istantaneo di "morte" $\hat{\lambda}(t_{(j)})$	Probabilità cumulativa di sopravvivere $t_{(j)}$ $\hat{P}_j$
9		20	1	1/20 = .050	(1 – 1/20) × 1 = .9500
13		19	1	1/19 = .053	(1 – 1/19) × .9500 = .8996
20		18	1	1/18 = .055	(1 – 1/18) × .8996 = .8501
26		17	1	1/17 = .059	(1 – 1/17) × .8501 = .7999
27		16	1	1/16 = .062	(1 – 1/16) × .7999 = .7503
28		15	1	1/15 = .067	(1 – 1/15) × .7503 = .7000
30		14	1	1/14 = .071	(1 – 1/14) × .7000 = .6503
32		13	2	2/13 = .154	(1 – 2/13) × .6503 = .5502
75		11	1	1/11 = .091	(1 – 1/11) × .5502 = .5001
79		10	1	1/10 = .100	(1 – 1/10) × .5001 = .4501
91		9	1	1/9 = .111	(1 – 1/9) × .4501 = .4001
	177*	8	0	0/8 = .0	(1 – 0/8) × .4001 = .4001
193		7	1	1/7 = .143	(1 – 1/7) × .4001 = .3429
541		6	1	1/6 = .167	(1 – 1/6) × .3429 = .2856
1129		5	1	1/5 = .200	(1 – 1/5) × .2856 = .2285
	1499*	4	0	0/4 = .0	(1 – 0/4) × .2285 = .2285
1585		3	1	1/3 = .333	(1 – 1/3) × .2285 = .1524

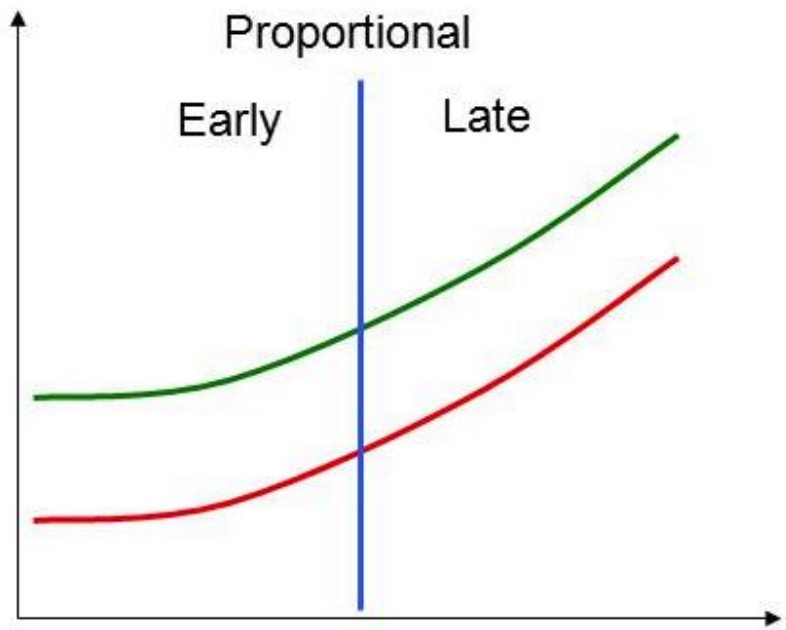
TABELLA 10.
Calcolo secondo Kaplan e Meier della curva di sopravvivenza del rene trapiantato nei pazienti di tabella 6.

Indicatori comparativi di effetto di variabili tempo-a-evento

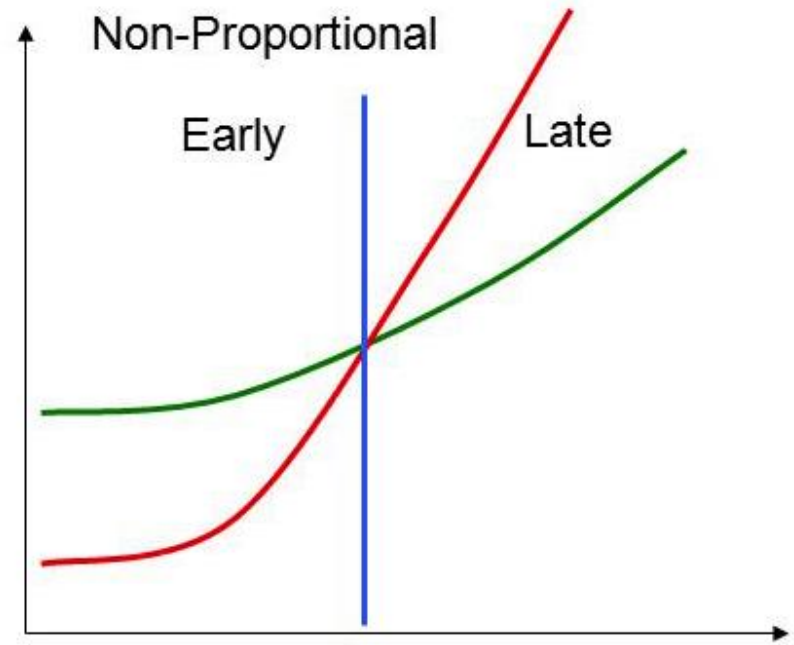
- Differenza tra stime della mediana di sopravvivenza (KM)
- Differenza media di sopravvivenza (*restricted means*) al tempo *t*
- Differenza tra stime di sopravvivenza (KM) al tempo *t* (*Milestone* Survival)
- Hazard Ratio (KM+Cox)

Proportional Hazard Assumption

If we are comparing a new treatment with the standard treatment, it is assumed that the ratio of the hazard for an individual on a new treatment to that for an individual on the standard treatment remains constant over time

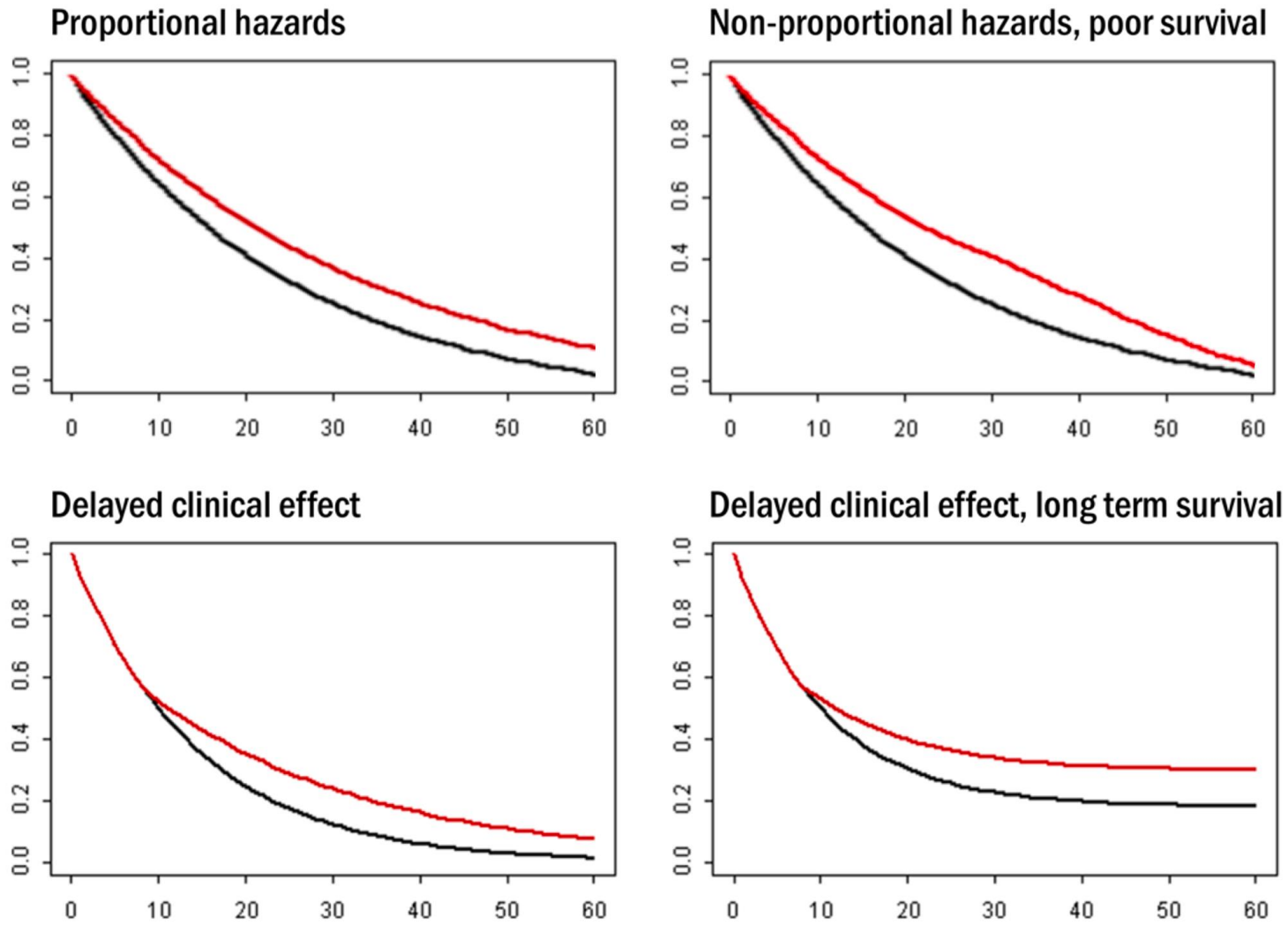


Here, the effect is the same in both time periods



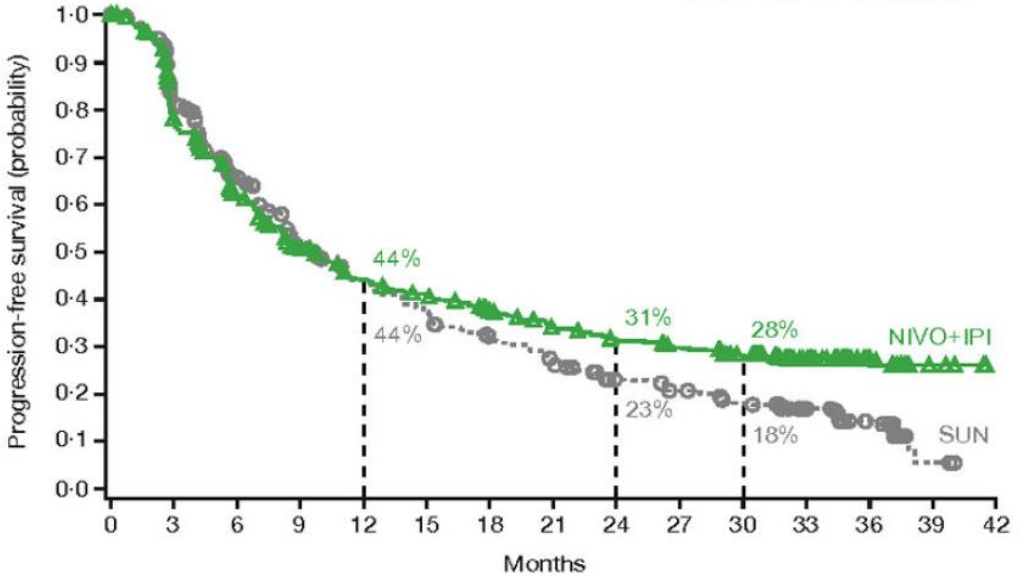
Here, the effect is negative in the early period and positive in the late period

Hazard Ratio ALWAYS a proper outcome measure?



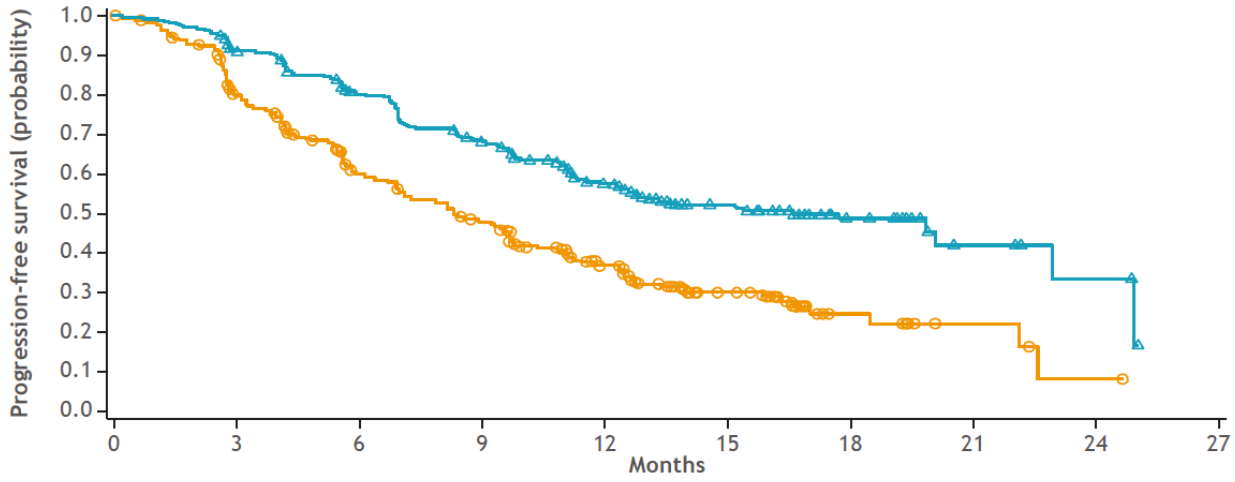
modif. from Chen et al. J Immunother Cancer 2013, 1:18

Hazard Ratio ALWAYS a proper outcome measure?



CM 214: PFS (ITT)

- **non proportional hazards**
- **hazard ratio likely underestimated**



CM 9ER: PFS (ITT)

- **proportional hazards**
- **hazard ratio properly estimated**

Variabili di Risposta e Indicatori Comparativi di effetto

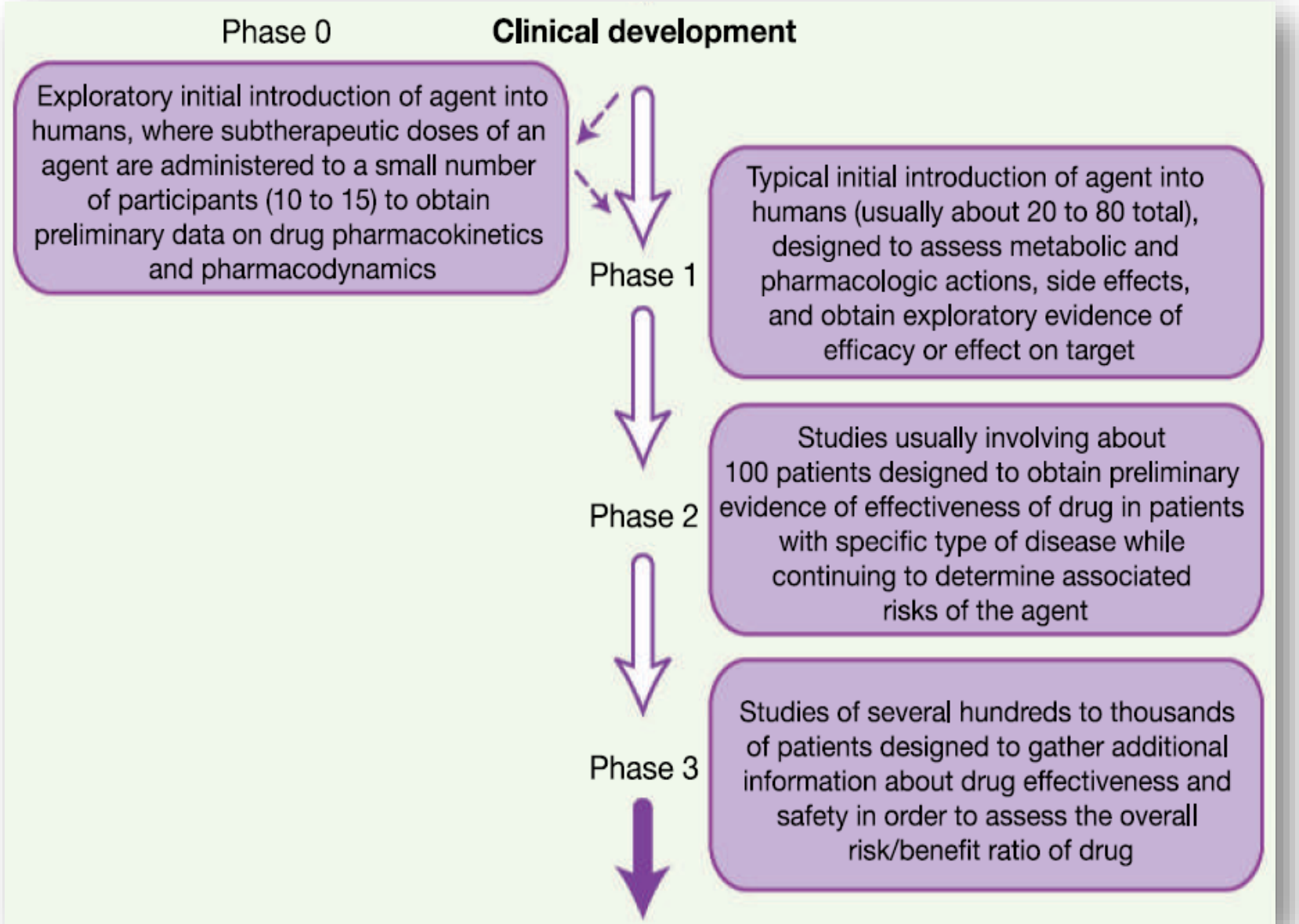
Endpoint	Tipo di variabile statistica	Indicatore comparativo di effetto	
		relativo	assoluto
Attività (CR, PR, SD, ecc.) Tossicità (eventi avversi)	nominale	RR (relative risk) OR (odds ratio)	RD (risk difference)
Punteggi di qualità di vita	intervallare	-	MD (mean difference)
OS (overall survival) PFS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

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- **Endpoints per le diverse Fasi della sperimentazione**
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

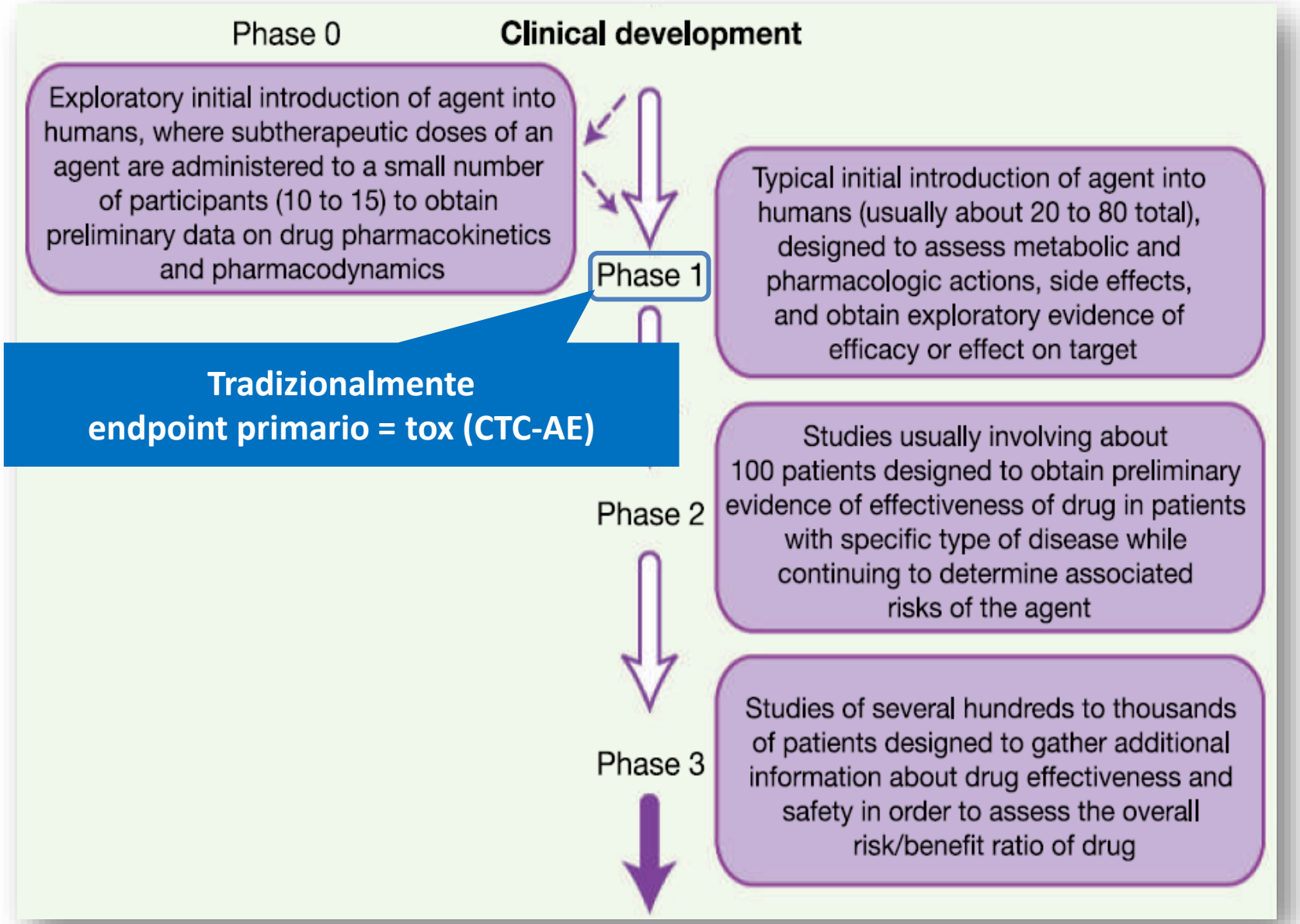
Endpoints (ICH E9)

- The primary variable ('target' variable, primary endpoint)
- Capable of providing the most clinically relevant and convincing evidence directly related to the primary objective of the trial
- A reliable and validated variable measuring some clinically relevant and important treatment benefit in the patient population

International Conference for Harmonization (ICH). Efficacy Guidelines. ICH Topic E 9, Statistical Principles for Clinical Trials



Clin Cancer Res; 16(6) March 15, 2010



Clin Cancer Res; 16(6) March 15, 2010

Common Terminology Criteria for Adverse Events v4.0 (CTCAE)

Publish Date: May 28, 2009

Quick Reference

The NCI Common Terminology Criteria for Adverse Events is a descriptive terminology which can be utilized for Adverse Event (AE) reporting. A grading (severity) scale is provided for each AE term.

Components and Organization

SOC

System Organ Class, the highest level of the MedDRA hierarchy, is identified by anatomical or physiological system, etiology, or purpose (e.g., SOC Investigations for laboratory test results). CTCAE terms are grouped by MedDRA Primary SOCs. Within each SOC, AEs are listed and accompanied by descriptions of severity (Grade).

CTCAE Terms

An Adverse Event (AE) is any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medical treatment or procedure that may or may ***not*** be considered related to the medical treatment or procedure. An AE is a term that is a unique representation of a specific event used for medical documentation and scientific analyses. Each CTCAE v4.0 term is a MedDRA LLT (Lowest Level Term).

Definitions

A brief definition is provided to clarify the meaning of each AE term.

Grades

Grade refers to the severity of the AE. The CTCAE displays Grades 1 through 5 with unique clinical descriptions of severity for each AE based on this general guideline:

Grade 1	Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.
Grade 2	Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental ADL*.
Grade 3	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care ADL**.
Grade 4	Life-threatening consequences; urgent intervention indicated.
Grade 5	Death related to AE.

A Semi-colon indicates 'or' within the description of the grade.

A single dash (-) indicates a grade is not available.

Not all Grades are appropriate for all AEs. Therefore, some AEs are listed with fewer than five options for Grade selection.

Grade 5

Grade 5 (Death) is not appropriate for some AEs and therefore is not an option.

Activities of Daily Living (ADL)

*Instrumental ADL refer to preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.

**Self care ADL refer to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden.

Issues with Novel Targeted Non-Cytotoxics

- Dose-Toxicity and Dose-Effect relationships: may not be parallel
- May not cause regression of established tumours
- Thus, for newer agents:
  - phase I trials: endpoint should be ????
  - phase II trials: endpoint should be ????

National Cancer Institute Of Canada
Institut national du cancer du Canada
Clinical Trials Group
Groupe Des Essais Cliniques

E. Eisenhauer, ESMO 2002

A Practical Approach: Phase I Design Non-Cytotoxics

- Continue to limit dose using toxicity.
- Explore alternative endpoints as part of trial including:
  - Target inhibition
  - Blood levels
- Final dose decision may be based on a composite of these.
- Further exploration of dose effects may need (randomized) phase II designs.

National Cancer Institute Of Canada
Institut national du cancer du Canada
Clinical Trials Group
Groupe Des Essais Cliniques

E. Eisenhauer, ESMO 2002

Table 1. Overall responses for all possible combinations of tumor responses in target and nontarget lesions with or without the appearance of new lesions*

Target lesions	Nontarget lesions	New lesions	Overall response
CR	CR	No	CR
CR	Incomplete response/SD	No	PR
PR	Non-PD	No	PR
SD	Non-PD	No	SD
PD	Any	Yes or no	PD
Any	PD	Yes or no	PD
Any	Any	Yes	PD

*CR = complete response; PR = partial response; SD = stable disease; and PD = progressive disease. *See* text for more details.

cal initial introduction of agent into
ans (usually about 20 to 80 total),
signed to assess metabolic and
rmacologic actions, side effects,
d obtain exploratory evidence of
efficacy or effect on target

Studies usually involving about
100 patients designed to obtain preliminary
evidence of effectiveness of drug in patients
with specific type of disease while
continuing to determine associated
risks of the agent

Phase 2

Tradizionalmente
endpoint primario = risposta (RECIST)

Phase 3

Studies of several hundreds to thousands
of patients designed to gather additional
information about drug effectiveness and
safety in order to assess the overall
risk/benefit ratio of drug

Clin Cancer Res; 16(6) March 15, 2010

Guidelines for the Evaluation of Immune Therapy Activity in Solid Tumors: Immune-Related Response Criteria

Jedd D. Wolchok,[1] Axel Hoos,[2] Steven O'Day,[3] Jeffrey S. Weber,[4] Omid Hamid,[3] Celeste Lebbé,[5] Michele Maio,[6] Michael Binder,[7] Oliver Bohnsack,[8] Geoffrey Nichol,[9] Rachel Humphrey,[2] and F. Stephen Hodi[10]

(Clin Cancer Res 2009;15(23):7412–20)

Table 1. Comparison between WHO criteria and the irRC

	WHO	irRC
New, measurable lesions (i.e., ≥5 × 5 mm)	Always represent PD	Incorporated into tumor burden
New, nonmeasurable lesions (i.e., <5 × 5 mm)	Always represent PD	Do not define progression (but preclude irCR)
Non-index lesions	Changes contribute to defining BOR of CR, PR, SD, and PD	Contribute to defining irCR (complete disappearance required)
CR	Disappearance of all lesions in two consecutive observations not less than 4 wk apart	Disappearance of all lesions in two consecutive observations not less than 4 wk apart
PR	≥50% decrease in SPD of all index lesions compared with baseline in two observations at least 4 wk apart, in absence of new lesions or unequivocal progression of non-index lesions	≥50% decrease in tumor burden compared with baseline in two observations at least 4 wk apart
SD	50% decrease in SPD compared with baseline cannot be established nor 25% increase compared with nadir, in absence of new lesions or unequivocal progression of non-index lesions	50% decrease in tumor burden compared with baseline cannot be established nor 25% increase compared with nadir
PD	At least 25% increase in SPD compared with nadir and/or unequivocal progression of non-index lesions and/or appearance of new lesions (at any single time point)	At least 25% increase in tumor burden compared with nadir (at any single time point) in two consecutive observations at least 4 wk apart

Issues with Novel Targeted Non-Cytotoxics

- Dose-Toxicity and Dose-Effect relationships: may not be parallel
- May not cause regression of established tumours
- Thus, for newer agents:
  - phase I trials: endpoint should be ????
  - phase II trials: endpoint should be ????

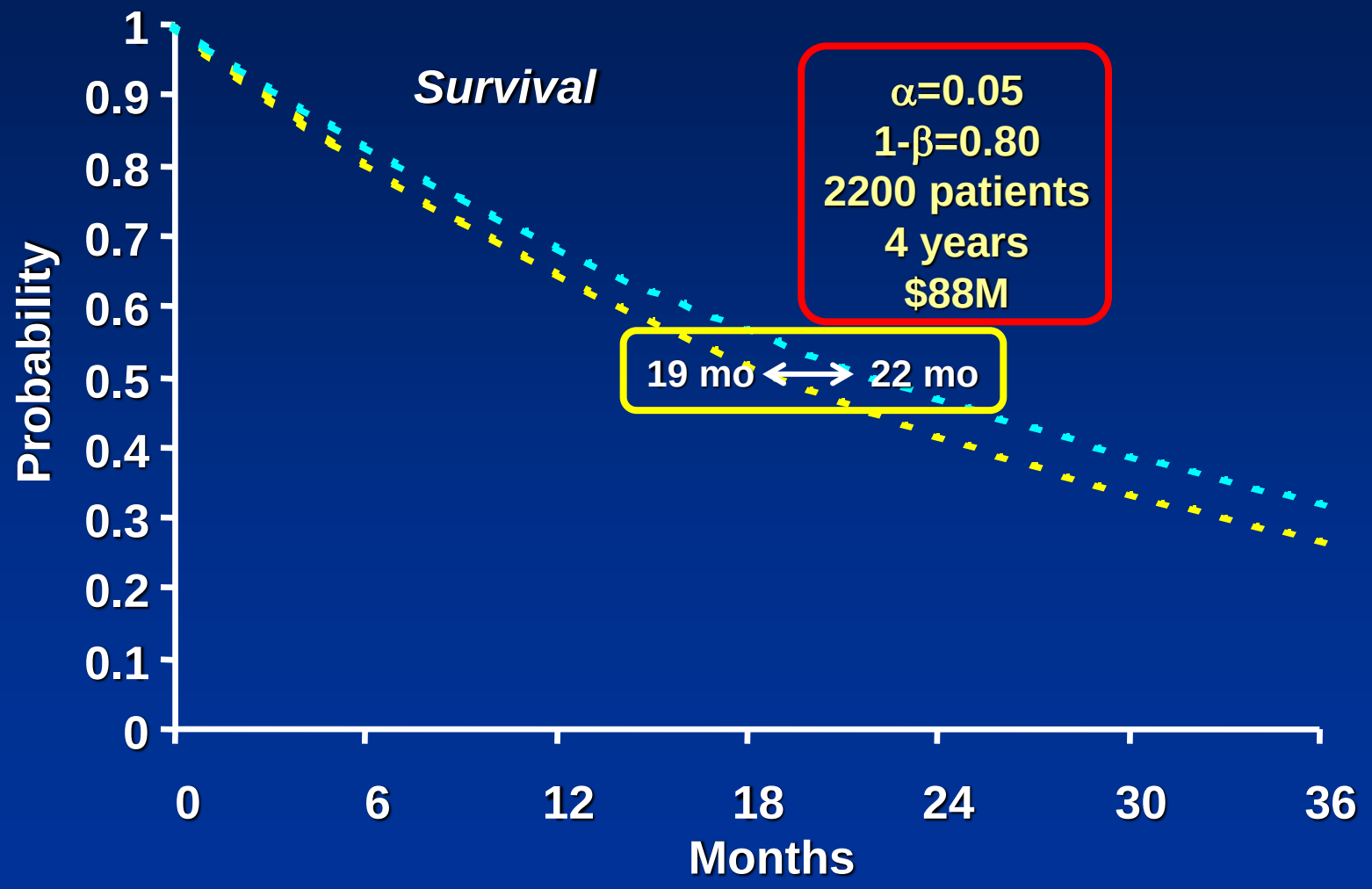
National Cancer Institute Of Canada
Institut national du cancer du Canada
Clinical Trials Group
Groupe Des Essais Cliniques

E. Eisenhauer, ESMO 2002

Regular Approval Basis – "Clinical Benefit"

- **Longer life**
- Better life
- Established Surrogate for one of above

Survival Superiority Study
Offers Too Little, Too Late, For Too Much

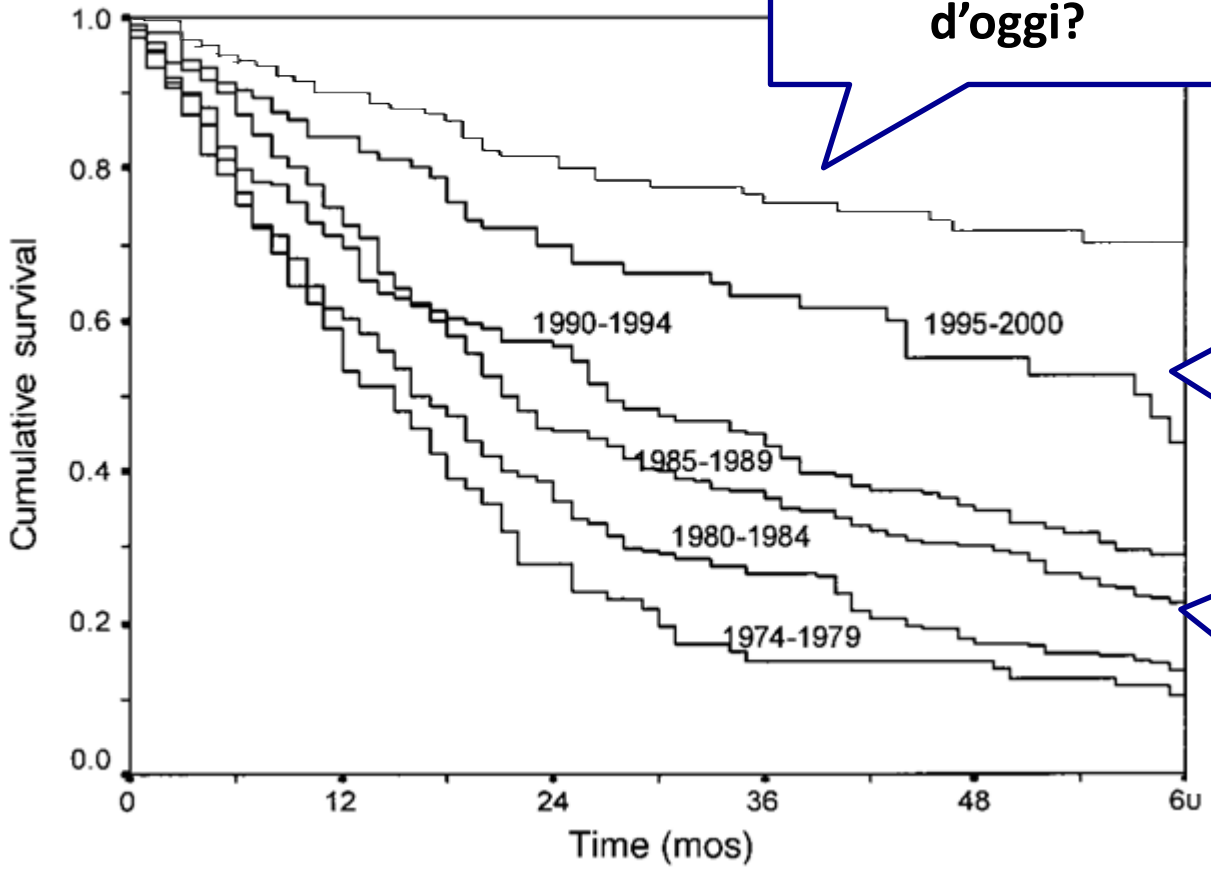


Assumes accrual = 100 patients/mo; follow-up = largest median + 2 mo (TTP) or 4 mo (survival)

L.L. Miller, 2003

Trends in Breast Cancer Survival/Giordano et al.
CANCER January 1, 2004 / Volume 100 / Number 1

Overall survival from time of recurrence

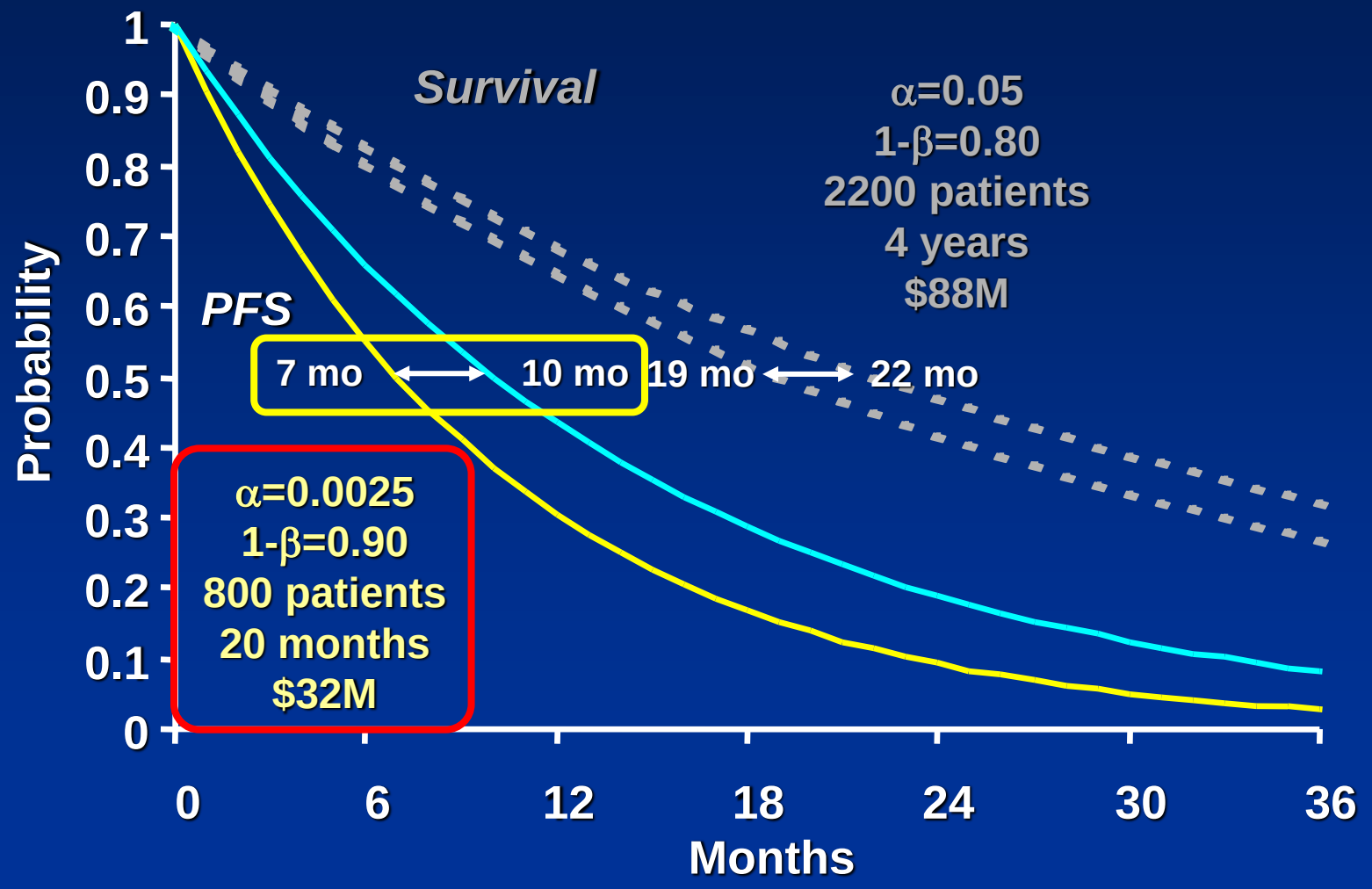


… e al giorno d'oggi?

Più difficile ottenere un beneficio relativo di Sopravvivenza

(abbastanza) facile ottenere un beneficio relativo di Sopravvivenza

Single Superiority Study Can Offer Highly Robust PFS Assessment (α=0.0025)



Assumes accrual = 100 patients/mo; follow-up = largest median + 2 mo (TTP) or 4 mo (survival)

L.L. Miller, 2003

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

The primary endpoint was the percent change from baseline in calculated LDL-C at 24 weeks with alirocumab compared [illegible]h ezetimibe.

LDL-C…

- **Endpoint di attività?**
- **Endpoint di efficacia?**
- **Endpoint "intermedio" (surrogato?)**

The Treat-to-Target Trial

Randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients

MATTHEW C. RIDDLE, MD[1] JOHN GERICH, MD[2]
JULIO ROSENSTOCK, MD[2]
ON BEHALF OF THE INSULIN GLARGINE 4002 STUDY INVESTIGATORS*

DIABETES CARE, VOLUME 26, NUMBER 11, NOVEMBER 2003

The primary outcome measure was the percentage of subjects achieving HbA_{1c} ≤7.0% without a single instance of [illegible]mp-tomatic nocturnal hypogl[illegible]on-

HbA_{1c}…

- **Endpoint di attività?**
- **Endpoint di efficacia?**
- **Endpoint "intermedio" (surrogato?)**

Regular Approval Basis – "Clinical Benefit"

- **Longer life**
- Better life
- **Established Surrogate** for one of above


Cremino
Cremino Gandola
GANDOLA
5
Cremini
Cremino

Surrogate outcome markers in research and clinical practice

Scott Twaddell

(Aust Prescr 2009;32:47–50)

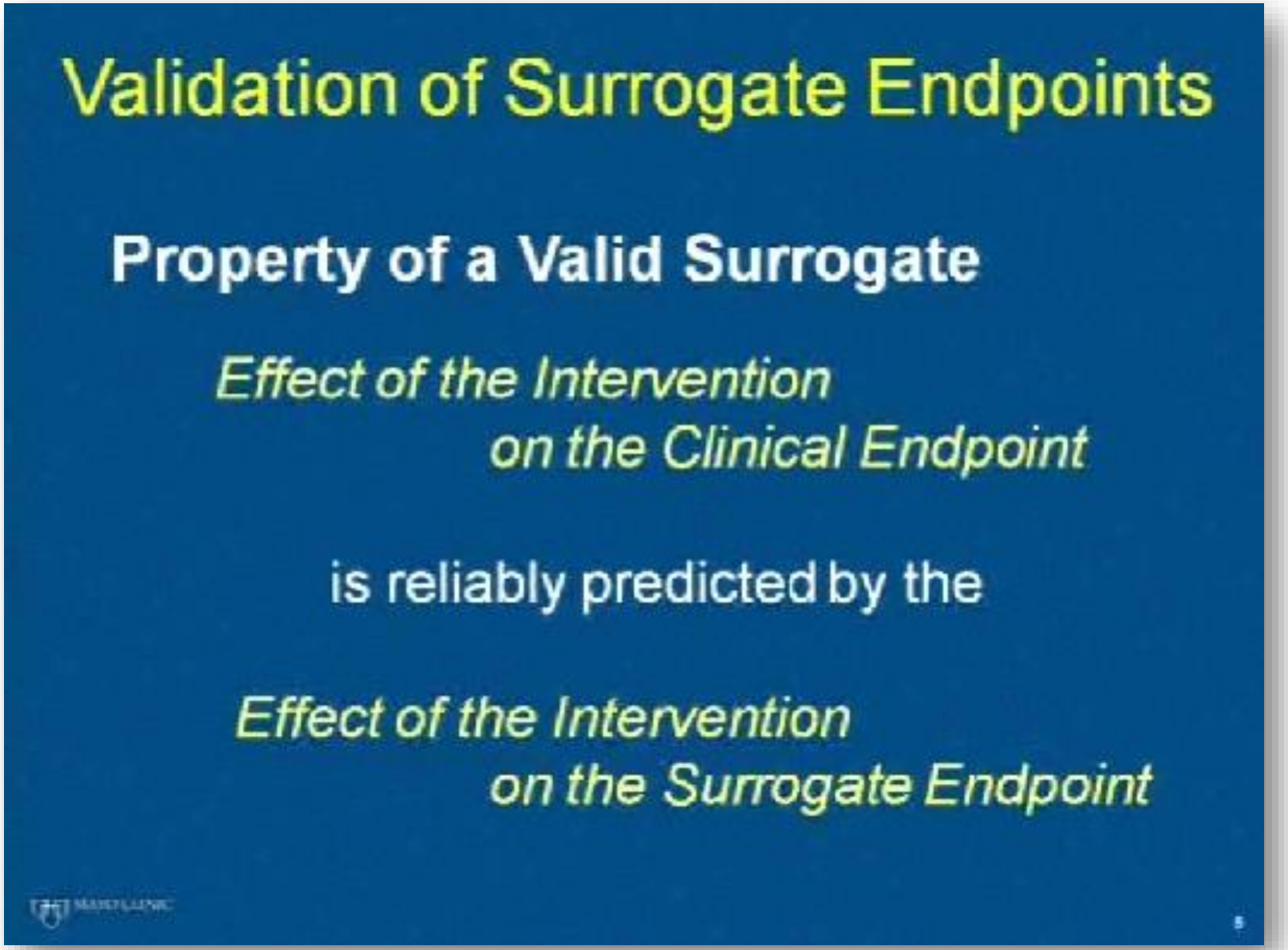
Table 1

Surrogate markers often used in clinical practice

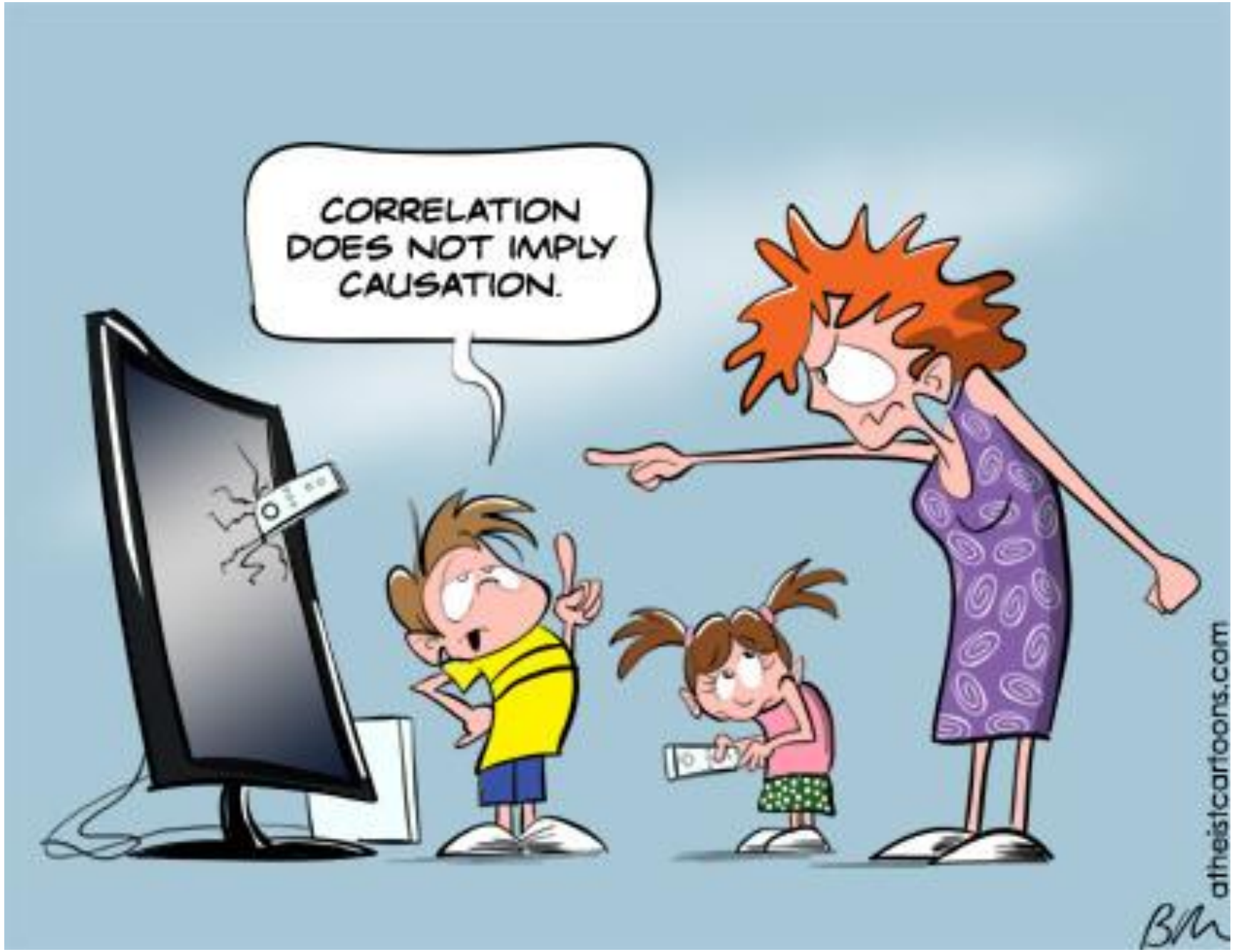
Generally accepted as valid		Doubt still exists about validity	
Surrogate marker	**Predicts**	**Surrogate marker**	**Predicts**
HbA1c	Diabetic microvascular complications	HbA1c	Diabetic macrovascular complications
FEV_1	Mortality in chronic obstructive pulmonary disease	Bone mineral density	Fracture risk
Blood pressure	Primary and secondary cardiovascular events	Prostate specific antigen	Prognosis of prostate cancer
Viral load	Survival in HIV infection	Suppression of arrhythmia	Long-term survival
Cholesterol concentration	Primary and secondary cardiovascular events	Carotid intima-media thickness	Coronary artery disease
Intraocular pressure	Visual loss in glaucoma	Albuminuria	Cardiovascular events

HbA1c glycated haemoglobin

FEV_1 forced expiratory volume in one second



D. Sargent, ASCO 2011

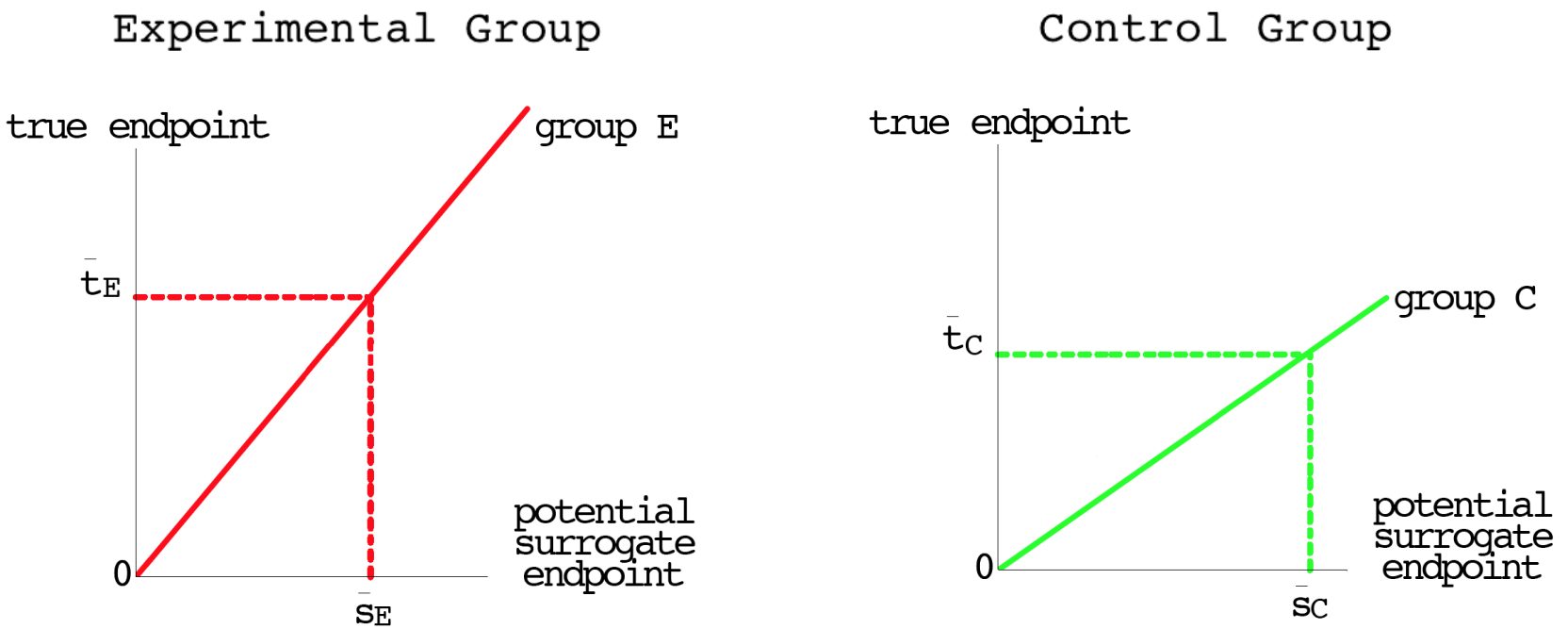

CORRELATION DOES NOT IMPLY CAUSATION.
atheistcartoons.com

A perfect correlate does not a surrogate make

Stuart G Baker*[1] and Barnett S Kramer[2]

BMC Medical Research Methodology 2003, **3**:16

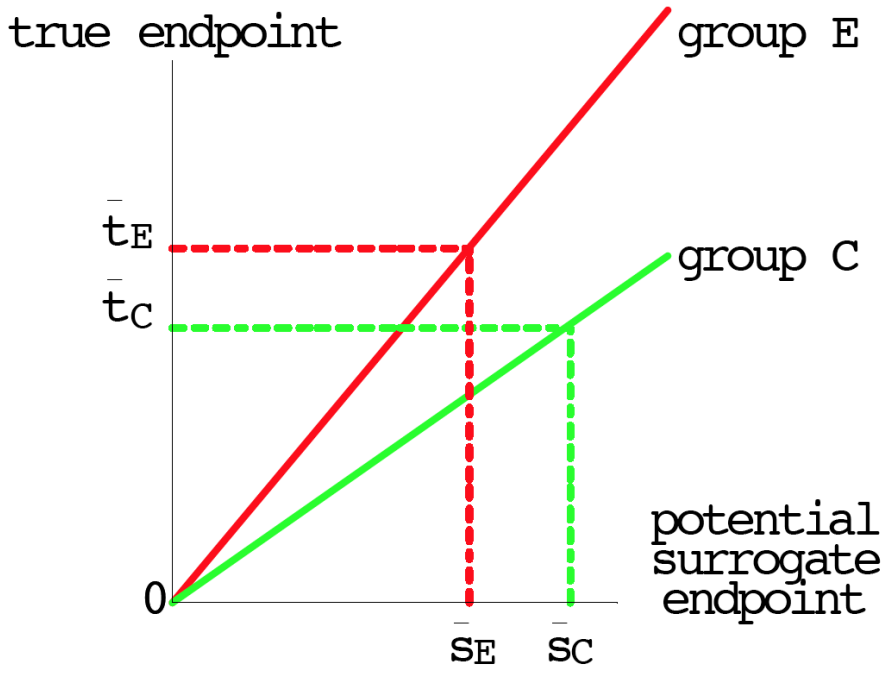
Background: There is common belief among some medical researchers that if a potential surrogate endpoint is highly correlated with a true endpoint, then a positive (or negative) difference in potential surrogate endpoints between randomization groups would imply a positive (or negative) difference in unobserved true endpoints between randomization groups.



A perfect correlate does not a surrogate make

Stuart G Baker*[1] and Barnett S Kramer[2]

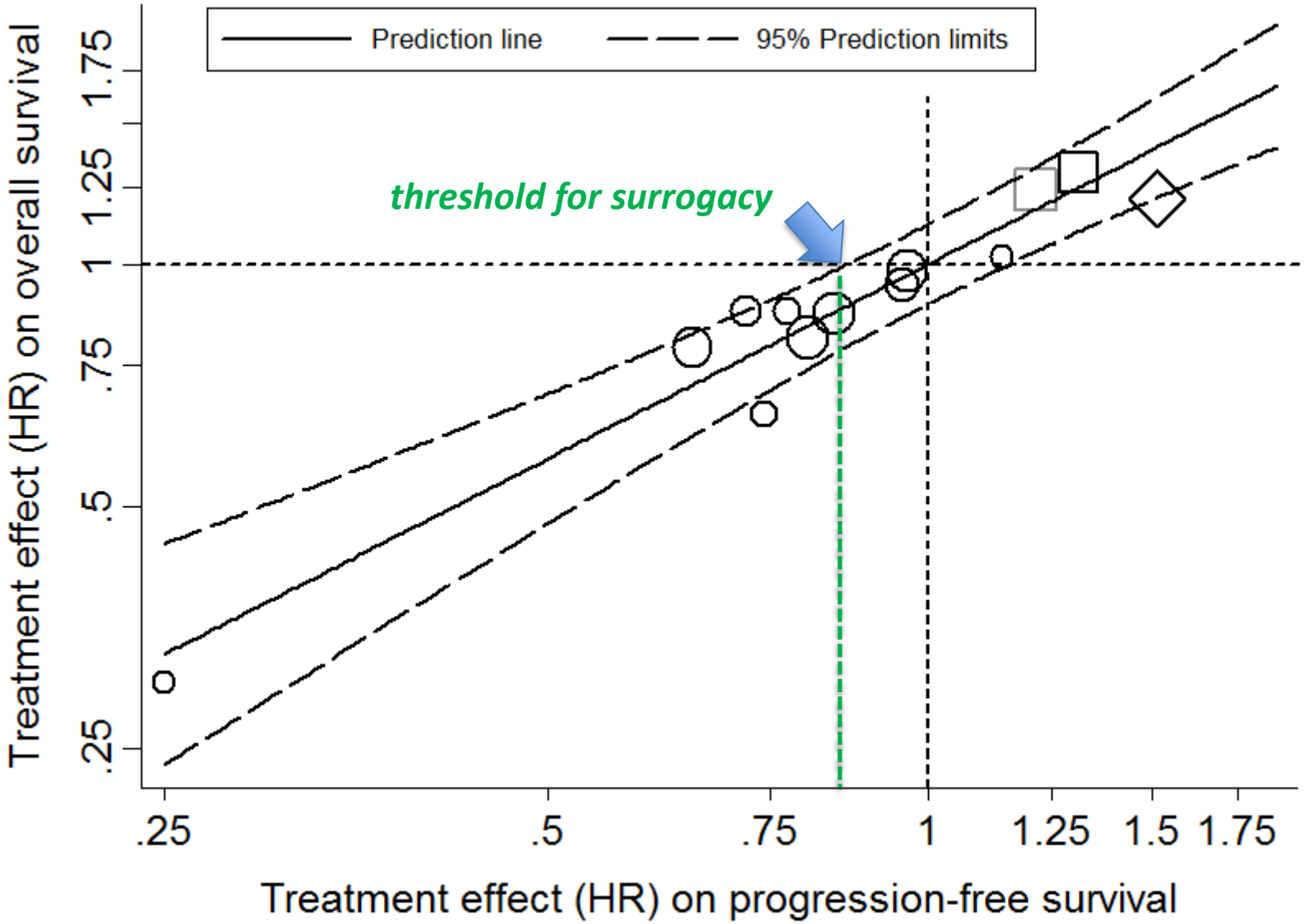
BMC Medical Research Methodology 2003, **3**:16



The mean surrogate outcome in the E group $\bar{s}_E$ is smaller than the mean surrogate outcome in the C group $\bar{s}_C$. However the mean true outcome in the E group $\bar{t}_E$ is larger than the mean true outcome in the C group $\bar{t}_C$, yielding the opposite conclusion for the effect of experimental intervention.

Conclusion: Perfect correlation between potential surrogate and unobserved true outcomes within randomized groups does not guarantee correct inference based on a potential surrogate endpoint.

TRIAL LEVEL CORRELATION BETWEEN EFFECTS

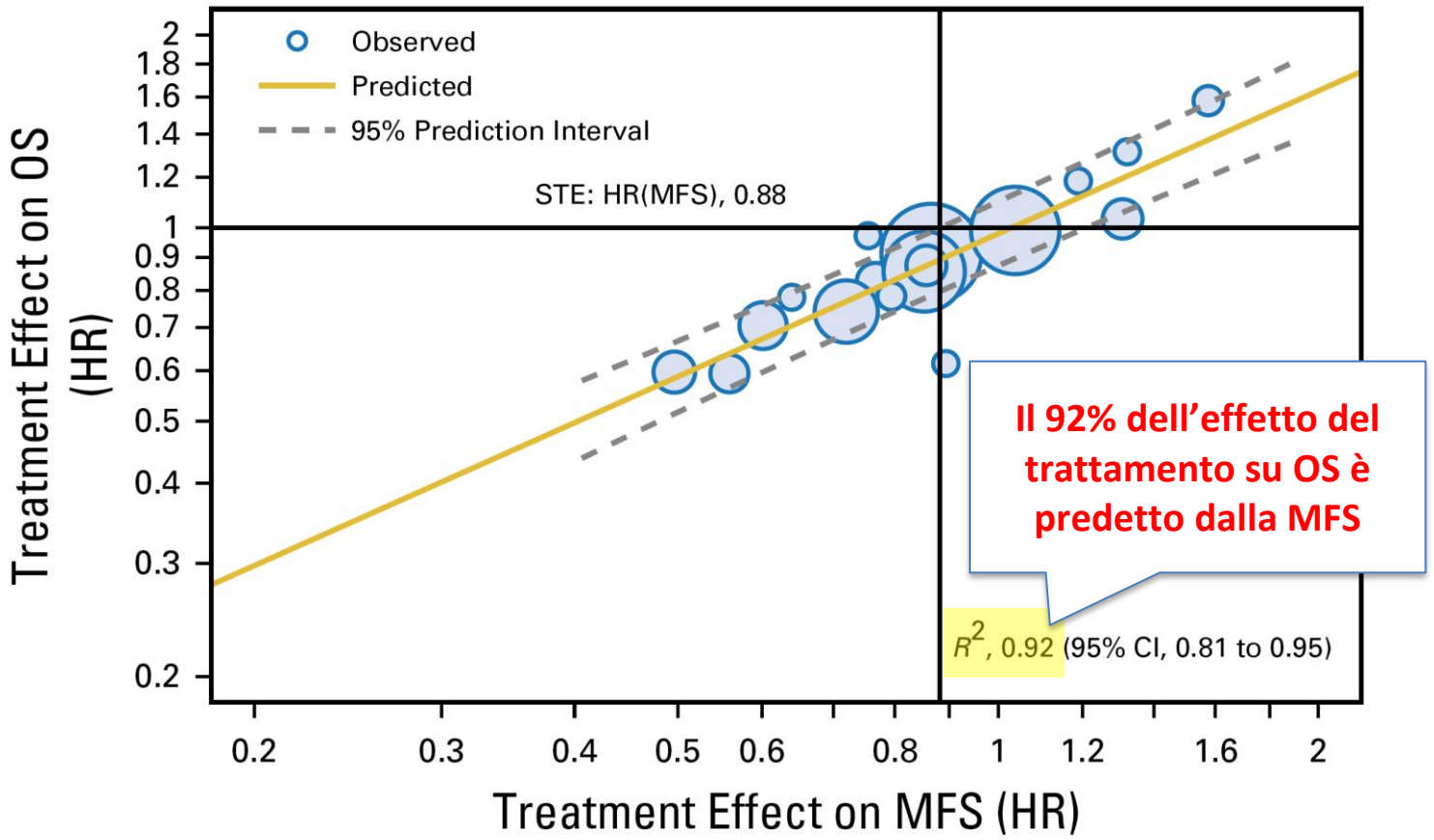


Burzykowski and Buyse, Pharmaceutical Statist 2006;5:173

Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer

Wanling Xie, Meredith M. Regan, Marc Buyse, Susan Halabi, Philip W. Kantoff, Oliver Sartor, Howard Soule, Noel W. Clarke, Laurence Collette, James J. Dignam, Karim Fizazi, Wendy R. Paruleker, Howard M. Sandler, Matthew R. Sydes, Bertrand Tombal, Scott G. Williams, and Christopher J. Sweeney, on behalf of the ICECaP Working Group

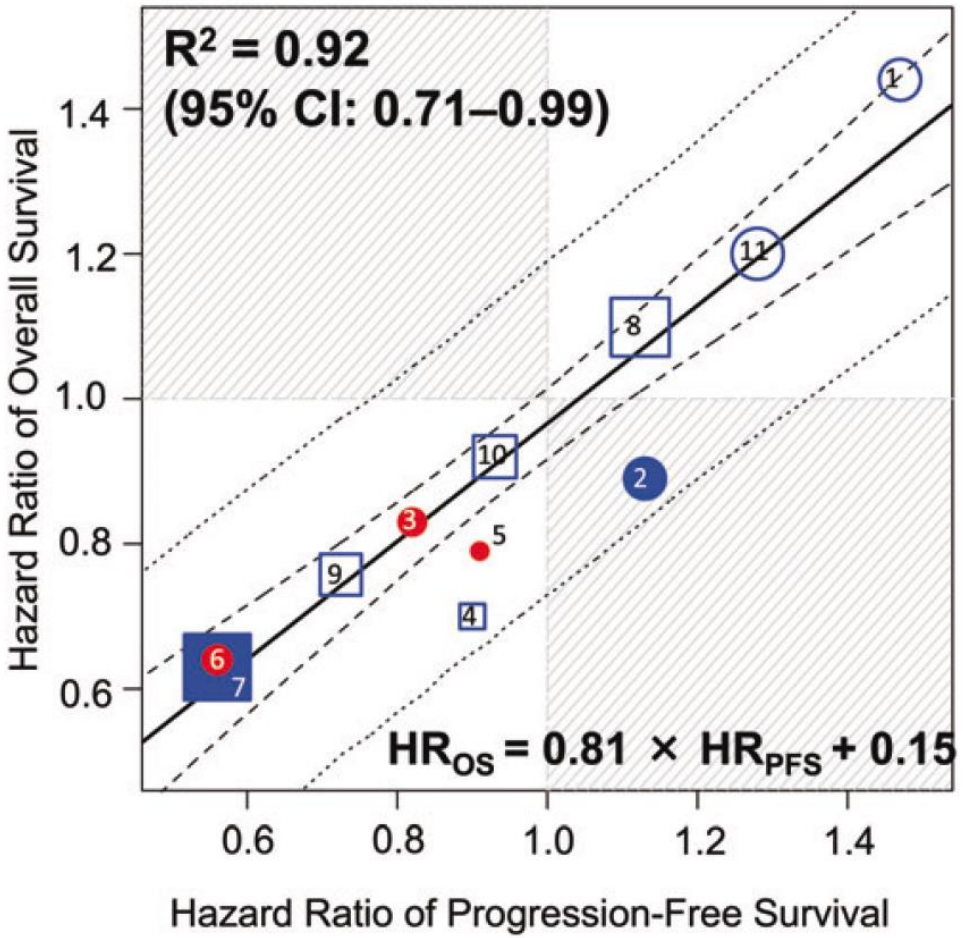
J Clin Oncol 35:3097-3104. © 2017 by American Society of Clinical Oncology



Progression-free survival as a surrogate endpoint for overall survival in glioblastoma: a literature-based meta-analysis from 91 trials

Kelong Han, Melanie Ren, Wolfgang Wick, Lauren Abrey, Asha Das, Jin Jin, and David A. Reardon

Neuro-Oncology 16(5), 696–706, 2014



Round symbols:	Glioblastoma
Square symbols:	Mixed high-grade glioma
Solid symbols:	Newly diagnosed
Hollow symbols:	Recurrent
———	Weighted linear fit
– – – – –	95% confidence interval
----------	95% prediction interval
Blue symbols:	Head-to-head comparison
Red symbols:	Single-arm trials using historical data as control

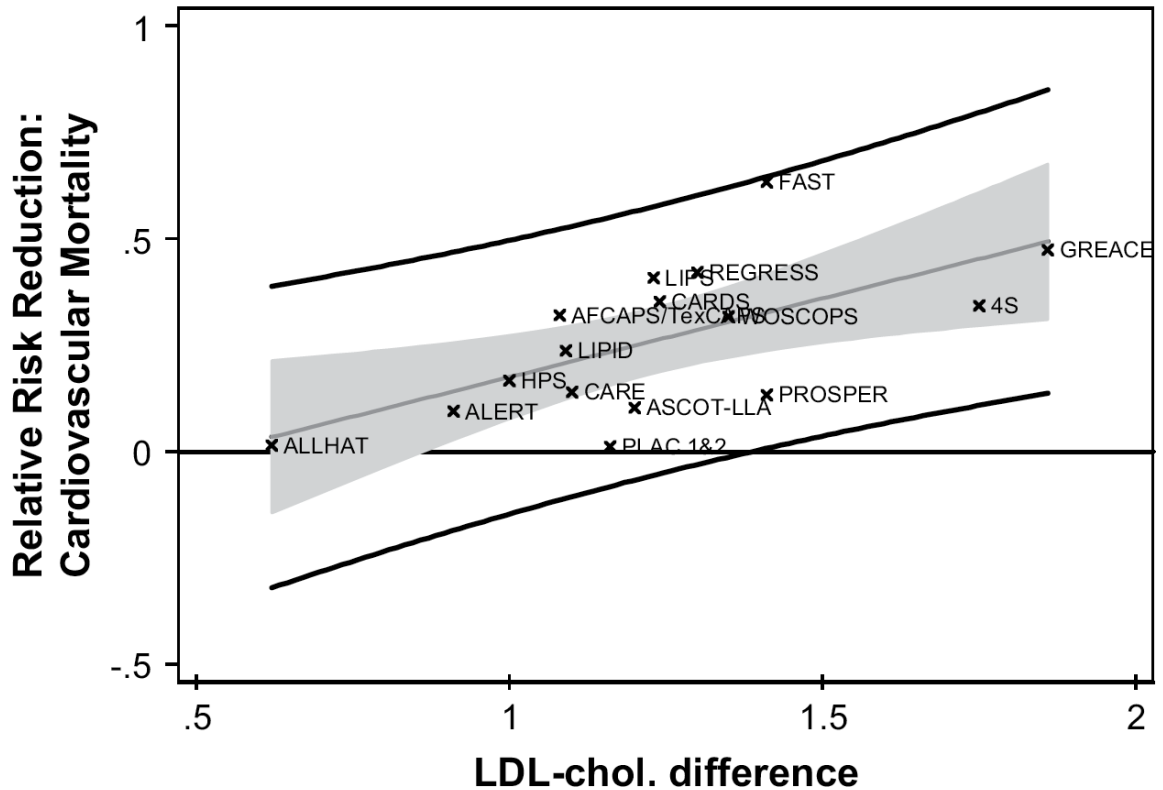
Linear regression determined that a 10% PFS risk reduction would yield an 8.1% ± 0.8% OS risk reduction.

LDL-cholesterol differences predicted survival benefit in statin trials by the surrogate threshold effect (STE)

Kent R. Johnson[a,*], Nick Freemantle[b], Danielle M. Anthony[a], Marissa N.D. Lassere[c]

[a]*Department of Clinical Pharmacology, University of Newcastle, Mater Hospital, Waratah NSW 2298, Australia*
[b]*Department of Primary Care and General Practice, University of Birmingham, Birmingham B15 2TT, UK*
[c]*Department of Rheumatology, University of New South Wales, St. George Hospital, Kogarah NSW 2217, Australia*

Journal of Clinical Epidemiology 62 (2009) 328–336

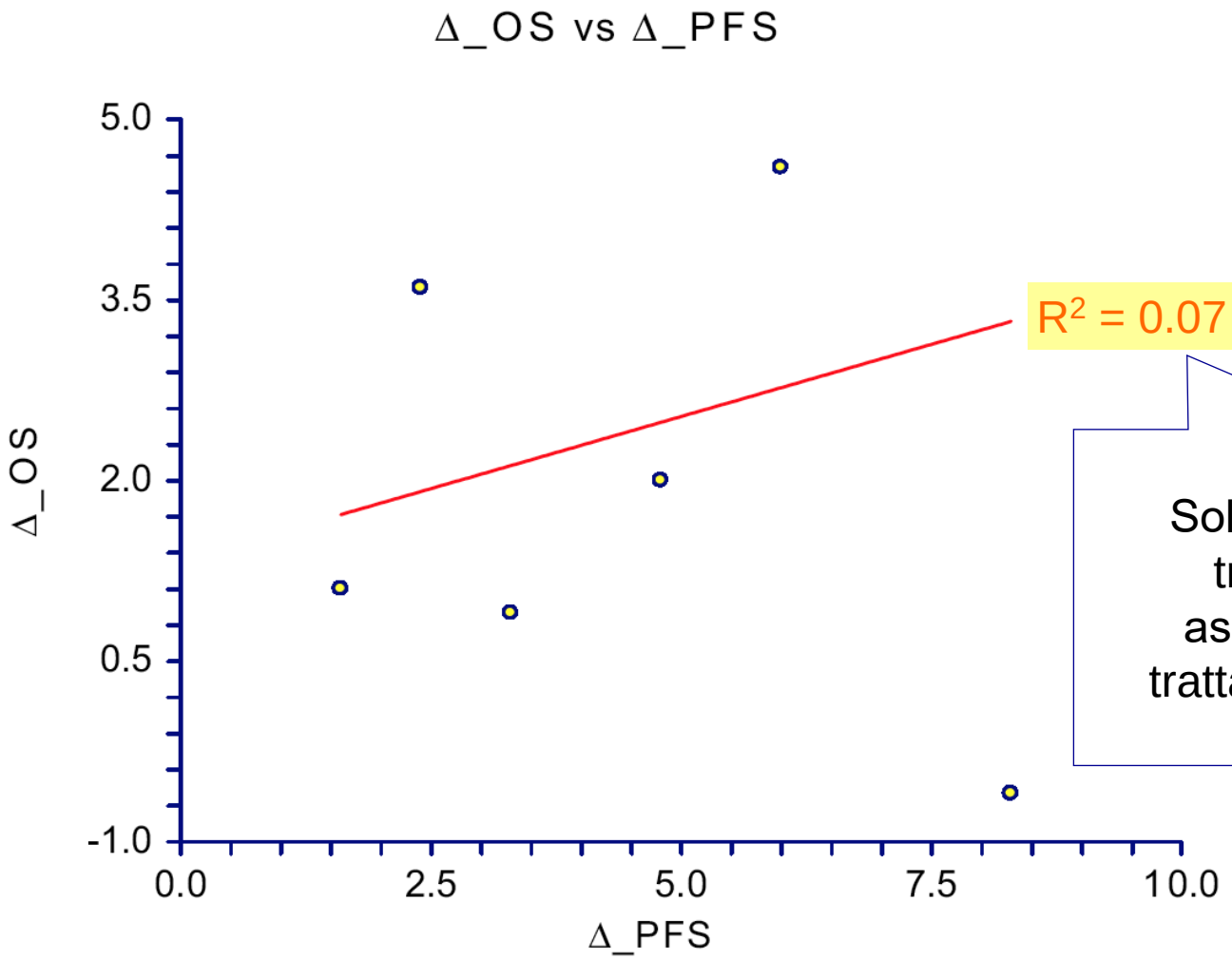


In 16 qualifying trials, regression analysis yielded a cardiovascular mortality model whose prediction bands demonstrated no cardiovascular survival benefit with LDL-cholesterol difference values below 1.4 mmol/L.

Surrogate End Points in Renal Cell Carcinoma: An Analysis of First-Line Trials With Targeted Therapies

Fausto Petrelli, Sandro Barni

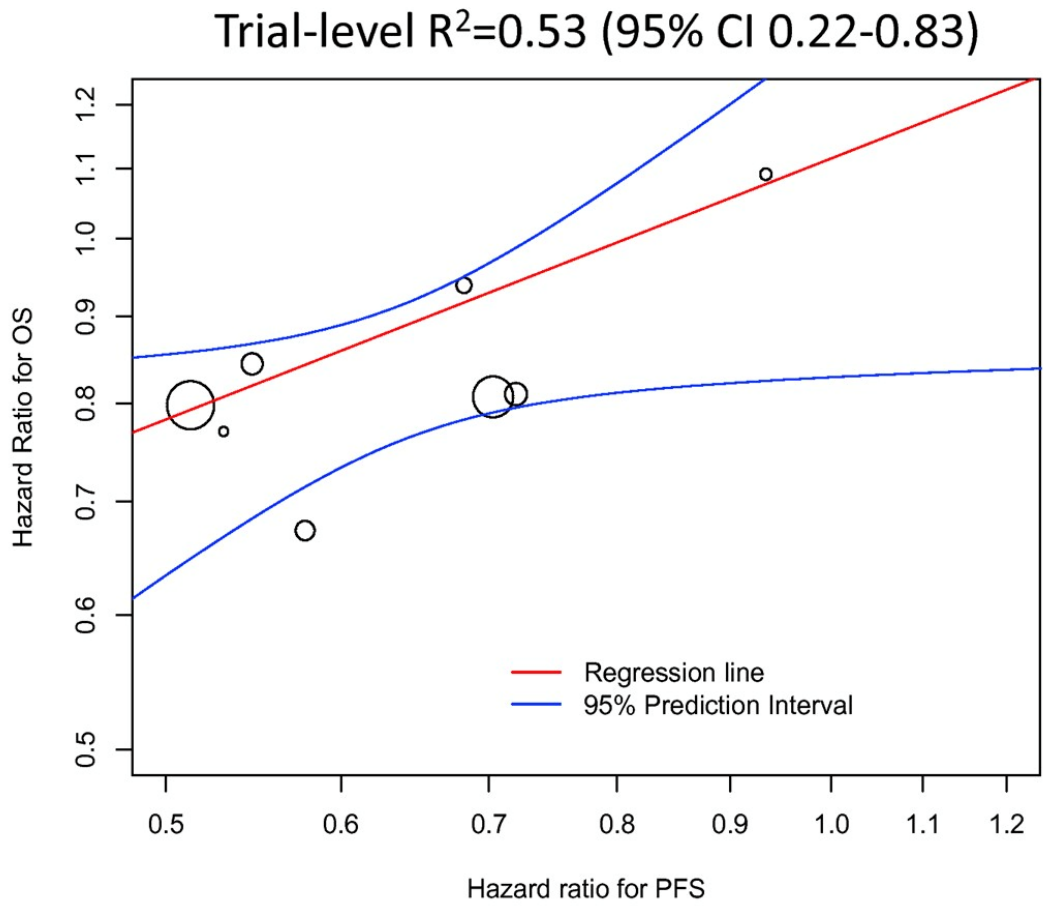
Clinical Genitourinary Cancer, Vol. 11, No. 4, 385-9 © 2013 Elsevier Inc.



Solo il 7% dell'effetto del trattamento su OS è ascrivibile all'effetto del trattamento stesso su PFS

S. Michiels[1,2], L. Pugliano[1,3], D. Grun[1], S. Marguet[2], J. Barinoff[4], D. Cameron[5], M. Cobleigh[6], A. Di Leo[7], S. Johnston[8], G. Gasparini[9], B.Kaufman[10], M. Marty[11], V. Nekjudova[12], S. Paluch-Shimon[13], F.Penault-Llorca[14], D. Slamon[15], C. Vogel[16], G. von Minckwitz[12], M. Buyse[17], M. Piccart[1,3]

ASCO 2013 CHICAGO 30 MAY - 04 JUN



- At the trial level, only 53% of the variation in treatment effects on OS can be explained by effects on PFS (trial-level R^2=0.53).

Gli Endpoint surrogati

- Rivestono un ruolo molto importante
  - nella pratica
  - nella ricerca clinica
- la sola dimostrazione di associazione tra un biomarker e l'endpoint clinico
  - non fornisce una evidenza sufficiente di surrogacy
- data l'alta variabilità e la scarsa conoscenza sui meccanismi biologici che collegano il trattamento con l'outcome clinico,
  - il ruolo dei SE deve essere dimostrato e validato per ogni specifica malattia e per ogni singola classe di farmaco

Take home messages

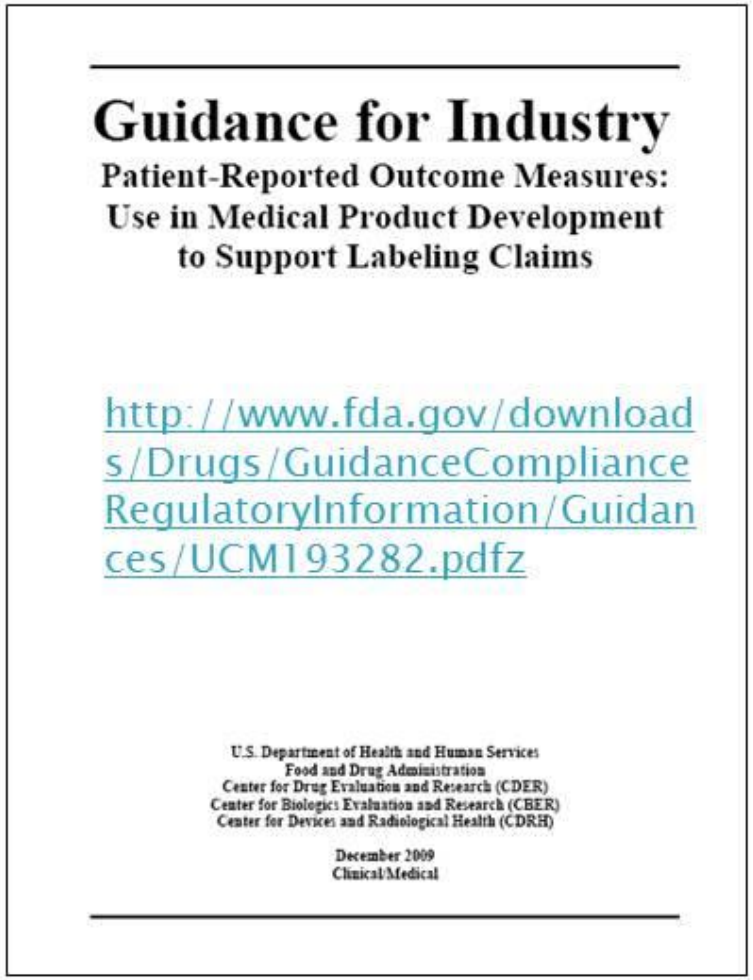
V. Torri, AIOM 2014

Regular Approval Basis – "Clinical Benefit"

- Longer life
- **Better life**
- Established Surrogate for one of above

Definition of PROs

- "Any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else"



Guidance for Industry

Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdfz

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)
Center for Devices and Radiological Health (CDRH)

December 2009
Clinical/Medical

Symptom Endpoints (Patient-Reported Outcomes)

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

Standards for PRO Development

- **Reliability**
  - **Test-retest**
  - **Internal consistency**
- **Validity**
  - **Content validity (qualitative)**
  - **Construct validity (discriminant)**
- **Ability to detect change**

Ethan Basch, 2010

Standards for PRO Development

- **Reliability**
  - **Test-retest**
- **Validity**
  - Content validity (qualitative)
  - Construct validity (discriminant)
- **Ability to detect change**

Reliability means the consistency or repeatability of the measure.

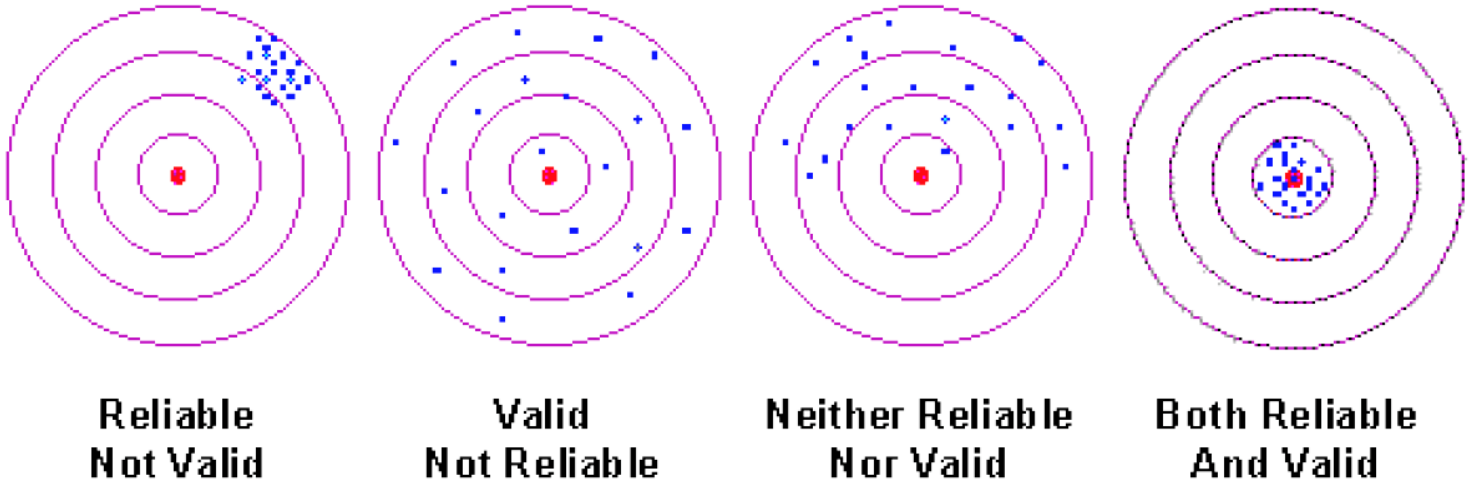
Ethan Basch, 2010

Standards for PRO Development

- Reliability
  - Test-retest
  - Internal consistency
- **Validity**
  - **C tent validity (qualitative)**
  - **iscriminant)**
- nge

Validity means measuring what you claim to be measuring.

Ethan Basch, 2010



- In the first situation, you are consistently and systematically measuring the wrong value for all respondents. This measure is reliable, but no valid.
- In the second situation, you get a valid group estimate, but you are inconsistent. Here, you can clearly see that reliability is directly related to the variability of your measure.
- The third scenario shows a case where your hits are spread across the target and you are consistently missing the center. Your measure in this case is neither reliable nor valid.
- Finally, we see the "Robin Hood" scenario – you consistently hit the center of the target. Your measure is both reliable and valid.

Standards for PRO Development

- Reliability
  - Test-retest
  - Internal consistency
- Validit
  - Cont
  - Constru idity (discriminant)
- **Ability to detect change**

The PRO instrument can identify differences in scores over time

Ethan Basch, 2010

NCI- PRO-CTCAE™ ITEMS-ITALIAN

Item Library Version 1.0

Quando un individuo è in terapia per un tumore, talvolta può sviluppare diversi sintomi ed effetti collaterali. Per ciascuna domanda, fare un segno o una X nella casella che meglio corrisponde all'esperienza vissuta negli ultimi sette giorni…

1. PRO-CTCAE™ Symptom Term: Dry mouth				
SENSAZIONE DI BOCCA SECCA				
Negli ultimi 7 giorni, quanto è stata GRAVE la SENSAZIONE DI BOCCA SECCA nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

2. PRO-CTCAE™ Symptom Term: Difficulty swallowing				
DIFFICOLTÀ A DEGLUTIRE				
Negli ultimi 7 giorni, quanto è stata GRAVE la DIFFICOLTÀ A DEGLUTIRE nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

3. PRO-CTCAE™ Symptom Term: Mouth/throat sores				
PIAGHE IN BOCCA O IN GOLA				
Negli ultimi 7 giorni, quanto sono state GRAVI le PIAGHE IN BOCCA O IN GOLA nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo
Negli ultimi 7 giorni, in che misura le PIAGHE IN BOCCA O IN GOLA HANNO INTERFERITO con le Sue attività abituali o quotidiane?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

4. PRO-CTCAE™ Symptom Term: Cracking at the corners of the mouth (cheilosis/cheilitis)				
SCREPOLATURE AGLI ANGOLI DELLA BOCCA				
Negli ultimi 7 giorni, quanto sono state GRAVI le SCREPOLATURE AGLI ANGOLI DELLA BOCCA, nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

Version date: 6/3/2017

https://healthcaredelivery.cancer.gov/pro-ctcae/pro-ctcae_italian.pdf

Validity and Reliability of the US National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)

Amylou C. Dueck, PhD; Tito R. Mendoza, PhD; Sandra A. Mitchell, PhD, CRNP, AOCN; Bryce B. Reeve, PhD; Kathleen M. Castro, RN, MS, AOCN; Lauren J. Rogak, MA; Thomas M. Atkinson, PhD; Antonia V. Bennett, PhD; Andrea M. Denicoff, MS, RN, ANP; Ann M. O'Mara, PhD, RN, FAAN; Yuelin Li, PhD; Steven B. Clauser, PhD, MPA; Donna M. Bryant, MSN, ANP-BC, OCN, CCRC; James D. Bearden III, MD, FACP; Theresa A. Gillis, MD; Jay K. Harness, MD; Robert D. Siegel, MD, FACP; Diane B. Paul, AAS; Charles S. Cleeland, PhD; Deborah Schrag, MD, MPH; Jeff A. Sloan, PhD; Amy P. Abernethy, MD, PhD; Deborah W. Bruner, RN, PhD, FAAN; Lori M. Minasian, MD, FACP; Ethan Basch, MD, MSc; for the National Cancer Institute PRO-CTCAE Study Group

JAMA Oncol. 2015;1(8):1051-1059.

NATIONAL CANCER INSTITUTE

Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) Item Library (version 1)

Certified Translation

This is to certify that the Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) Item Library (version 1) was translated from English to Italian using a universal translation methodology including representation from major Italian speaking regions by qualified translators employed by Italian National Cancer Institute, under the direction of the US National Cancer Institute.

This translation followed the translation methodology recommended by International Society of Pharmacoeconomic and Outcomes Research (ISPOR) a methodology established to ensure that resulting translations of patient-reported outcome measures reflect conceptual equivalence with the source document rendered in language that is culturally acceptable and relevant to the target population. This rigorous methodology requires two forward translations into the target language by native speakers, a reconciled version of the two forward translations done by a third independent translator who is a native speaker of the target language, a back translation of the reconciled version by a native English speaker fluent in the target language, and an independent review by a native speaker trained as a physician or nurse in oncology.

After the translation phase was completed, the PRO-CTCAE Item Library (version 1) was linguistically validated by testing with patients to confirm suitability of the translations for Italian speaking patients. All translation work was performed by members of the Italian translation team to the best of their abilities as native speakers of Italian (or English in the case of the back-translator), and as translators and researchers experienced in the field of health-related quality of life and patient-reported outcomes survey research under the direction of the undersigned.

This translation is, to the best of my knowledge, a valid and accurate translation of the corresponding original English language version of the PRO-CTCAE Item Library (version 1).

Name: Sandra A. Mitchell, PhD, CRNP

Title: Research Scientist and Program Director; Outcomes Research Branch; US National Cancer Institute

Signature:

Date: June 5, 2017

Lise Rene Tamsett NOTARY PUBLIC Montgomery County State of Maryland My Commission Expires

6/07/2017

NCI- PRO-CTCAE™ ITEMS-ITALIAN

Item Library Version 1.0

Quando un individuo è in terapia per un tumore, talvolta può sviluppare diversi sintomi ed effetti collaterali. Per ciascuna domanda, fare un segno o una (X) nella casella che meglio corrisponde all'esperienza vissuta negli ultimi sette giorni…

1. PRO-CTCAE™ Symptom Term: Dry mouth				
SENSAZIONE DI BOCCA SECCA				
Negli ultimi 7 giorni, quanto è stata GRAVE la SENSAZIONE DI BOCCA SECCA nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

2. PRO-CTCAE™ Symptom Term: Difficulty swallowing				
DIFFICOLTÀ A DEGLUTIRE				
Negli ultimi 7 giorni, quanto è stata GRAVE la DIFFICOLTÀ A DEGLUTIRE nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

3. PRO-CTCAE™ Symptom Term: Mouth/throat sores				
PIAGHE IN BOCCA O IN GOLA				
Negli ultimi 7 giorni, quanto sono state GRAVI le PIAGHE IN BOCCA O IN GOLA nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo
Negli ultimi 7 giorni, in che misura le PIAGHE IN BOCCA O IN GOLA HANNO INTERFERITO con le Sue attività abituali o quotidiane?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

4. PRO-CTCAE™ Symptom Term: Cracking at the corners of the mouth (cheilosis/cheilitis)				
SCREPOLATURE AGLI ANGOLI DELLA BOCCA				
Negli ultimi 7 giorni, quanto sono state GRAVI le SCREPOLATURE AGLI ANGOLI DELLA BOCCA, nel momento PEGGIORE?				
O Per nulla	O Un po'	O Abbastanza	O Molto	O Moltissimo

The PRO-CTCAE™ items and information herein were developed by the Division of Cancer Control and Population Sciences in the NATIONAL CANCER INSTITUTE at the NATIONAL INSTITUTES OF HEALTH, in Bethesda, Maryland, U.S.A. Use of the PRO- CTCAE™ is subject to NCI's Terms of Use.

Version date: 6/3/2017

https://healthcaredelivery.cancer.gov/pro-ctcae/pro-ctcae_italian.pdf

Co-primary endpoints can be different medical assessments angled at different aspects of a disease, therefore, are used collectively to strengthen evidence for the treatment effect.

Li QH. Evaluating co-primary endpoints collectively in clinical trials. Biom J. 2009 Feb;51(1):137-45.

Guidance for Industry

E9 Statistical Principles for Clinical Trials

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)
September 1998
ICH

5. *Multiple Primary Variables (2.2.5)*

It may sometimes be desirable to use more than one primary variable, each of which (or a subset of which) could be sufficient to cover the range of effects of the therapies.

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Abiraterone in Metastatic Prostate Cancer without Previous Chemotherapy

Charles J. Ryan, M.D., Matthew R. Smith, M.D., Ph.D., Johann S. de Bono, M.B., Ch.B., Ph.D., Arturo Molina, M.D., Christopher J. Logothetis, M.D., Paul de Souza, M.B., Ph.D., Karim Fizazi, M.D., Ph.D., Paul Mainwaring, M.D., Josep M. Piulats, M.D., Ph.D., Siobhan Ng, M.D., Joan Carles, M.D., Peter F.A. Mulders, M.D., Ph.D., Ethan Basch, M.D., Eric J. Small, M.D., Fred Saad, M.D., Dirk Schrijvers, M.D., Ph.D., Hendrik Van Poppel, M.D., Ph.D., Som D. Mukherjee, M.D., Henrik Suttmann, M.D., Winald R. Gerritsen, M.D., Ph.D., Thomas W. Flaig, M.D., Daniel J. George, M.D., Evan Y. Yu, M.D., Eleni Efstathiou, M.D., Ph.D., Allan Pantuck, M.D., Eric Winquist, M.D., Celestia S. Higano, M.D., Mary-Ellen Taplin, M.D., Youn Park, Ph.D., Thian Kheoh, Ph.D., Thomas Griffin, M.D., Howard I. Scher, M.D., and Dana E. Rathkopf, M.D., for the COU-AA-302 Investigators*

METHODS

In this double-blind study, we randomly assigned 1088 patients to receive abiraterone acetate (1000 mg) plus prednisone (5 mg twice daily) or placebo plus prednisone. The coprimary end points were radiographic progression-free survival and overall survival.

Encorafenib, Binimetinib, and Cetuximab in *BRAF* V600E–Mutated Colorectal Cancer

S. Kopetz, A. Grothey, R. Yaeger, E. Van Cutsem, J. Desai, T. Yoshino, H. Wasan, F. Ciardiello, F. Loupakis, Y.S. Hong, N. Steeghs, T.K. Guren, H.-T. Arkenau, P. Garcia-Alfonso, P. Pfeiffer, S. Orlov, S. Lonardi, E. Elez, T.-W. Kim, J.H.M. Schellens, C. Guo, A. Krishnan, J. Dekervel, V. Morris, A. Calvo Ferrandiz, L.S. Tarpgaard, M. Braun, A. Gollerkeri, C. Keir, K. Maharry, M. Pickard, J. Christy-Bittel, L. Anderson, V. Sandor, and J. Tabernero

N Engl J Med 2019;381:1632-43.

END POINTS

The original sole primary end point was overall survival in the triplet-therapy group as compared with the control group. The protocol was amended to include an additional primary end point of the objective response rate in the triplet-therapy group as compared with the control group, as assessed by independent central reviewers who were unaware of the treatment assignments, and the initial interim analysis of overall survival (the results of which are reported here) was added in an attempt to expeditiously assess efficacy.

Applied Clinical Trials Online

Composite Endpoints: Proceed with Caution

By Peter Kleist May 1, 2006

Multiple single endpoints are combined in order to confront an investigational drug with a higher number of events expected during the trial.

Statistical precision and efficiency will be increased, trials become smaller, less costly, and the results of promising new treatments will be available earlier.

The selected individual components of a composite endpoint, as reported in the biomedical literature, are not always clinically meaningful.

Trial of invasive versus medical therapy in elderly patients with chronic symptomatic coronary-artery disease (TIME): a randomised trial

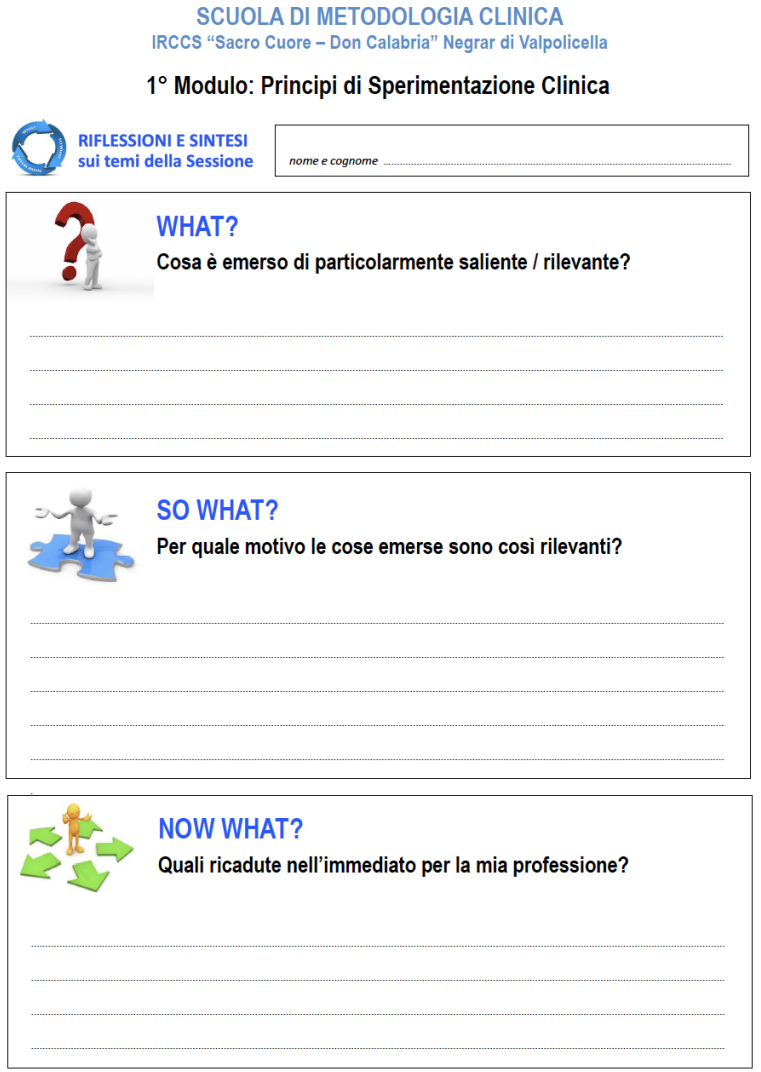
The TIME Investigators

Lancet 2001; **358:** 951–957

The frequency of the composite endpoint (death, nonfatal myocardial infarction, and hospital admission for ACS) was much lower with revascularization; however, this was due to a marked difference in hospital admissions, which accounted for 75% of the events in the medical treatment group. In contrast, there were twice as many deaths in the invasive treatment group.

The question remains how to interpret the results and inform a patient who has to decide between conservative or surgical therapy.

1. Riflettete da soli per 10 min. e compilate il form →
2. Confrontatevi con i Colleghi del Vostro tavolo per 15 min., declinate un W^3 condiviso e delegate un portavoce
3. Riportate sulla lavagna il Vostro W^3 condiviso su almeno due aspetti ritenuti rilevanti e impattanti sulla professione (in 5 min.)
4. Presentate ai Colleghi degli altri tavoli il Vostro W^3 condiviso



SCUOLA DI METODOLOGIA CLINICA

IRCCS "Sacro Cuore – Don Calabria" Negrar di Valpolicella

1° Modulo: Principi di Sperimentazione Clinica

RIFLESSIONI E SINTESI sui temi della Sessione

nome e cognome .....

WHAT?

Cosa è emerso di particolarmente saliente / rilevante?

SO WHAT?

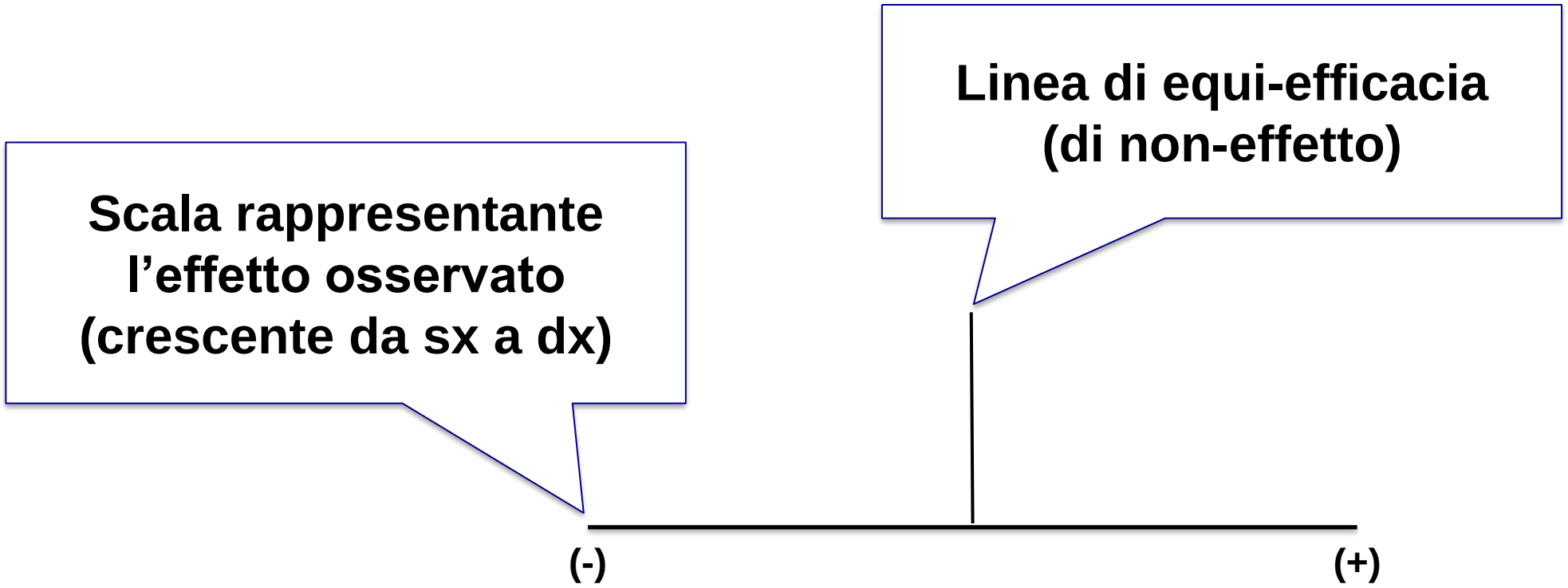
Per quale motivo le cose emerse sono così rilevanti?

NOW WHAT?

Quali ricadute nell'immediato per la mia professione?

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- **Il Forest Plot**
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

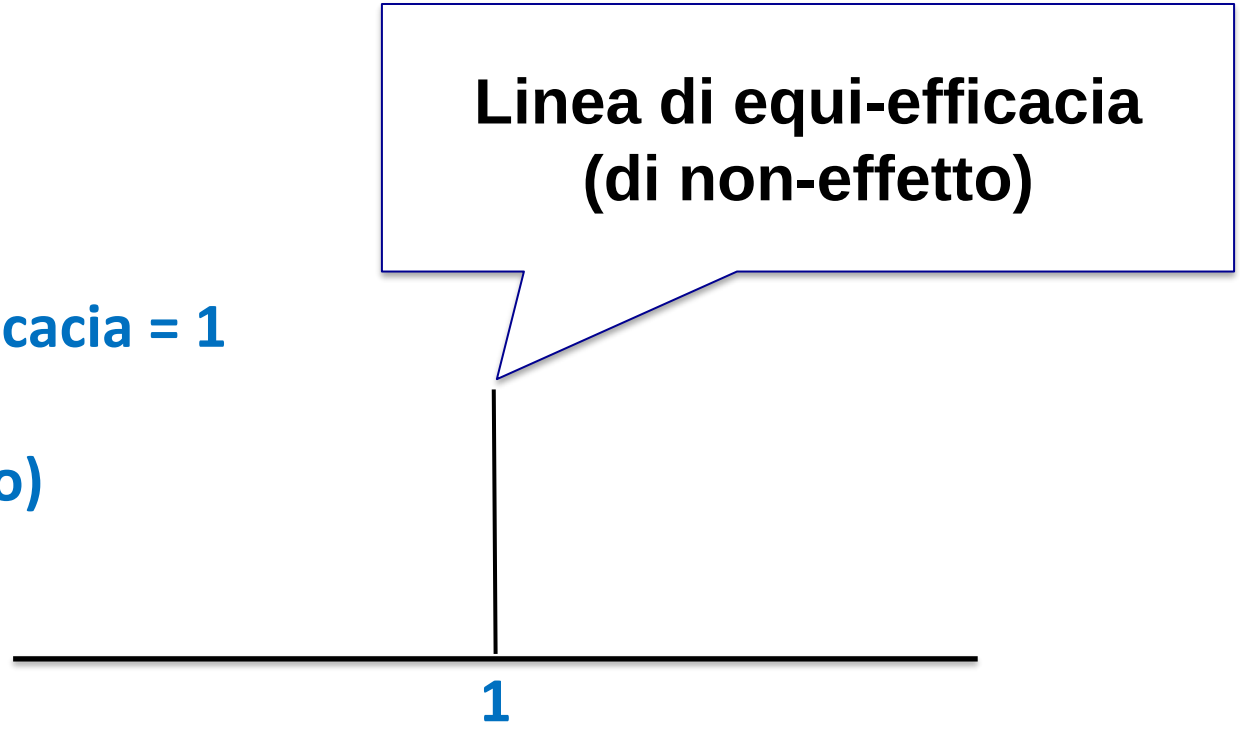
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione degli studi clinici mediante Forest (Forrest?) Plot

Misure di effetto con equi-efficacia = 1

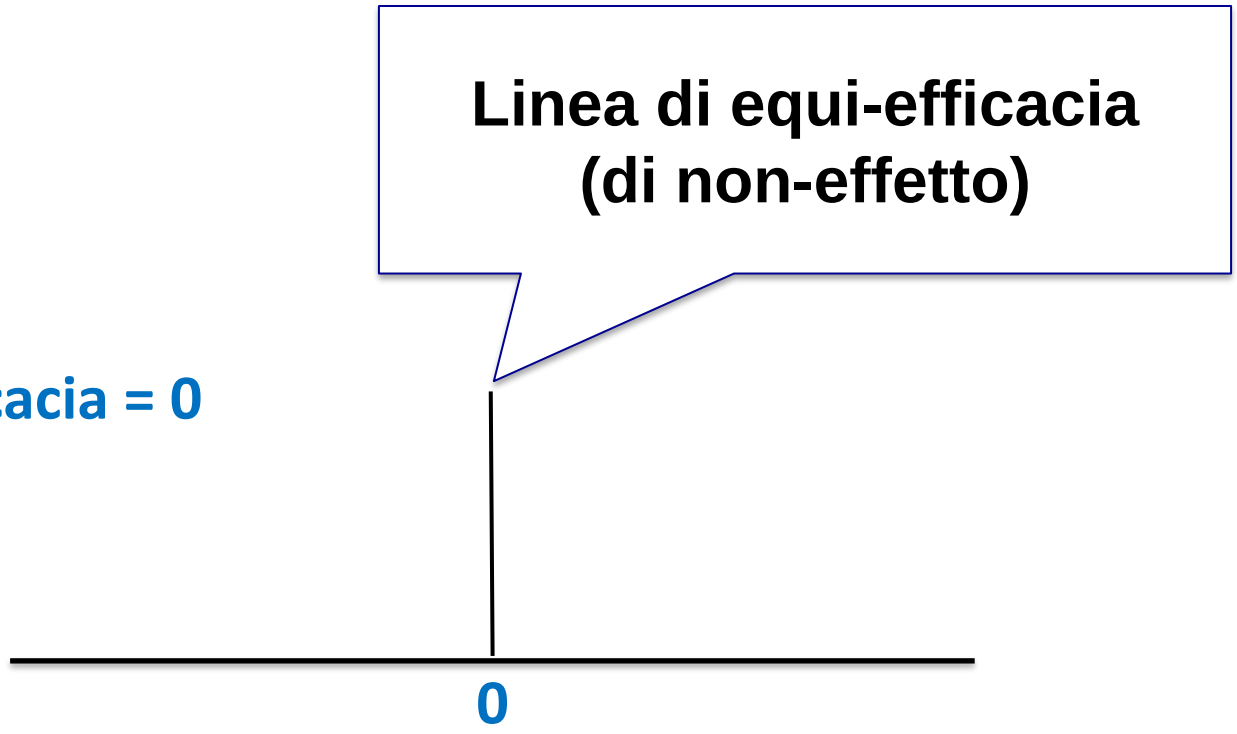
- RR, OR (variabili nominali)
- HR (variabili tempo a evento)



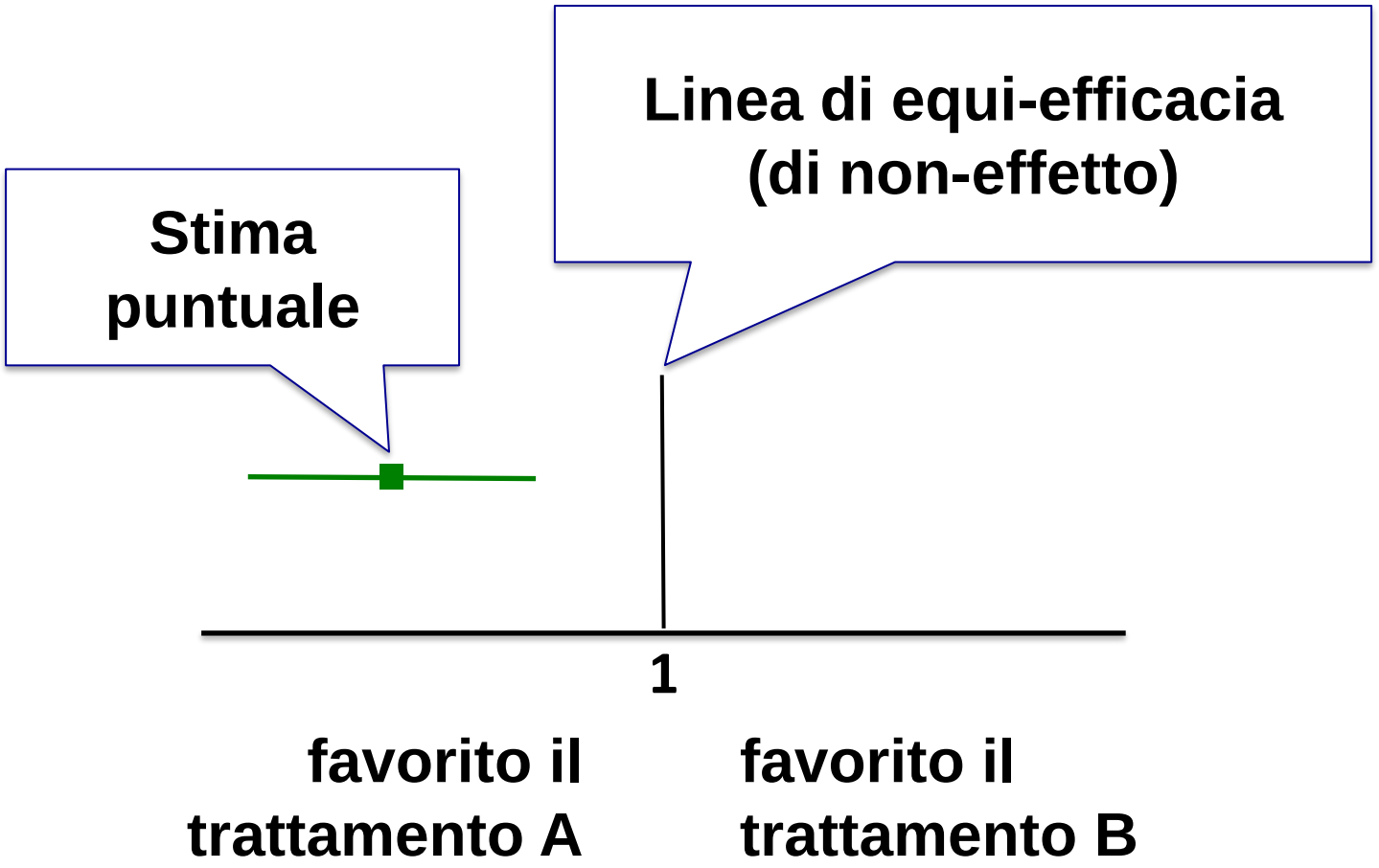
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot

Misure di effetto con equi-efficacia = 0

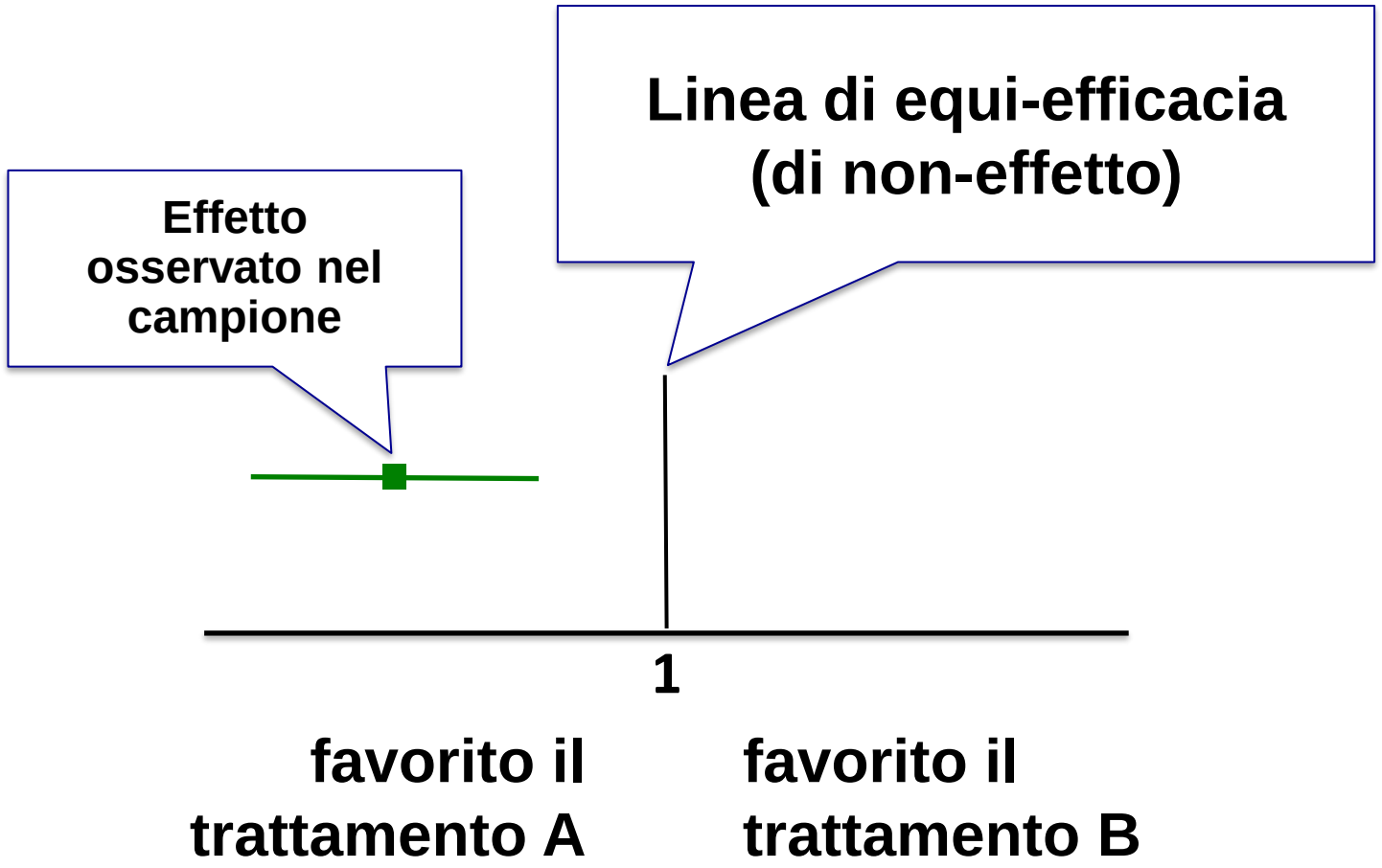
- **MD (variabili intervallari)**



Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



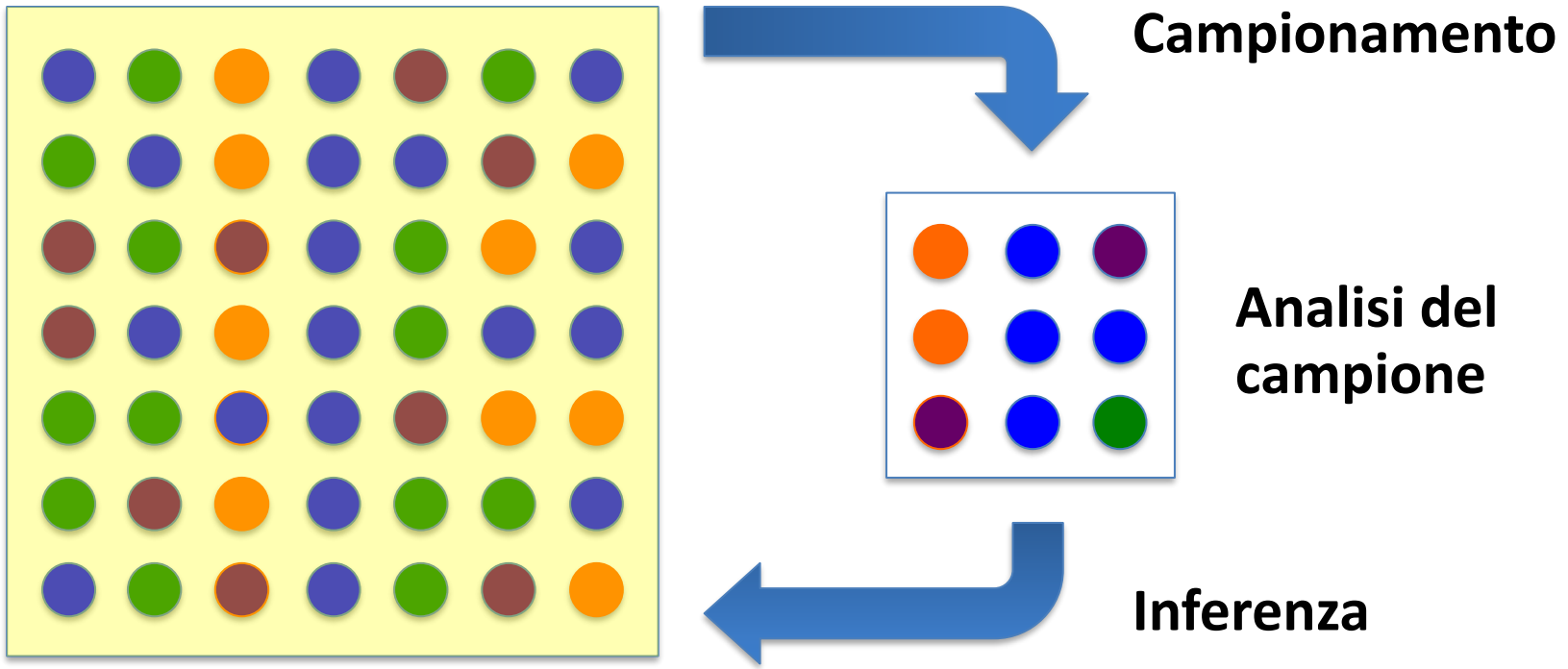
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- **Interpretazione statistica dei risultati dello studio**
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

Popolazione Vs Campione

(ovvero: perché è consigliato non fidarsi…)



- L’effetto osservato è solo uno dei possibili risultati dello studio legati all’assegnazione casuale dei soggetti ai trattamenti.
- Se non c’è un reale effetto del trattamento, è comunque possibile che per caso si osservino differenze tra i gruppi.
- Se c’è un reale effetto del trattamento, è comunque possibile che per caso non si osservi quella differenza fra i trattamenti.

effetto osservato = effetto del trattamento + effetto del caso

(in assenza di errore sistematico!)

Hypothesis Testing and Jury Trials

In tribunale	Test statistico
Presunzione di innocenza	Ipotesi di non differenza tra i trattamenti (H_0)
Evidenza di colpevolezza	Ipotesi di efficacia del trattamento in esame (H_1)

Susan Hilsenbeck, SABCS 2007 (modif. GLP)

# of Hypothesis

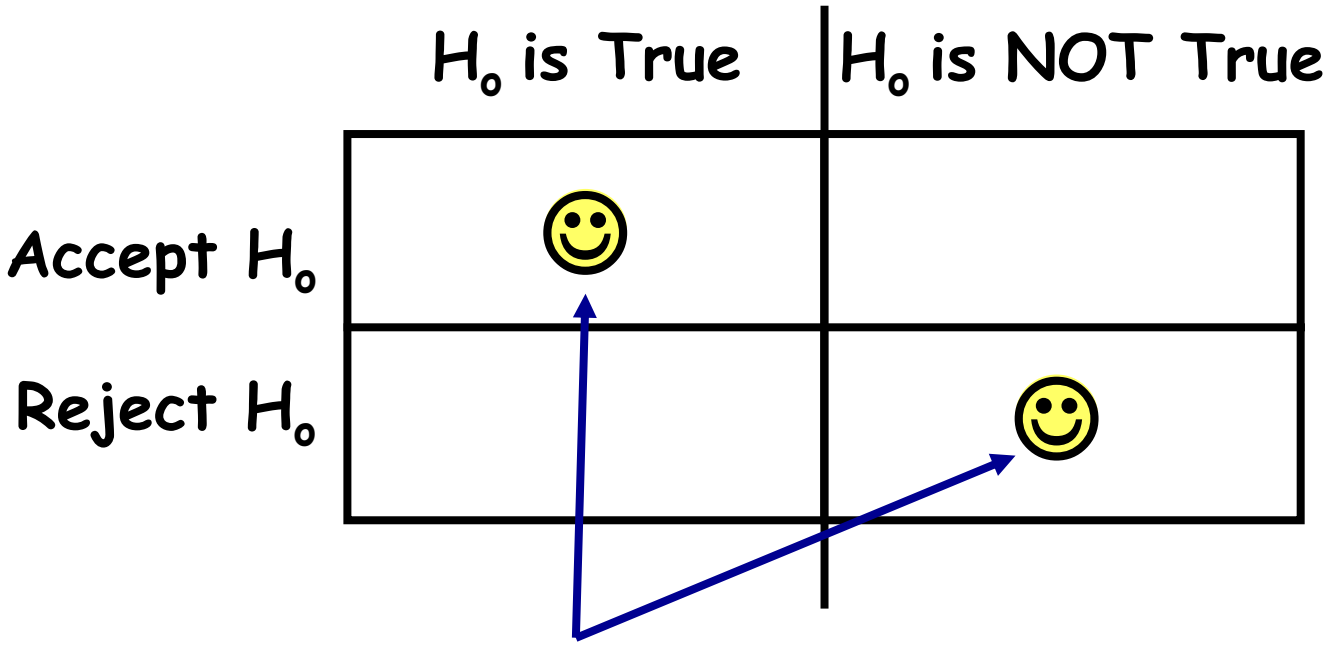
Ma noi non sappiamo quale delle due ipotesi sia quella vera nella realtà (cioè nella popolazione) …

	H_o is True	H_o is NOT True
Accept H_o		
Reject H_o		

Il test statistico può risultare positivo o negativo

Susan Hilsenbeck, SABCS 2007 (modif. GLP)

Making Decisions: Test of Hypothesis



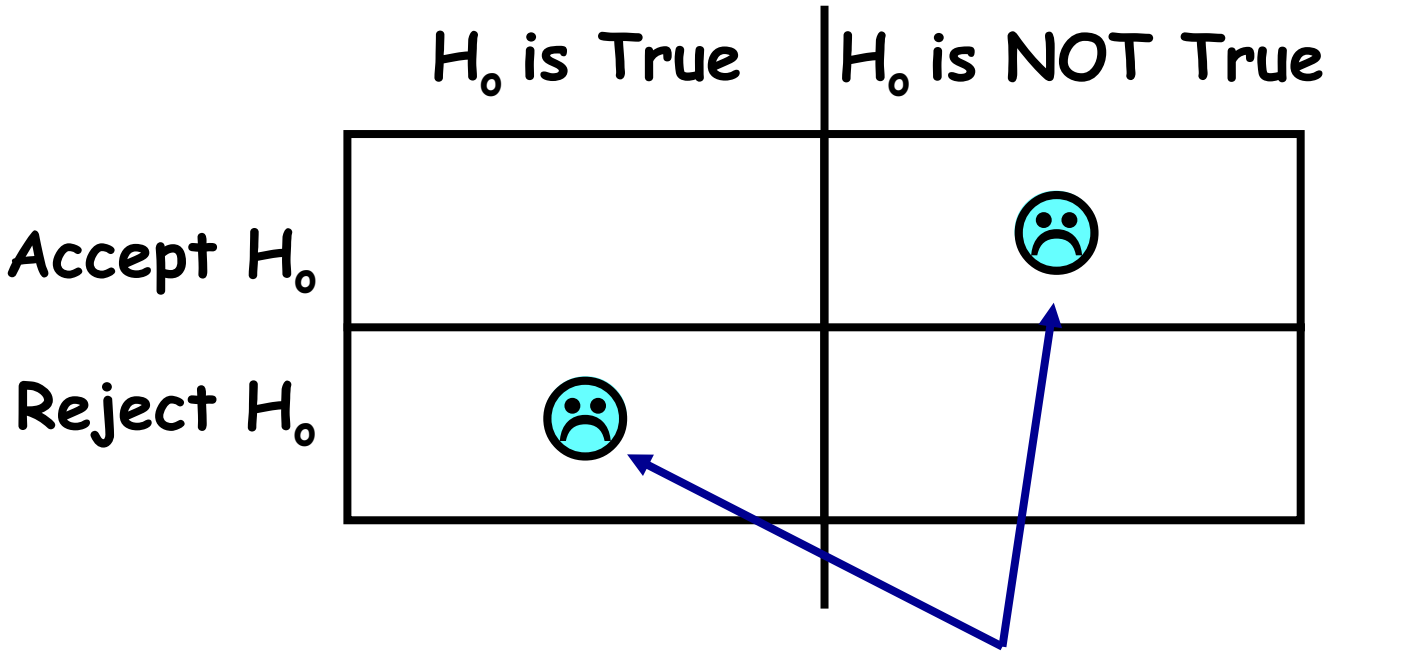
Il test potrebbe quindi dare
la risposta corretta…

Hypothesis Testing and Jury Trials

In tribunale	Test statistico
Presunzione di innocenza	Ipotesi di non differenza tra i trattamenti (H_0)
Evidenza di colpevolezza	Ipotesi di efficacia del trattamento in esame (H_1)
Assolvere l'innocente	Accettazione corretta di H_0
Condannare il colpevole	Corretto rifiuto di H_0
Condannare l'innocente	
Assolvere il colpevole	

Susan Hilsenbeck, SABCS 2007 (modif. GLP)

Making Decisions: Test of Hypothesis



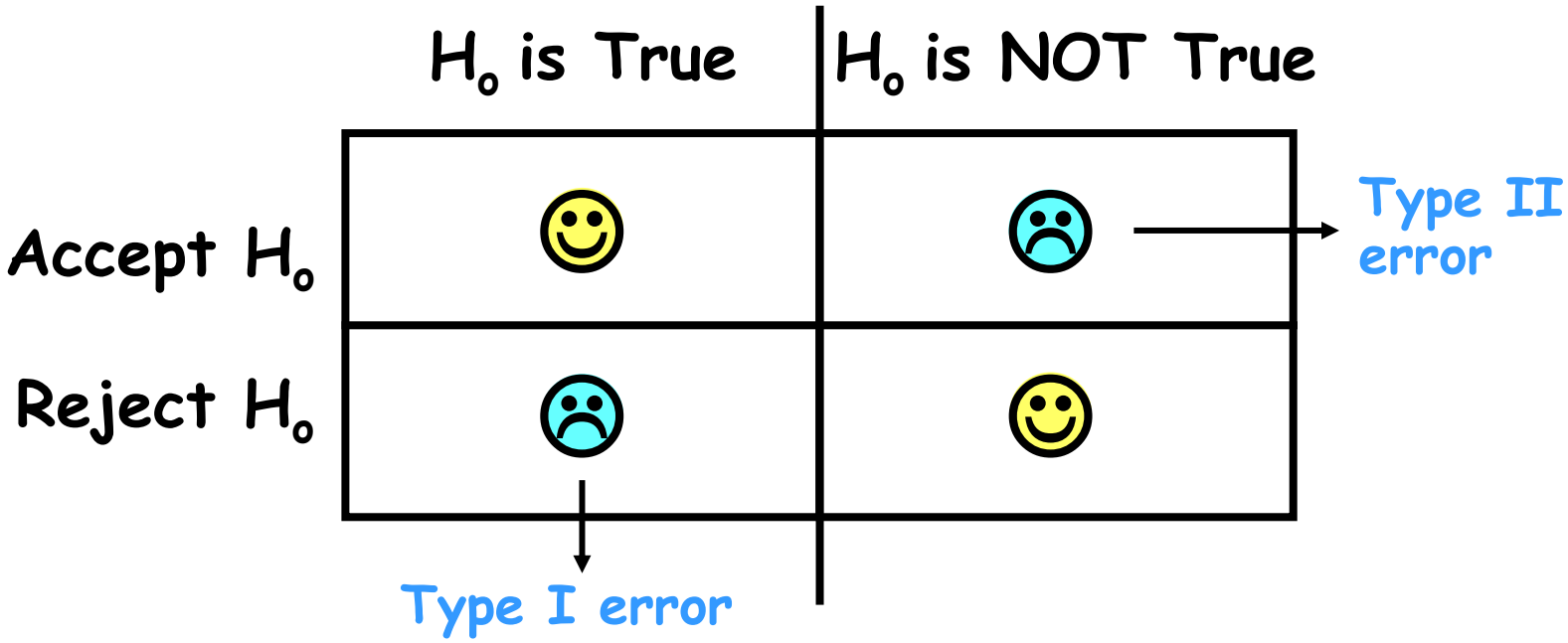
… oppure potrebbero essere stati commessi degli errori.

Hypothesis Testing and Jury Trials

In tribunale	Test statistico
Presunzione di innocenza	Ipotesi di non differenza tra i trattamenti (H_0)
Evidenza di colpevolezza	Ipotesi di efficacia del trattamento in esame (H_1)
Assolvere l'innocente	Accettazione corretta di H_0
Condannare il colpevole	Corretto rifiuto di H_0
Condannare l'innocente	Errore di 1° tipo
Assolvere il colpevole	Errore di 2° tipo

Susan Hilsenbeck, SABCS 2007 (modif. GLP)

Making Decisions: Test of Hypothesis



La probabilità di commettere un errore statistico è strettamente dipendente dalle dimensioni del campione

Susan Hilsenbeck, SABCS 2007 (modif. GLP)

ERRORI STATISTICI

- ✓ **Errore di 1° tipo** (errore *alfa*)
  - quando si conclude per un'efficacia del trattamento sperimentale, quando non lo è nella realtà; *lo studio è falsamente positivo*.
- ✓ **Errore di 2° tipo** (errore *beta*)
  - quando si conclude per una non efficacia del trattamento sperimentale, quando invece lo è nella realtà; *lo studio è falsamente negativo*.
- ✓ Il calcolo delle dimensioni del campione mira a contenere la dimensione degli errori statistici entro valori accettabili (5% per l'errore di 1° tipo e 20% per l'errore di 2° tipo)

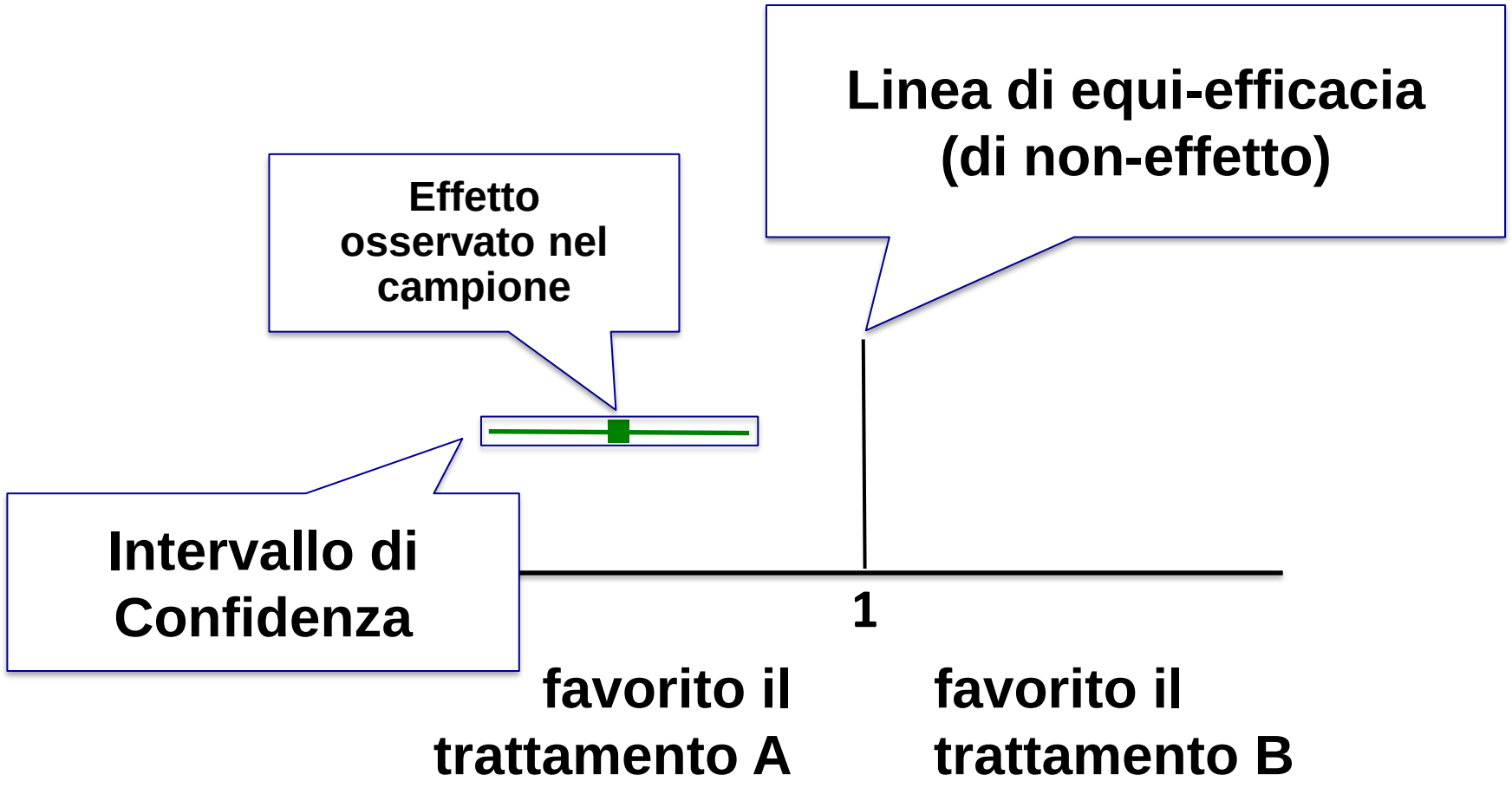
False Positive (type 1 error)

False Negative (type 2 error)

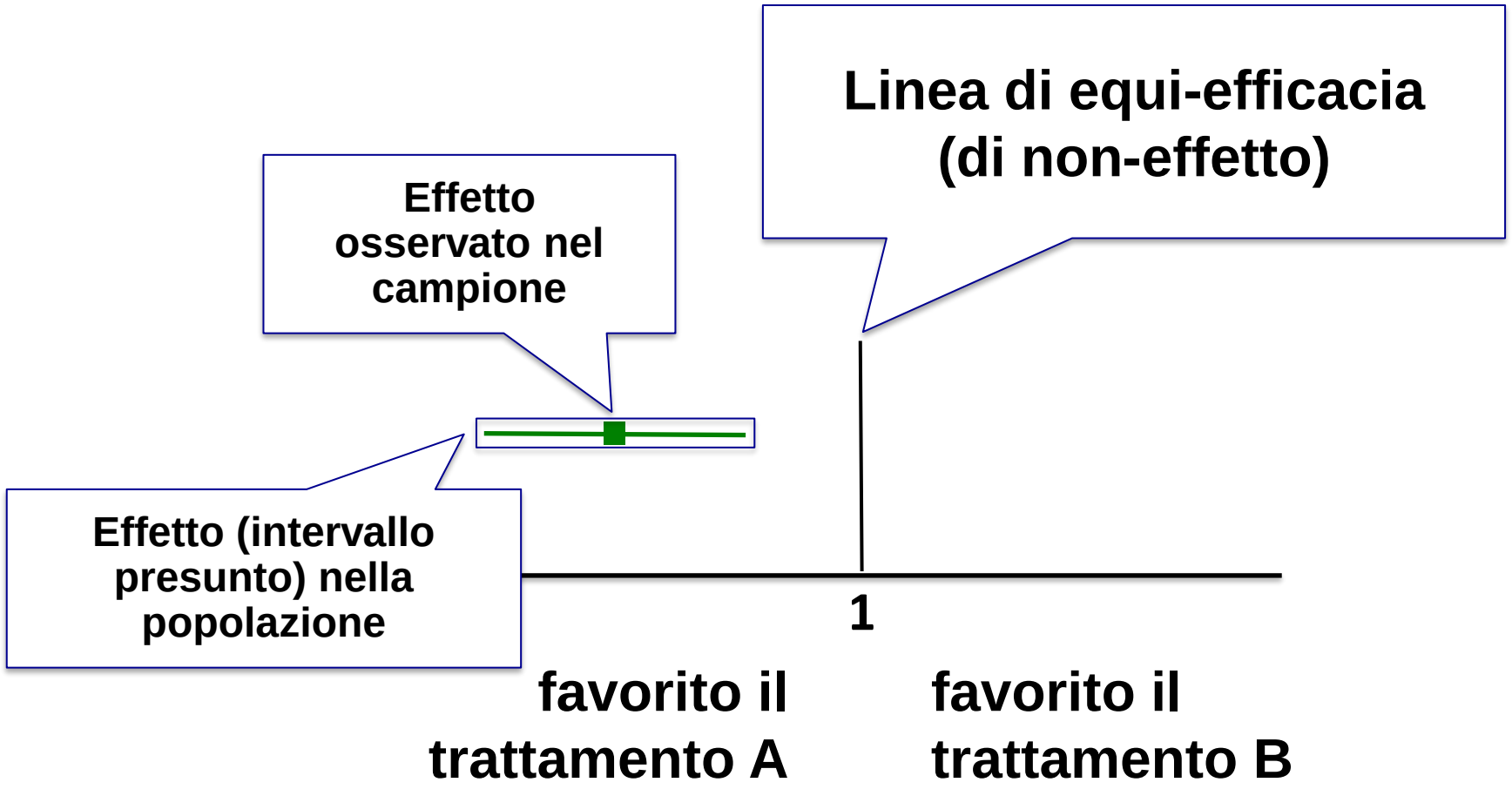
ERRORI STATISTICI

- ✓ **Errore di 1° tipo** (errore *alfa*)
  - quando si conclude per un'efficacia del trattamento sperimentale, quando non lo è nella realtà; *lo studio è falsamente positivo*.
- ✓ **Errore di 2° tipo** (errore *beta*)
  - quando si conclude per una non efficacia del trattamento sperimentale, quando invece lo è nella realtà; *lo studio è falsamente negativo*.
- ✓ Il calcolo delle dimensioni del campione mira a contenere la dimensione degli errori statistici entro valori accettabili (5% per l'errore di 1° tipo e 20% per l'errore di 2° tipo)

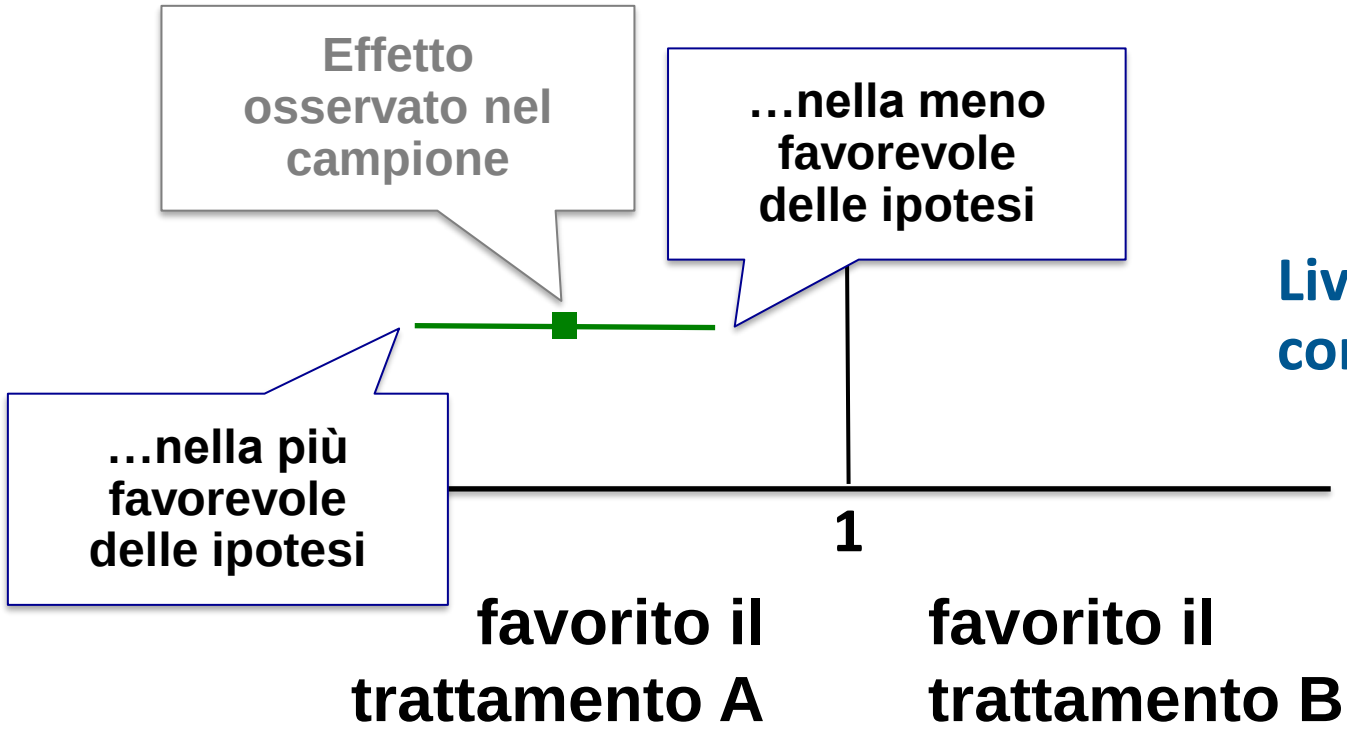
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Livello di confidenza convenzionale: 95%

Statistical Vs Clinical Significance

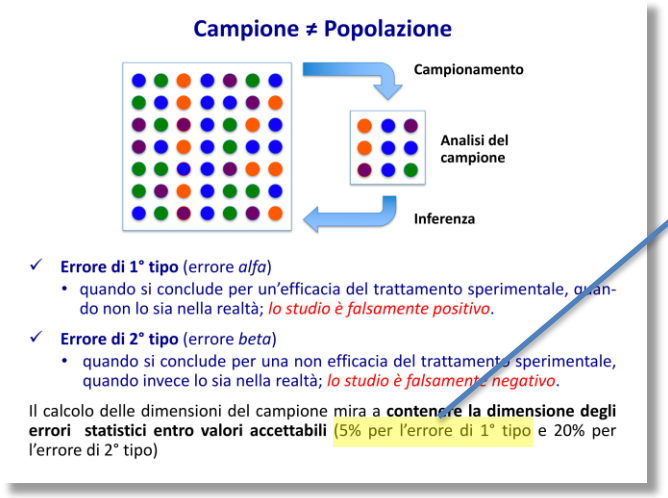
- **Statistical Significance**

  "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

Multiplicity



Probability of at least one false significant result

Number of tests	Probability
1	0.050
2	0.098
5	0.226
10	0.401
50	0.923

Multiple comparisons, **multiplicity** or multiple testing problem occurs when one **considers a set of statistical inferences simultaneously or infers a subset of parameters selected based on the observed values**

modif from Schulz & Grimes; Lancet 2005; 365: 1657–61

Multiplicity is everywhere…

- In subgroup analyses (also when pre-specified)
- In multiple endpoints (this is why the primary endpoint must be pre-specified)
- In interim analyses
- In reanalysis of the same study

Misure per contenere la Multiplicity

- ***Alpha-split, alpha-spending***
  ripartire (tra confronto primario e sotto-gruppi di interesse, oppure tra diversi endpoints) la probabilità convenzionale (5%) di risultato falsamente positivo
- ***Hierarchical approach***
  - stabilito preventivamente l'ordine in cui saranno effettuati i test statistici;
  - la sequenza di test si interromperà al primo risultato statisticamente NON significativo.

  La non indipendenza e la non simultaneità dei test permetterà di contenere il rischio di risultato falsamente positivo entro il convenzionale 5%.

Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma

R.J. Motzer, N.M. Tannir, D.F. McDermott, O. Arén Frontera, B. Melichar, T.K. Choueiri, E.R. Plimack, P. Barthélémy, C. Porta, S. George, T. Powles, F. Donskov, V. Neiman, C.K. Kollmannsberger, P. Salman, H. Gurney, R. Hawkins, A. Ravaud, M.-O. Grimm, S. Bracarda, C.H. Barrios, Y. Tomita, D. Castellano, B.I. Rini, A.C. Chen, S. Mekan, M.B. McHenry, M. Wind-Rotolo, J. Doan, P. Sharma, H.J. Hammers, and B. Escudier, for the CheckMate 214 Investigators*

N Engl J Med 2018;378:1277-90.

The overall alpha level was 0.05, split among three coprimary end points. The objective response rate was analyzed at an alpha level of 0.001. Progression-free survival was evaluated at an alpha level of 0.009, with a power of 80% or more. We evaluated overall survival at an alpha level of 0.04 with 90% power (independent of coprimary end points) on the basis of a hazard ratio of 0.77, accounting for two formal interim analyses after 51% (reported herein) and 75% of deaths had occurred, using a stratified log-rank test.

Multiplicity in randomised trials II: subgroup and interim analyses

Kenneth F Schulz, David A Grimes

Lancet 2005; 365: 1657–61

Number of planned interim analyses	Interim analysis	Pocock	Peto	O'Brien-Fleming
2	1	0·029	0·001	0·005
	2 (final)	0·029	0·05	0·048
3	1	0·022	0·001	0·0005
	2	0·022	0·001	0·014
	3 (final)	0·022	0·05	0·045
4	1	0·018	0·001	0·0001
	2	0·018	0·001	0·004
	3	0·018	0·001	0·019
	4 (final)	0·018	0·05	0·043
5	1	0·016	0·001	0·00001
	2	0·016	0·001	0·0013
	3	0·016	0·001	0·008
	4	0·016	0·001	0·023
	5 (final)	0·016	0·05	0·041

Overall α=0·05.

Table 2: Interim stopping levels (p values) for different numbers of planned interim analyses by group sequential design[14,15]

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]

0.05 for PFS

Single final analysis

0.05 for OS

0.011 at first interim

0.025 at second interim

0.041 at final

0.05 for ORR

Single final analysis

Multiplicity is everywhere…

- In **subgroup analyses** (also when pre-specified)
- In **multiple endpoints** (this is why the primary endpoint must be pre-specified)
- In **interim** analyses
- In **reanalysis** of the same study

Misure per contenere la Multiplicity

- ***Alpha-split, alpha-spending***
  ripartire (tra confronto primario e sotto-gruppi di interesse, oppure tra diversi endpoints) la probabilità convenzionale (5%) di risultato falsamente positivo

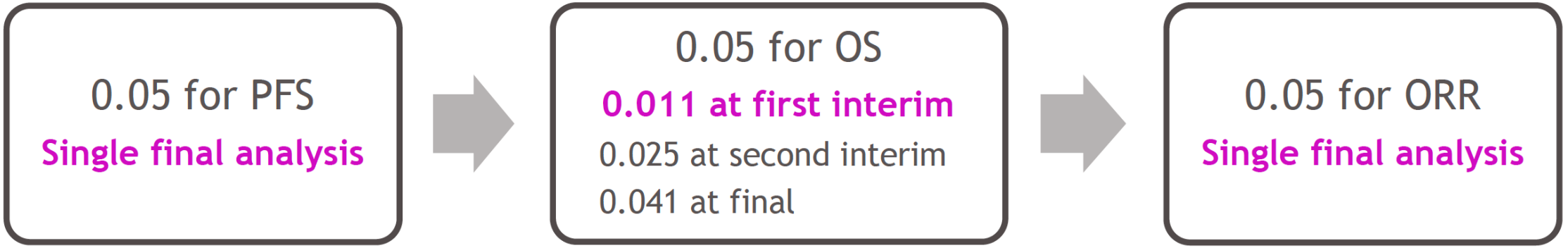
- ***Hierarchical approach***
  - stabilito preventivamente l'ordine in cui saranno effettuati i test statistici;
  - la sequenza di test si interromperà al primo risultato statisticamente NON significativo.

  La non indipendenza e la non simultaneità dei test permetterà di contenere il rischio di risultato falsamente positivo entro il convenzionale 5%.

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]



Annals of Oncology (2020) 31 (suppl_4): S1142-S1215. 10.1016/annonc/annonc325

Encorafenib plus binimetinib versus vemurafenib or encorafenib in patients with *BRAF*-mutant melanoma (COLUMBUS): a multicentre, open-label, randomised phase 3 trial

Reinhard Dummer, Paolo A Ascierto*, Helen J Gogas, Ana Arance, Mario Mandala, Gabriella Liszkay, Claus Garbe, Dirk Schadendorf, Ivana Krajsova, Ralf Gutzmer, Vanna Chiarion-Sileni, Caroline Dutriaux, Jan Willem B de Groot, Naoya Yamazaki, Carmen Loquai, Laure A Moutouh-de Parseval, Michael D Pickard, Victor Sandor, Caroline Robert†, Keith T Flaherty†*

Lancet Oncol 2018; 19: 603–15

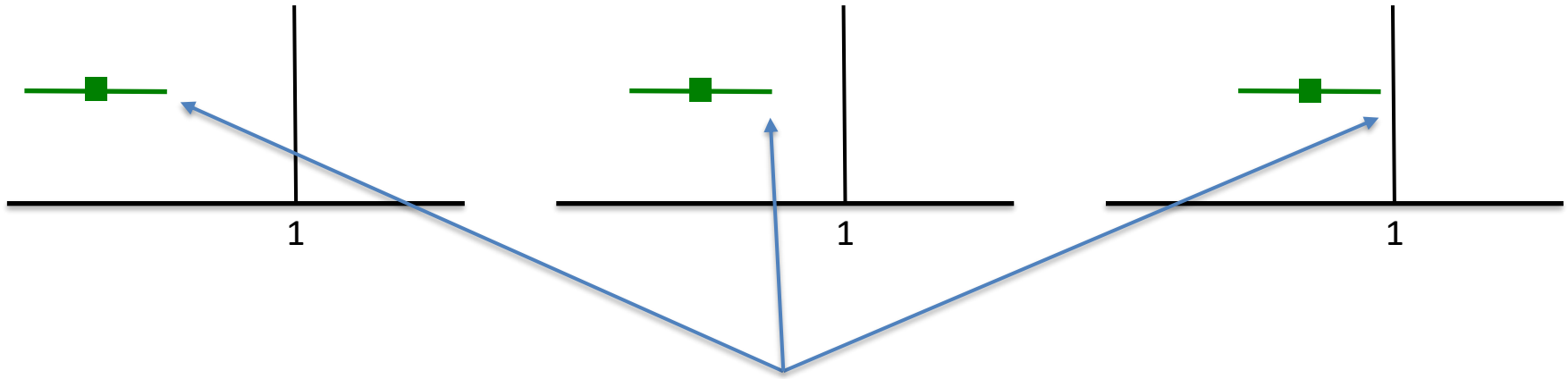
Statistical analysis

...

We used a hierarchical testing procedure to control type-I errors for the primary and key secondary endpoints. The key secondary endpoint of part one (progression-free survival of encorafenib plus binimetinib vs encorafenib) was to be tested if the primary efficacy endpoint (progression-free survival of encorafenib plus binimetinib vs vemurafenib) was significant.

Interpretazione statistica di uno Studio di Superiorità

Tutti e tre gli esempi indicano una differenza statisticamente significativa

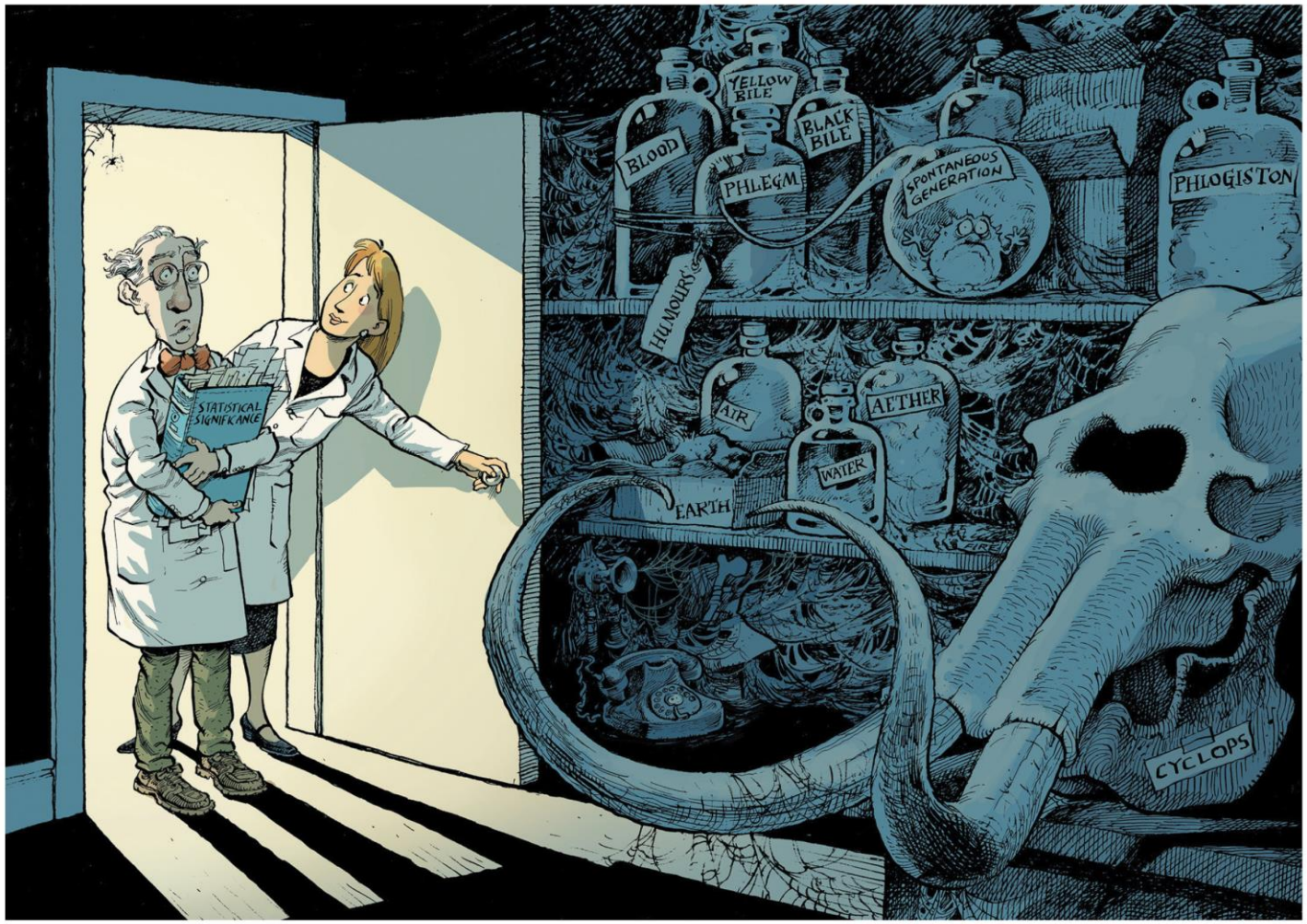


L'estremo dx dell'intervallo di confidenza (la "meno favorevole delle ipotesi") NON interseca la linea di non-effetto (P<0.05)

(ottenimento di un risultato statisticamente significativo)

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

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- **Interpretazione clinica dei risultati dello studio**
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

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

J Natl Cancer Inst 2011;103:16–20

When Are "Positive" Clinical Trials in Oncology Truly Positive?

Alberto Ocana, Ian F. Tannock

What Constitutes a Positive Clinical Trial in Oncology?

We would define a positive trial as one in which the predefined value of δ represents a clinically important difference in an end-point that directly reflects benefit (mainly OS or quality of life) to patients and for which the results provide a best estimate of the difference that exceeds that predefined value of δ.

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

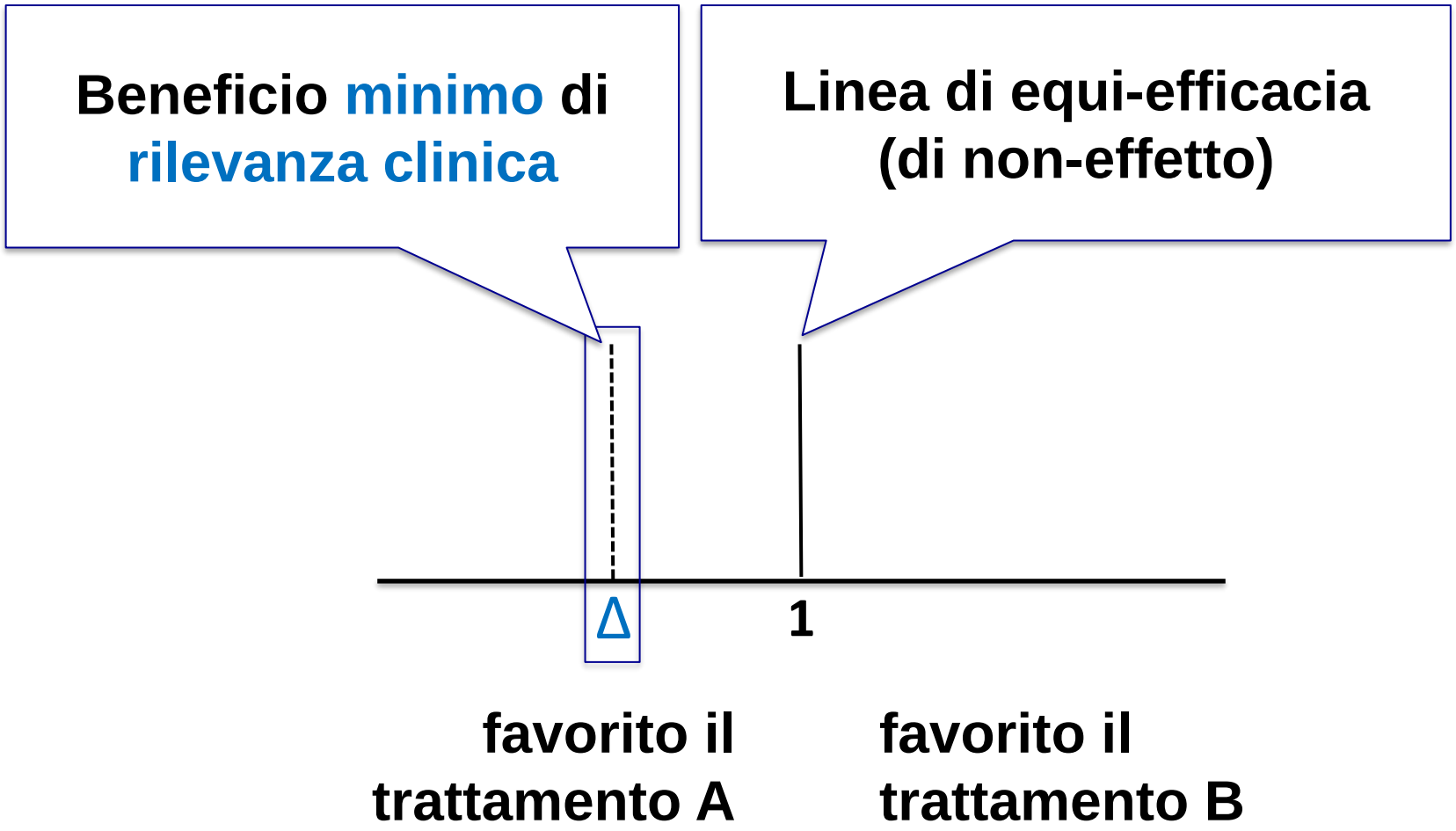
studio di superiorità

A > B di una quantità Δ di interesse clinico

studio di non inferiorità

A < B non oltre una quantità M di rilevanza clinica

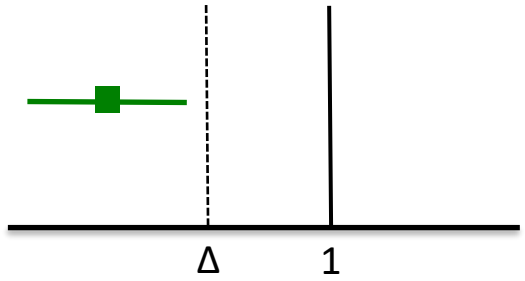
Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione clinica di uno Studio di Superiorità

Effetto (sempre) clinicamente rilevante?

(dato uno specifico Δ di interesse)



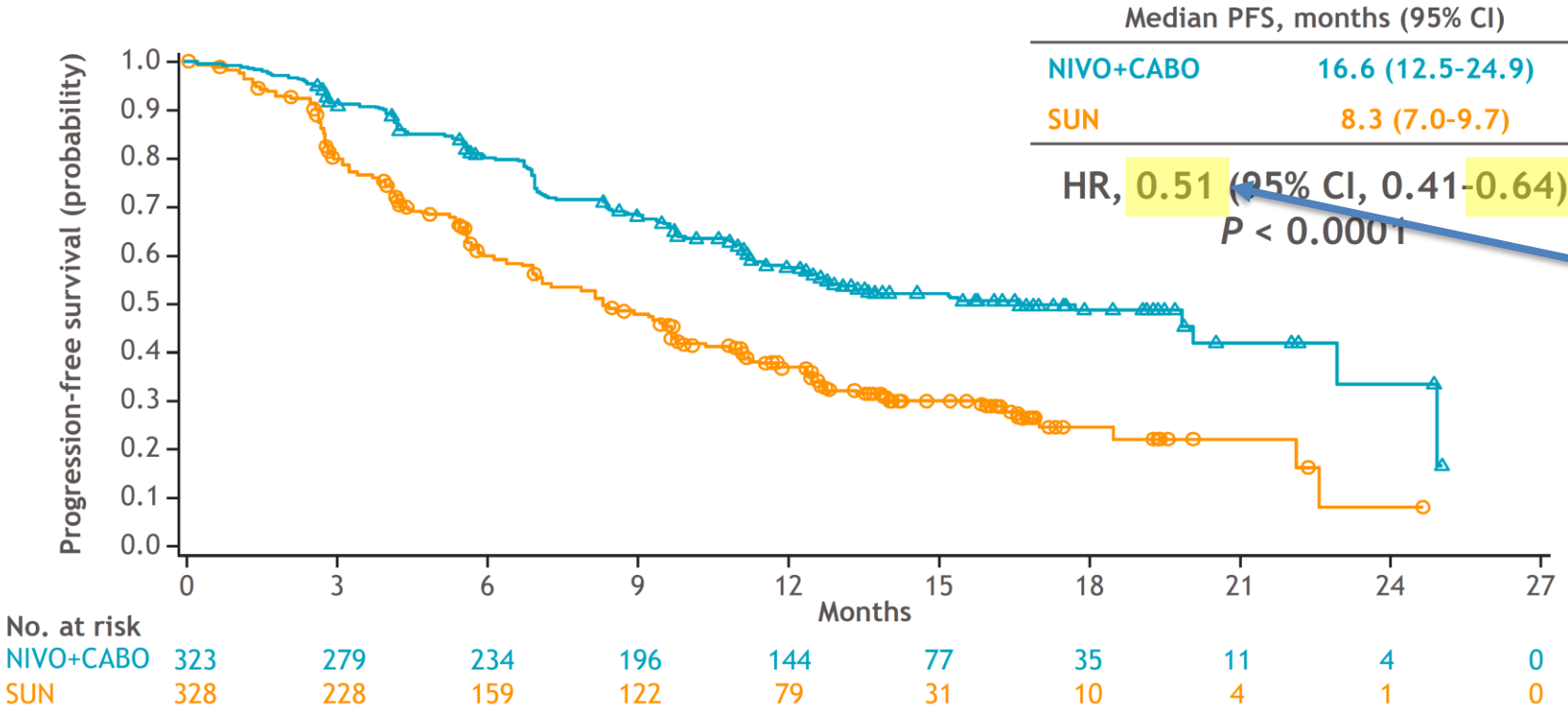
RILEVANTE e
(del tutto) **AFFIDABILE**

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

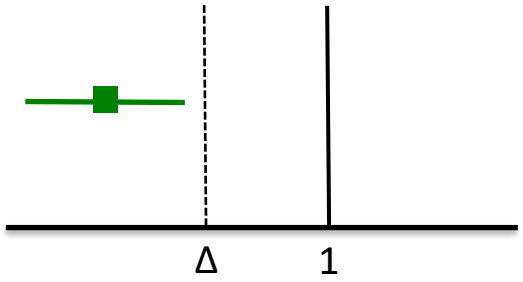


It was estimated that ~350 progression or death events would result at least 95% power to detect a hazard ratio of 0.68

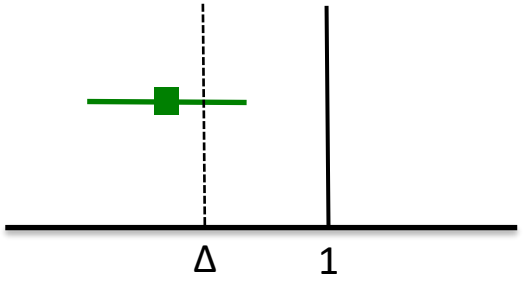
Interpretazione clinica di uno Studio di Superiorità

Effetto (sempre) clinicamente rilevante?

(dato uno specifico Δ di interesse)



RILEVANTE e
(del tutto) **AFFIDABILE**

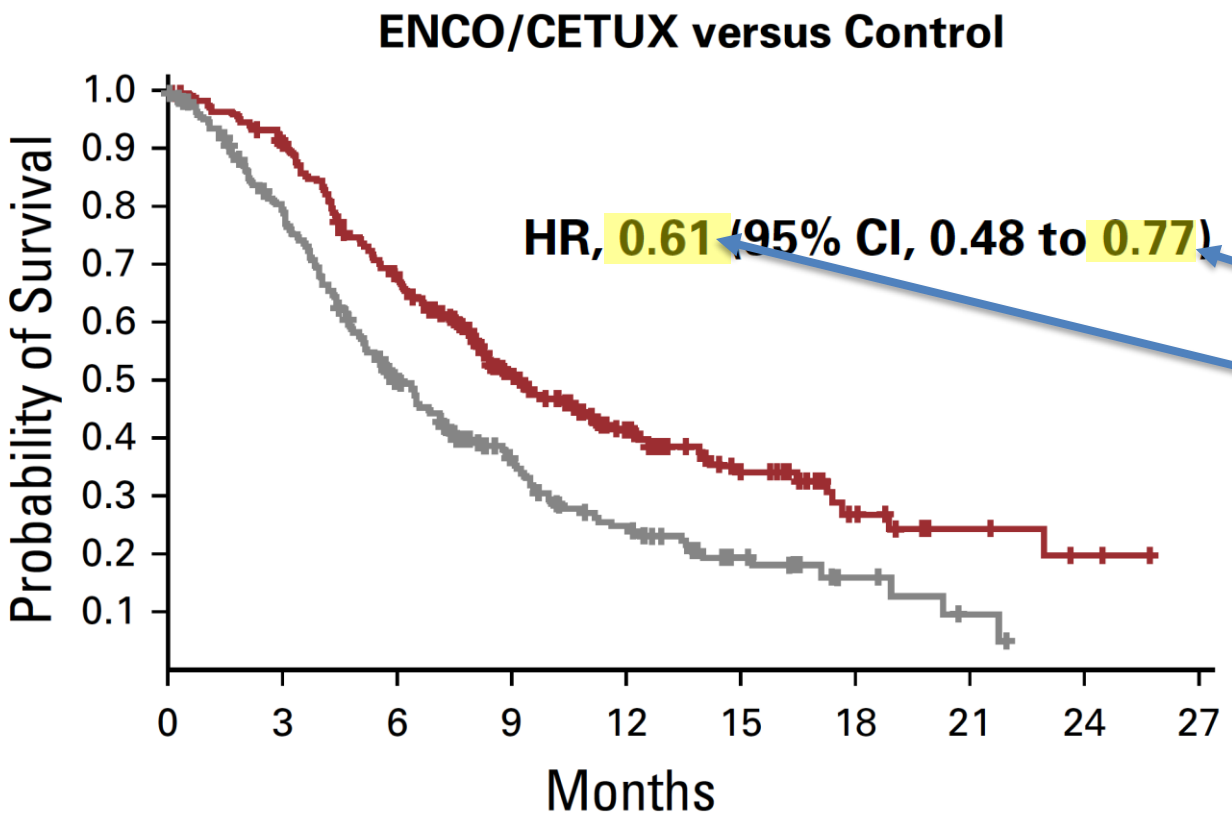


RILEVANTE e
(ragionevolmente) **AFFIDABILE**

Encorafenib Plus Cetuximab as a New Standard of Care for Previously Treated *BRAF* V600E–Mutant Metastatic Colorectal Cancer: Updated Survival Results and Subgroup Analyses from the BEACON Study

Josep Tabernero, MD, PhD[1]; Axel Grothey, MD[2]; Eric Van Cutsem, MD, PhD[3]; Rona Yaeger, MD[4]; Harpreet Wasan, MD[5]; Takayuki Yoshino, MD, PhD[6]; Jayesh Desai, MBBS[7]; Fortunato Ciardiello, MD, PhD[8]; Fotios Loupakis, MD, PhD[9]; Yong Sang Hong, MD, PhD[10]; Neeltje Steeghs, MD, PhD[11]; Tormod Kyrre Guren, MD, PhD[12]; Hendrik-Tobias Arkenau, MD, PhD[13]; Pilar Garcia-Alfonso, MD[14]; Elena Elez, MD, PhD[1]; Ashwin Gollerkeri, MD[15]; Kati Maharry, PhD[15]; Janna Christy-Bittel, MSN[15]; and Scott Kopetz, MD, PhD[16]

J Clin Oncol 39:273-284.

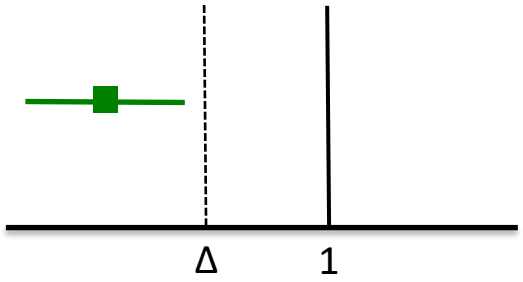


The sample size was driven by the secondary end point of overall survival in the doublet-therapy group as compared with the control group. For this comparison, we calculated that 338 deaths would be required to give the trial 90% power to detect a hazard ratio for death of 0.70, with the use of a stratified log-rank test at a one-sided significance level of 0.025. The number of

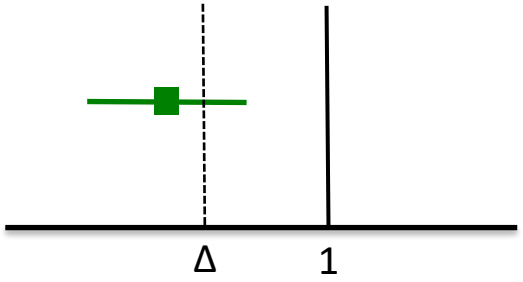
Interpretazione clinica di uno Studio di Superiorità

Effetto (sempre) clinicamente rilevante?

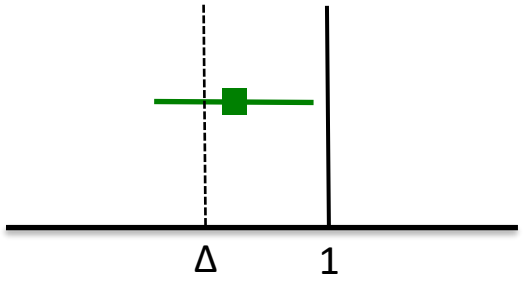
(dato uno specifico Δ di interesse)



RILEVANTE e
(del tutto) **AFFIDABILE**



RILEVANTE e
(ragionevolmente) **AFFIDABILE**



STATISTICAMENTE SIGNIFICATIVO

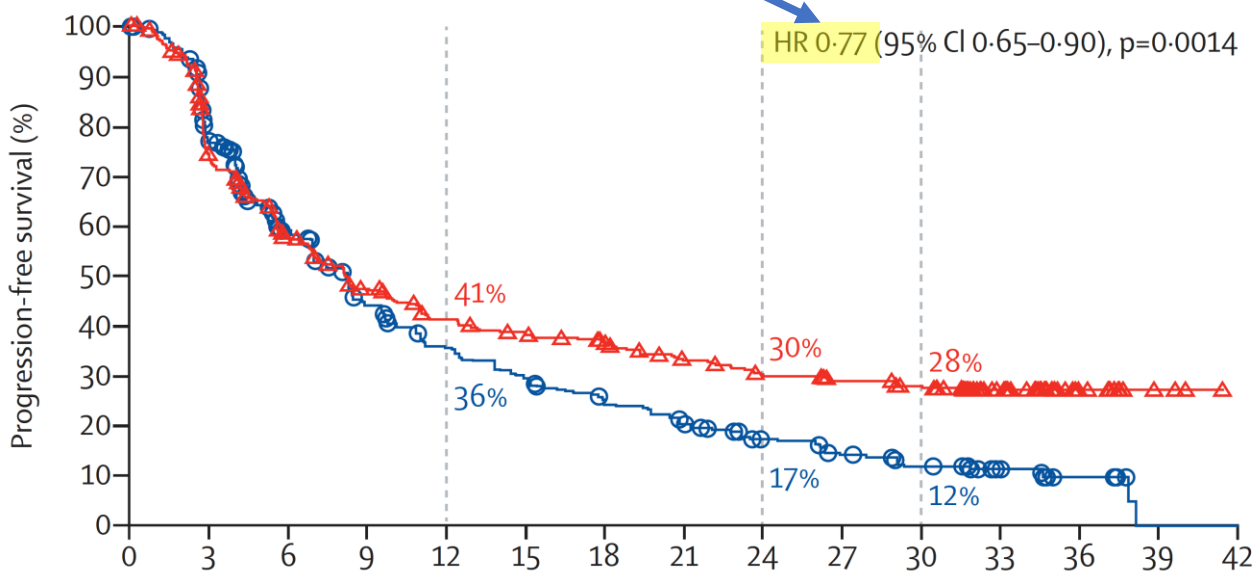
Sample size justification for PFS comparison

For this comparison of PFS, it will be required to observe at least 583 PFS events among the randomized intermediate/poor risk subjects in the two respective treatment arms for a two-sided experiment-wise $\alpha = 0.01$ log-rank test, to show a statistically significant difference in PFS between the treatment arms with at least 90% power when the true hazard ratio of the experimental arm to control arm is 0.73.

Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial

*Robert J Motzer, Brian I Rini, David F McDermott, Osvaldo Arén Frontera, Hans J Hammers, Michael A Carducci, Pamela Salman, Bernard Escudier, Benoit Beuselinck, Asim Amin, Camillo Porta, Saby George, Victoria Neiman, Sergio Bracarda, Scott S Tykodi, Philippe Barthélémy, Raya Leibowitz-Amit, Elizabeth R Plimack, Sjoukje F Oosting, Bruce Redman, Bohuslav Melichar, Thomas Powles, Paul Nathan, Stéphane Oudard, David Pook, Toni K Choueiri, Frede Donskov, Marc-Oliver Grimm, Howard Gurney, Daniel Y C Heng, Christian K Kollmannsberger, Michael R Harrison, Yoshihiko Tomita, Ignacio Duran, Viktor Grünwald, M Brent McHenry, Sabeen Mekan, Nizar M Tannir, on behalf of the CheckMate 214 investigators**

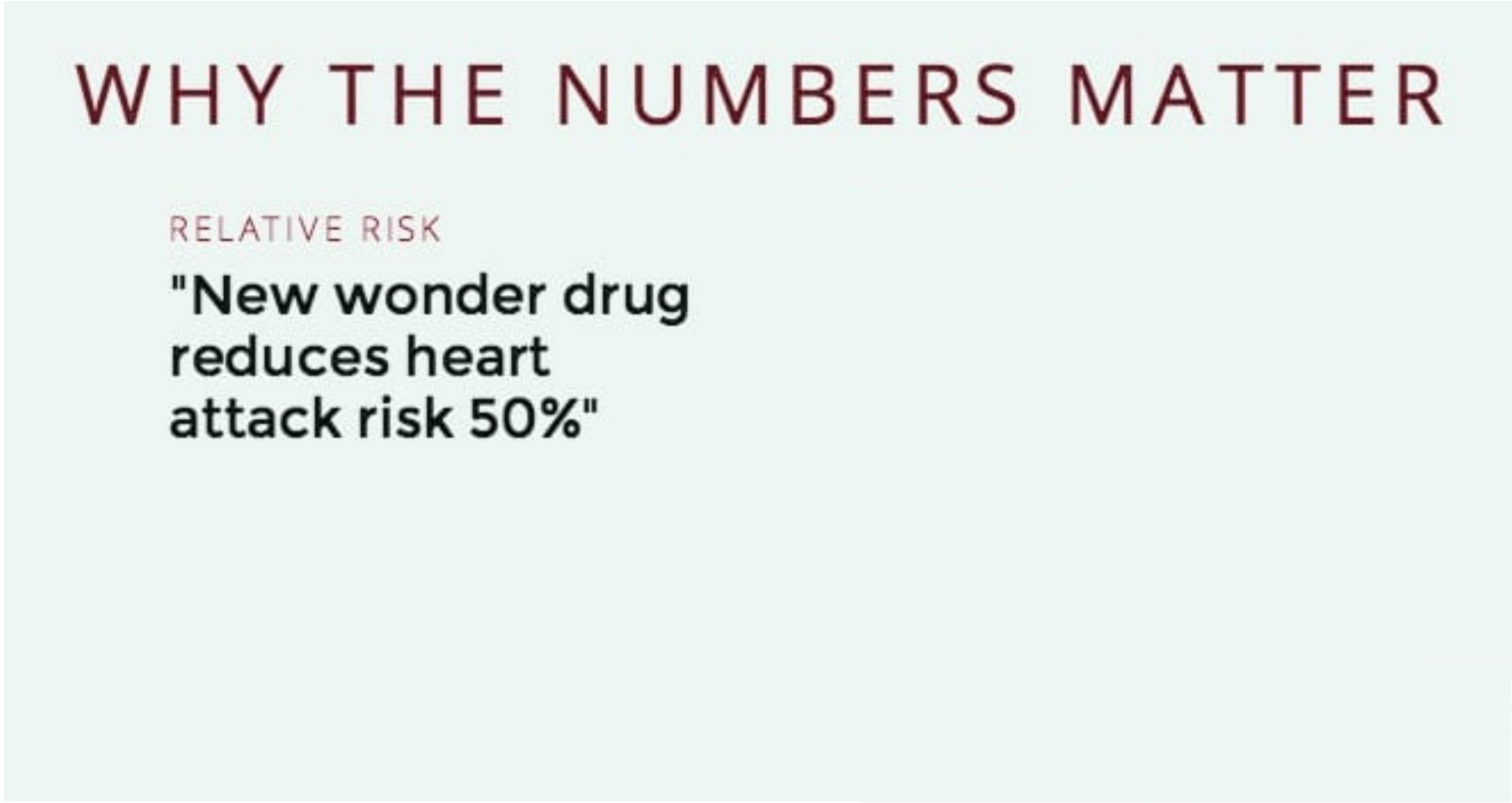
Lancet Oncol 2019; 20: 1370–85



Beneficio RELATIVO (RR, OR , HR) o ASSOLUTO (RD)?

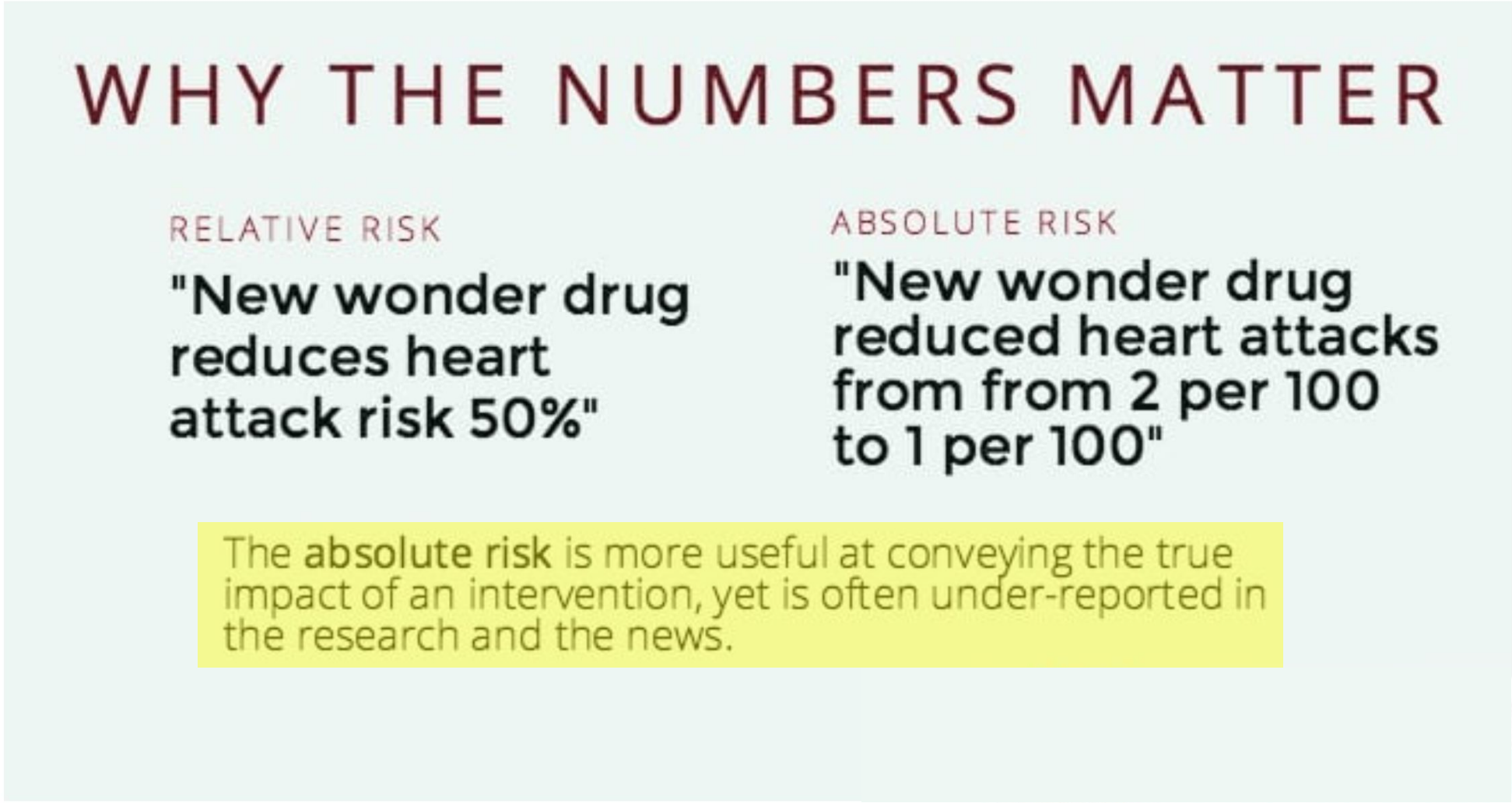
Endpoint	Tipo di variabile statistica	Indicatore comparativo di effetto	
		relativo	assoluto
Attività (CR, PR, SD, ecc.) Tossicità (eventi avversi)	nominale	RR (relative risk) OR (odds ratio)	RD (risk difference)
Punteggi di qualità di vita	intervallare	-	MD (mean difference)
OS (overall survival) PFS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

Beneficio RELATIVO (RR, OR , HR) o ASSOLUTO (RD)?



https://www.healthnewsreview.org/toolkit/tips-for-understanding-studies/absolute-vs-relative-risk/

Beneficio RELATIVO (RR, OR , HR) o ASSOLUTO (RD)?



https://www.healthnewsreview.org/toolkit/tips-for-understanding-studies/absolute-vs-relative-risk/

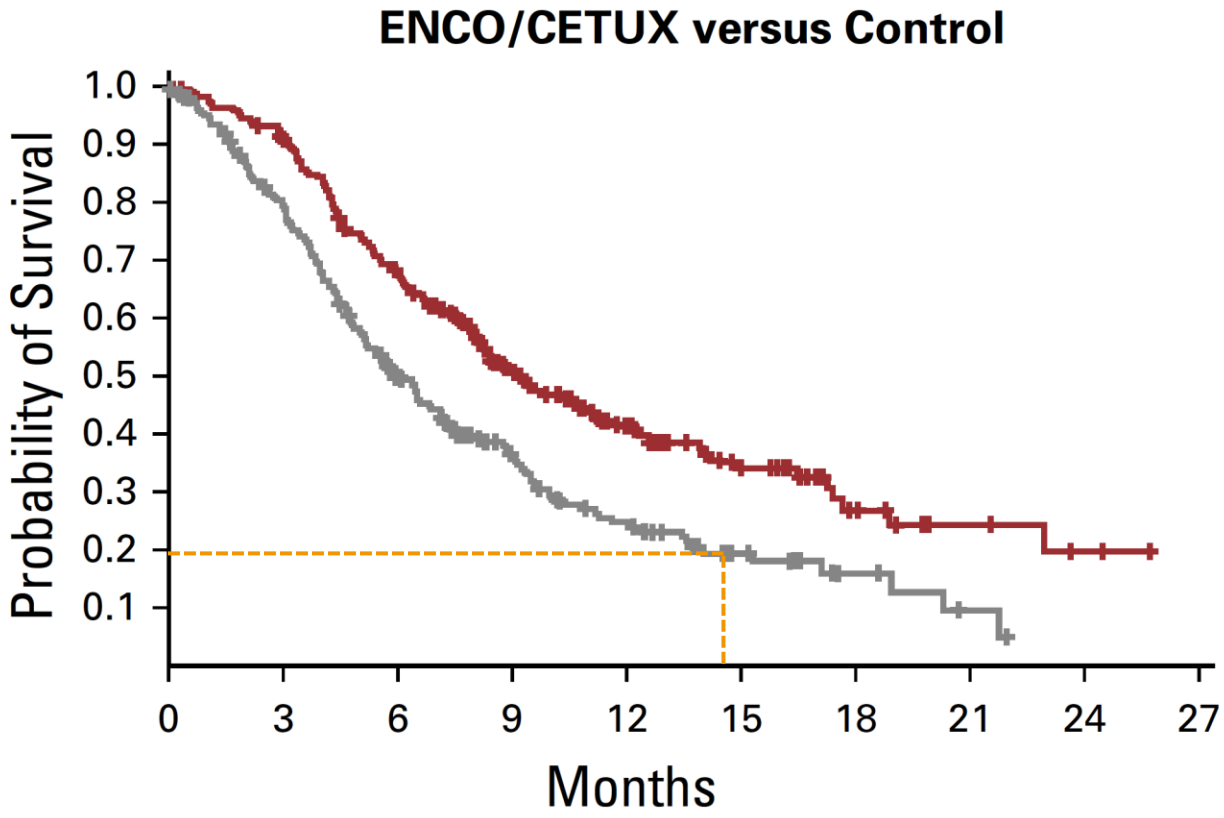
Beneficio RELATIVO (RR, OR , HR) o ASSOLUTO (RD)?

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) - - -	RD (risk difference) differenza tra mediane differenza tra stime al tempo t… RMST (restricted mean survival time)

Encorafenib Plus Cetuximab as a New Standard of Care for Previously Treated *BRAF* V600E–Mutant Metastatic Colorectal Cancer: Updated Survival Results and Subgroup Analyses from the BEACON Study

Josep Tabernero, MD, PhD[1]; Axel Grothey, MD[2]; Eric Van Cutsem, MD, PhD[3]; Rona Yaeger, MD[4]; Harpreet Wasan, MD[5]; Takayuki Yoshino, MD, PhD[6]; Jayesh Desai, MBBS[7]; Fortunato Ciardiello, MD, PhD[8]; Fotios Loupakis, MD, PhD[9]; Yong Sang Hong, MD, PhD[10]; Neeltje Steeghs, MD, PhD[11]; Tormod Kyrre Guren, MD, PhD[12]; Hendrik-Tobias Arkenau, MD, PhD[13]; Pilar Garcia-Alfonso, MD[14]; Elena Elez, MD, PhD[1]; Ashwin Gollerkeri, MD[15]; Kati Maharry, PhD[15]; Janna Christy-Bittel, MSN[15]; and Scott Kopetz, MD, PhD[16]

J Clin Oncol 39:273-284.



- Median f.u.: 14.7 months
- Baseline risk* at median f.u.: 80%
- **Risk Difference**: **17 events lower** / 100 pts (95%CI: 26 lower to 9 lower)
- **NNT ~6**

** J Clin Epidemiol 118 (2020) 124-131*

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

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.

European Journal of Cancer 152 (2021) 116–128

Quality of life in patients with BRAF-mutant melanoma receiving the combination encorafenib plus binimetinib: Results from a multicentre, open-label, randomised, phase III study (COLUMBUS)

Helen Gogas [a,*], Reinhard Dummer [b], Paolo A. Ascierto [c], Ana Arance [d], Mario Mandalà [e], Gabriella Liszkay [f], Claus Garbe [g], Dirk Schadendorf [h,i], Ivana Krajsová [j], Ralf Gutzmer [k], Vanna Chiarion Sileni [l], Caroline Dutriaux [m], Naoya Yamazaki [n], Carmen Loquai [o], Paola Queirolo [p,v], Groot Jan de Willem [q], Abir Tadmouri Sellier [r], Jeanne Suissa [r], Juliette Murris [r], Ashwin Gollerkeri [s], Caroline Robert [t], Keith T. Flaherty [u]

When all cycles were considered, the encorafenib plus binimetinib group was associated with an estimated 3.03 points higher post-baseline score for the FACT-M and an estimated 5.28 points higher postbaseline score for the EORTC QLQ-C30, compared to the vemurafenib group.

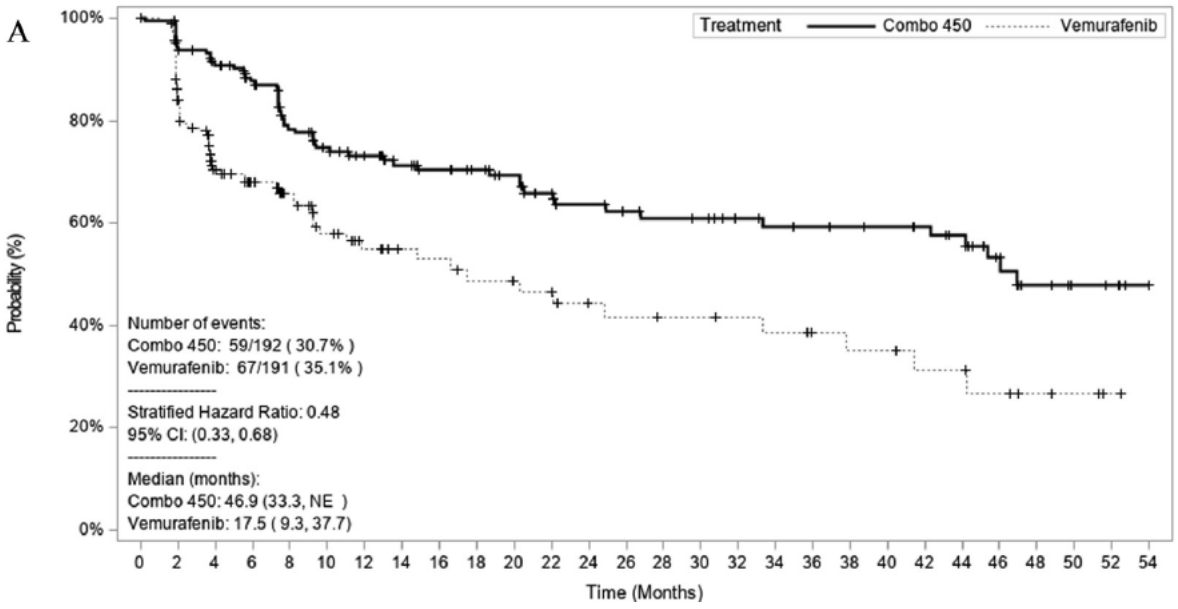
Being above the recommended cutoff values for MCID (2 points[1] for FACT-M subscale and 5 points[2] for EORTC QLQC30), the differences in post baseline FACT-M and EORTC QLQ-C30 scores between groups were considered clinically meaningful.

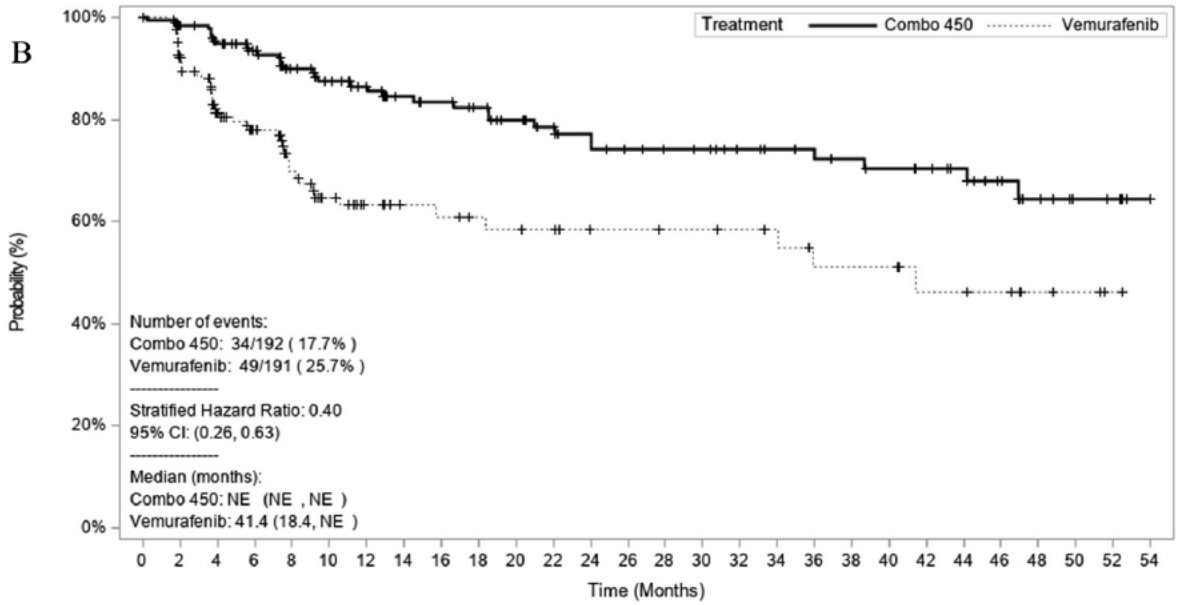
[1] Askew RL, et al. Value Health 2009; 12(8):1144e50. [19]
[2] Musoro JZ, et al. Eur J Canc 2018;104:169e81

European Journal of Cancer 152 (2021) 116–128

Quality of life in patients with BRAF-mutant melanoma receiving the combination encorafenib plus binimetinib: Results from a multicentre, open-label, randomised, phase III study (COLUMBUS)

Helen Gogas [a,*], Reinhard Dummer [b], Paolo A. Ascierto [c], Ana Arance [d], Mario Mandalà [e], Gabriella Liszkay [f], Claus Garbe [g], Dirk Schadendorf [h,i], Ivana Krajsová [j], Ralf Gutzmer [k], Vanna Chiarion Sileni [l], Caroline Dutriaux [m], Naoya Yamazaki [n], Carmen Loquai [o], Paola Queirolo [p,v], Groot Jan de Willem [q], Abir Tadmouri Sellier [r], Jeanne Suissa [r], Juliette Murris [r], Ashwin Gollerkeri [s], Caroline Robert [t], Keith T. Flaherty [u]





Time to definitive 10% deterioration. A, EORTC QLQ-C 30 (Global health status) and B, FACT-M

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

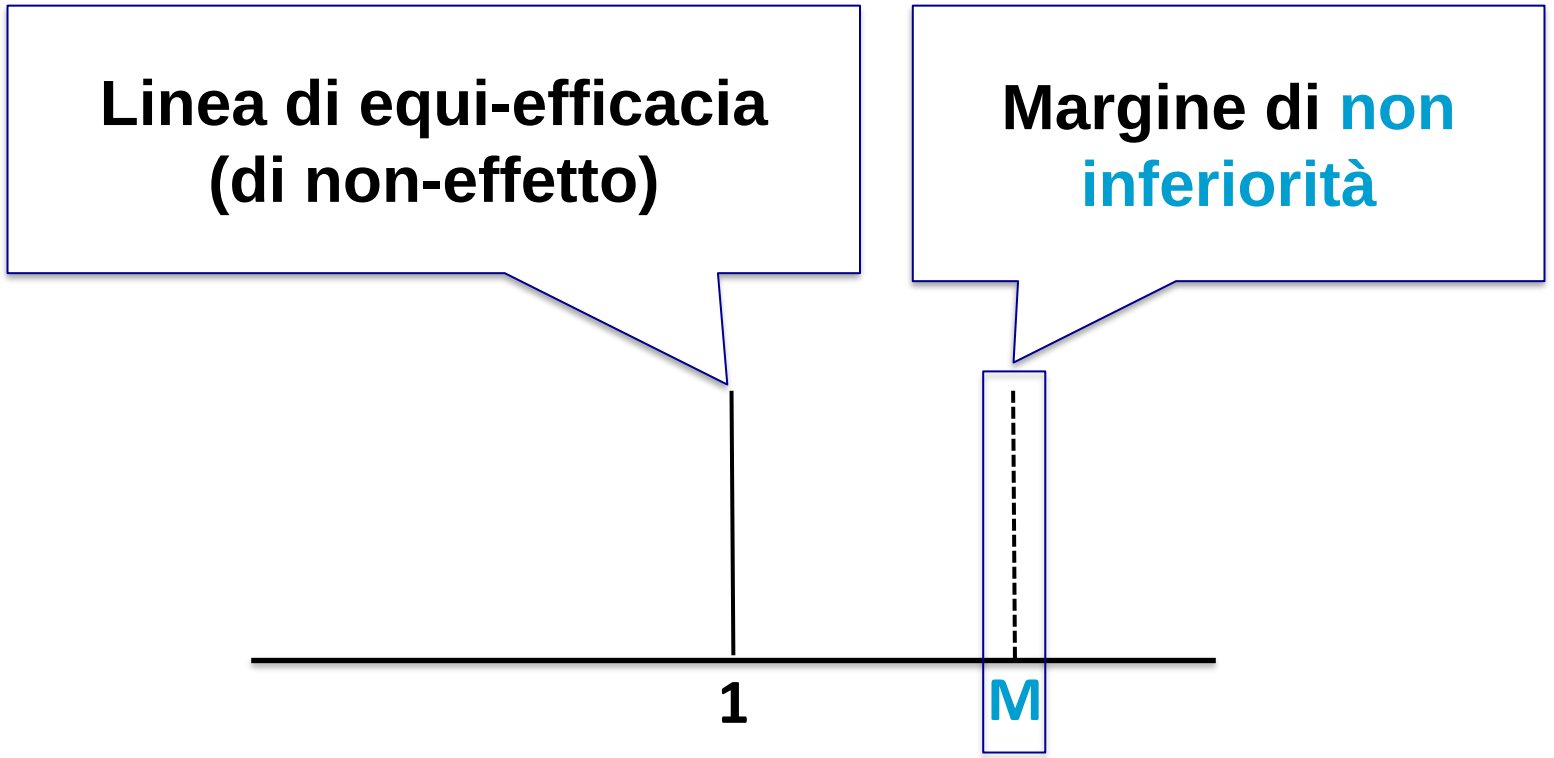
studio di superiorità

A > B di una quantità Δ di interesse clinico

studio di non inferiorità

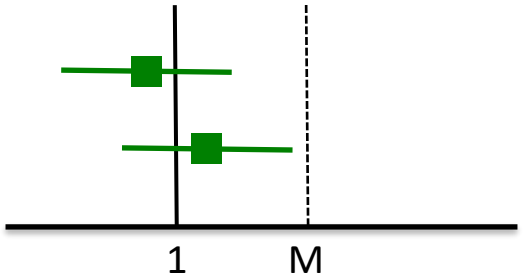
A < B non oltre una quantità M di rilevanza clinica

Interpretazione degli studi clinici mediante Forest (Forrest?) Plot



Interpretazione clinica di uno Studio di Non-Inferiorità

(dato uno specifico M di interesse)



Dimostrazione di Non-Inferiorità

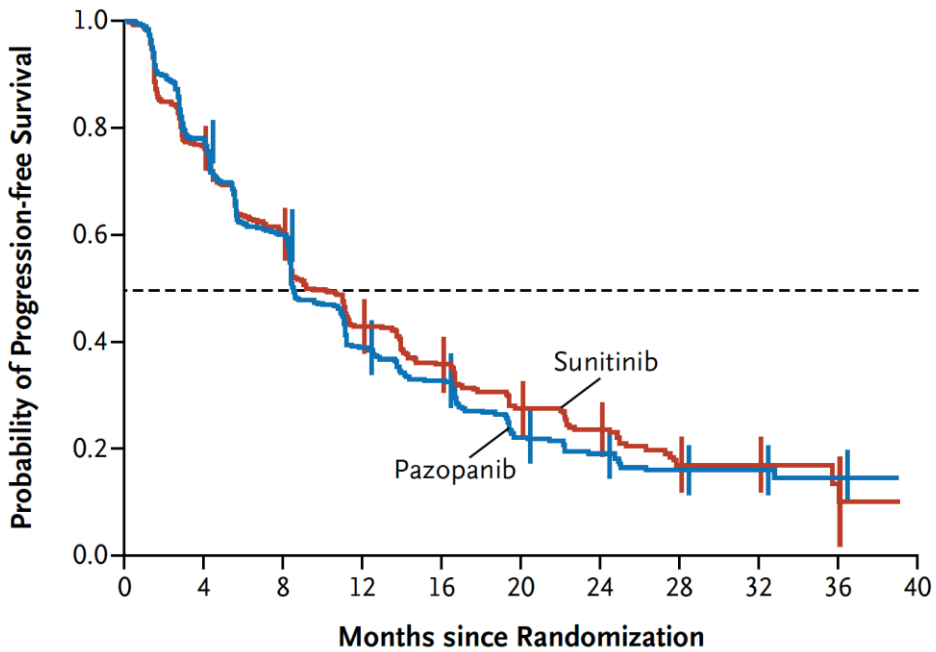
Il limite superiore (*quello a dx*) dell'intervallo di confidenza non interseca la linea corrispondente al margine di non inferiorità M…

…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



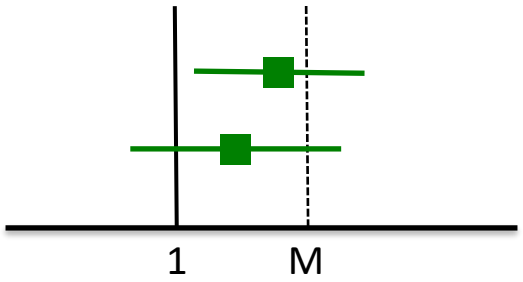
STATISTICAL ANALYSIS

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)

Interpretazione clinica di uno Studio di Non-Inferiorità

(dato uno specifico M di interesse)



Non Dimostrazione di Non-Inferiorità

Il limite superiore (*quello a dx*) dell'intervallo di confidenza interseca la linea corrispondente al margine di non inferiorità M…

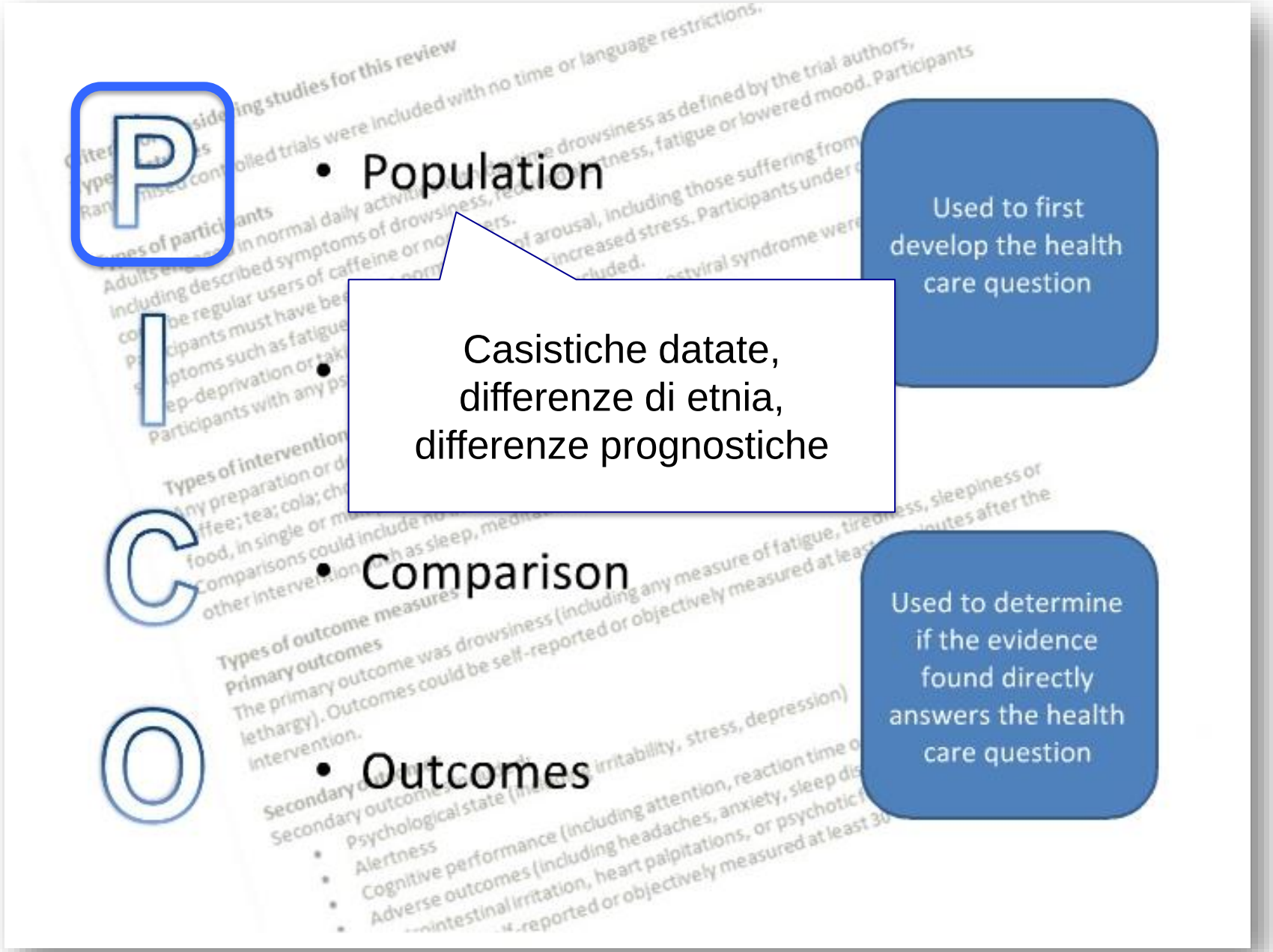
…indipendentemente da dove si colloca il limite inferiore (*quello a sx*)

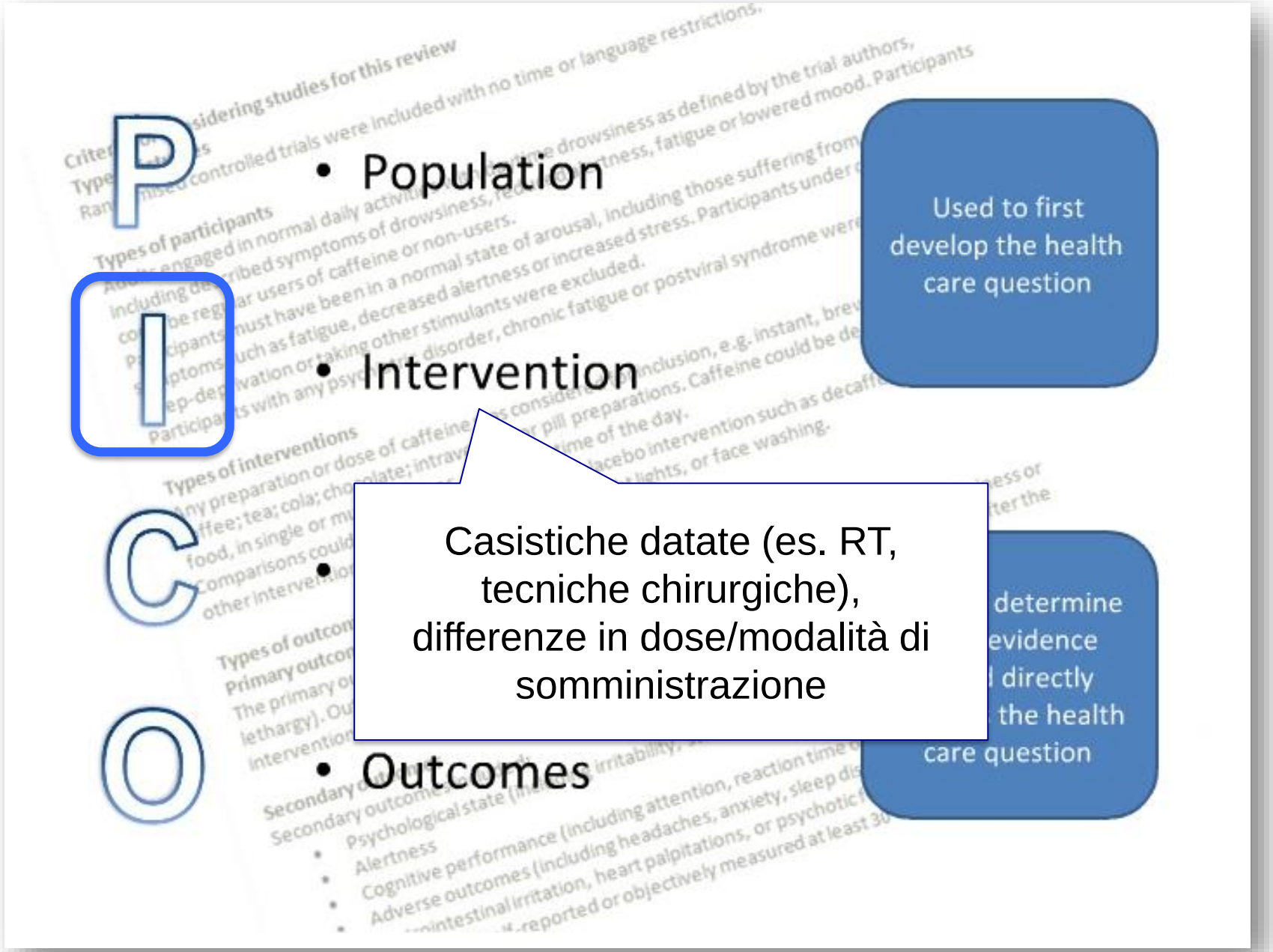
- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- **Trasferibilità dei risultati dello studio**
- Affidabilità dei risultati dello studio
- Analisi per sottogruppi

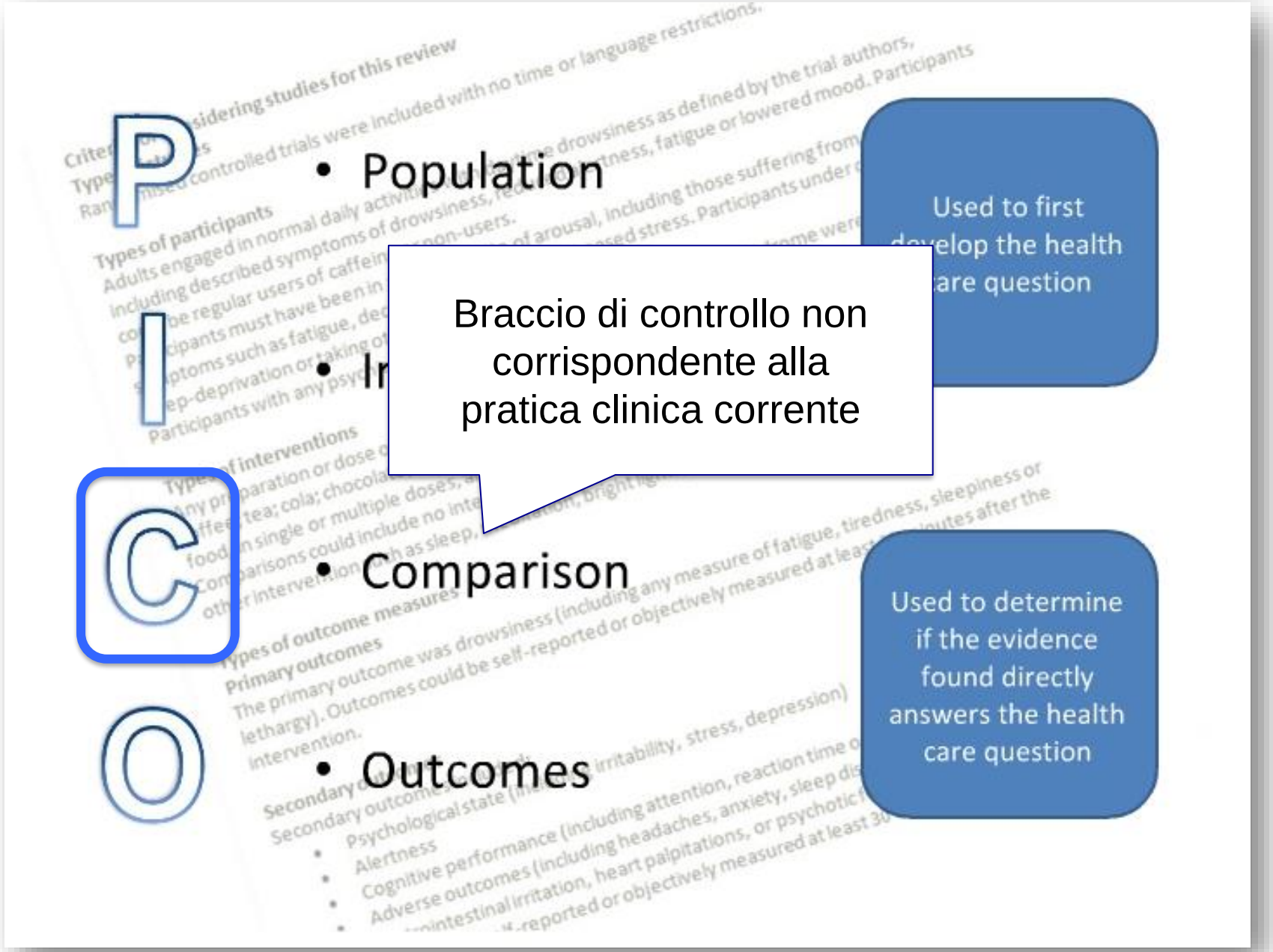
Direct evidence…

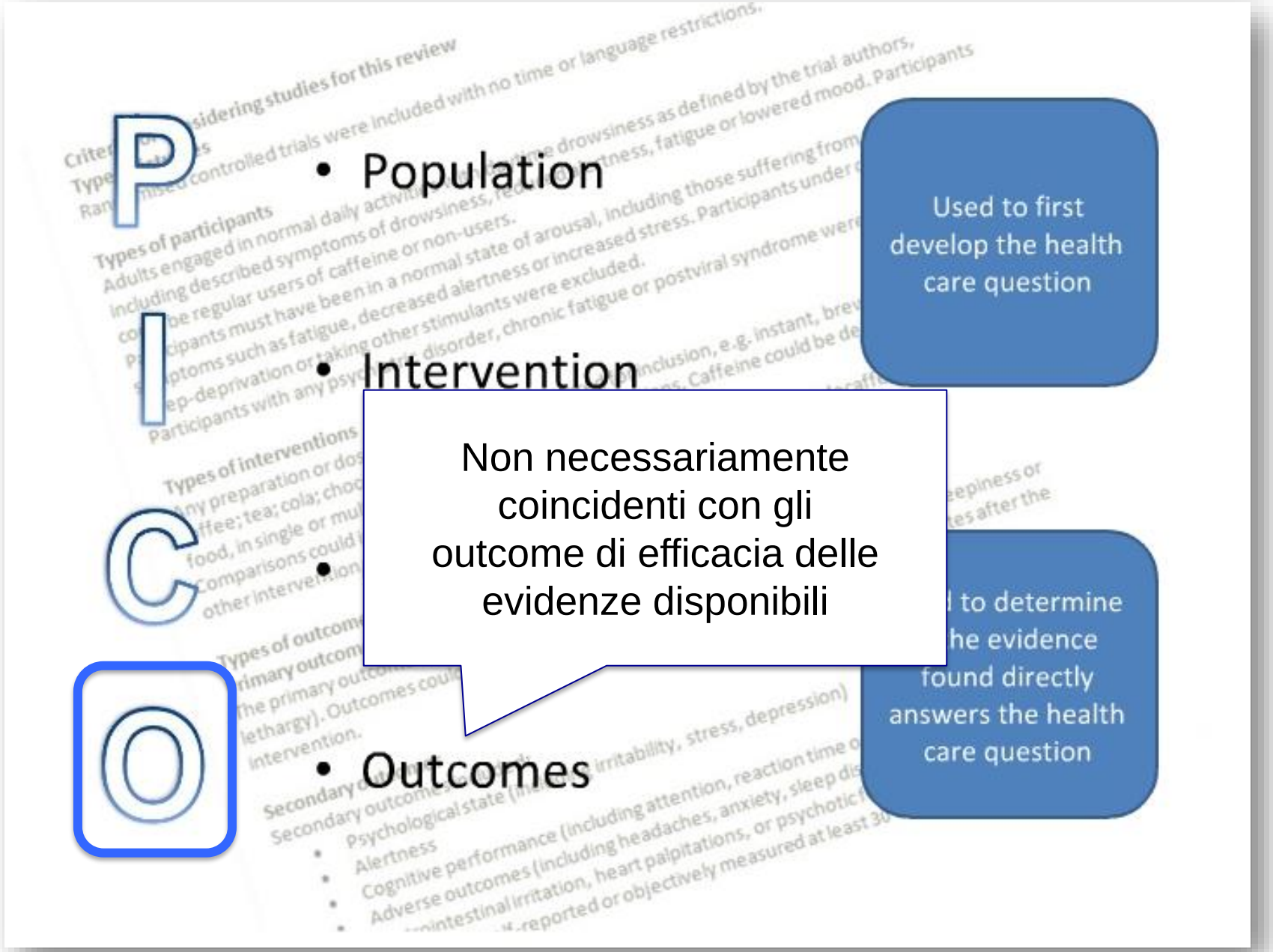
…comes from research that:

- 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


P
• Population
I
•
C
• Comparison
O
• Outcomes
Used to first develop the health care question
Casistiche datate, differenze di etnia, differenze prognostiche
Used to determine if the evidence found directly answers the health care question


P
• Population
I
• Intervention
C
•
O
• Outcomes
Used to first develop the health care question
Casistiche datate (es. RT, tecniche chirurgiche), differenze in dose/modalità di somministrazione
determine evidence directly the health care question


P
I
C
O
• Population
• In
• Comparison
• Outcomes
Braccio di controllo non corrispondente alla pratica clinica corrente
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question


P
I
C
O
• Population
• Intervention
•
• Outcomes
Used to first develop the health care question
Non necessariamente coincidenti con gli outcome di efficacia delle evidenze disponibili
to determine
he evidence
found directly
answers the health
care question

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- **Affidabilità dei risultati dello studio**
- Analisi per sottogruppi

Caso Vs Bias

CASO

Errore Random

Risultati Imprecisi

Errore in diminuzione con l'aumentare delle dimensioni del campione

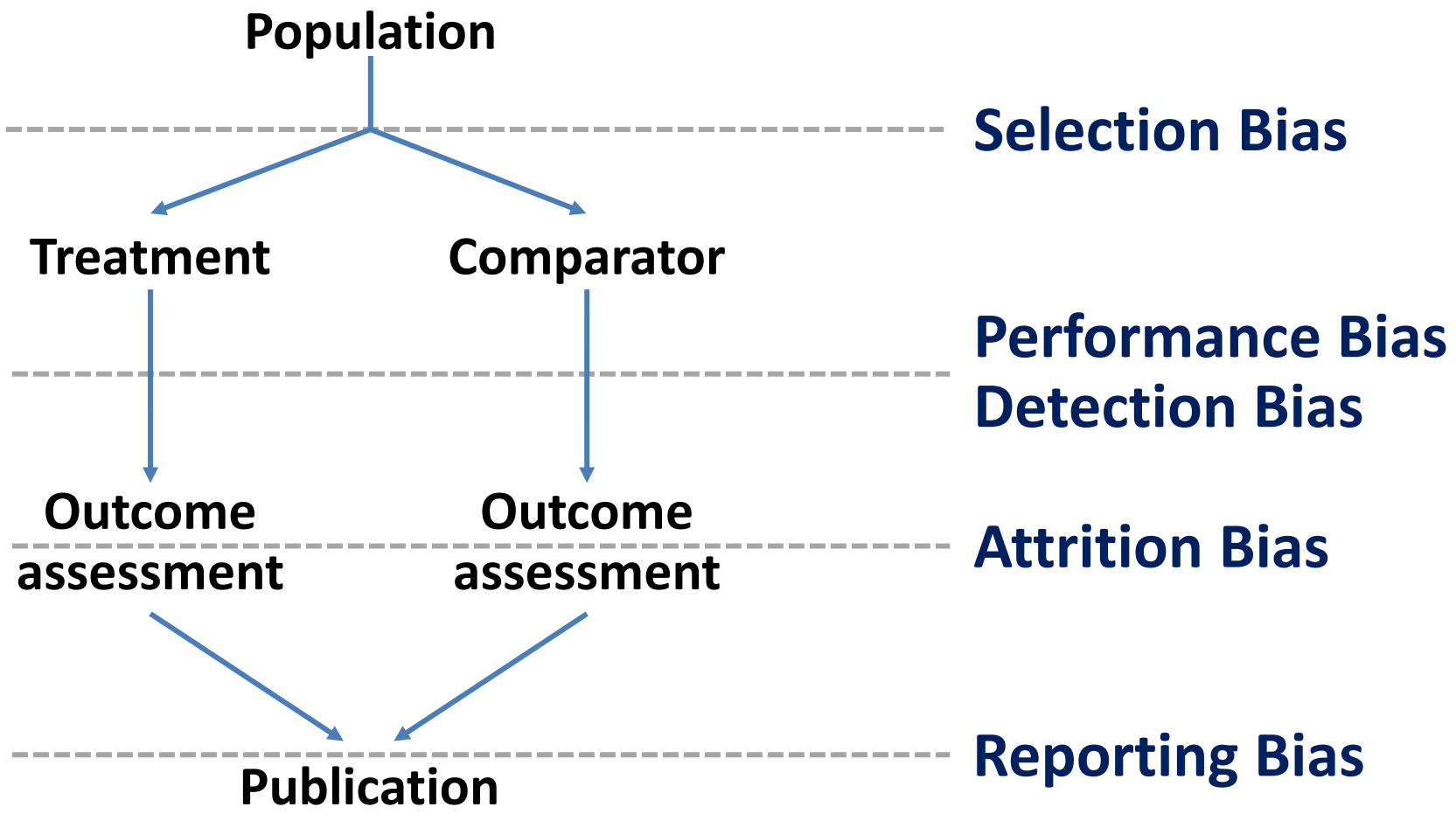
BIAS

Errore Sistematico

Risultati Inesatti

Errore non influenzato dalle dimensioni del campione

Study Flow & Risk of Bias…



SOURCES OF BIAS IN CLINICAL TRIALS

Type of bias	Description	
Selection bias.	Systematic differences between baseline characteristics of the groups that are compared.	• Sequence generation. • Allocation concealment.
Performance bias.	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.	• Blinding of participants and personnel. • Other potential threats to validity.
Detection bias.	Systematic differences between groups in how outcomes are determined.	• Blinding of outcome assessment. • Other potential threats to validity.
Attrition bias.	Systematic differences between groups in withdrawals from a study.	• Incomplete outcome data
Reporting bias.	Systematic differences between reported and unreported findings.	• Selective outcome reporting

Cochrane Handbook for Systematic Reviews of Interventions

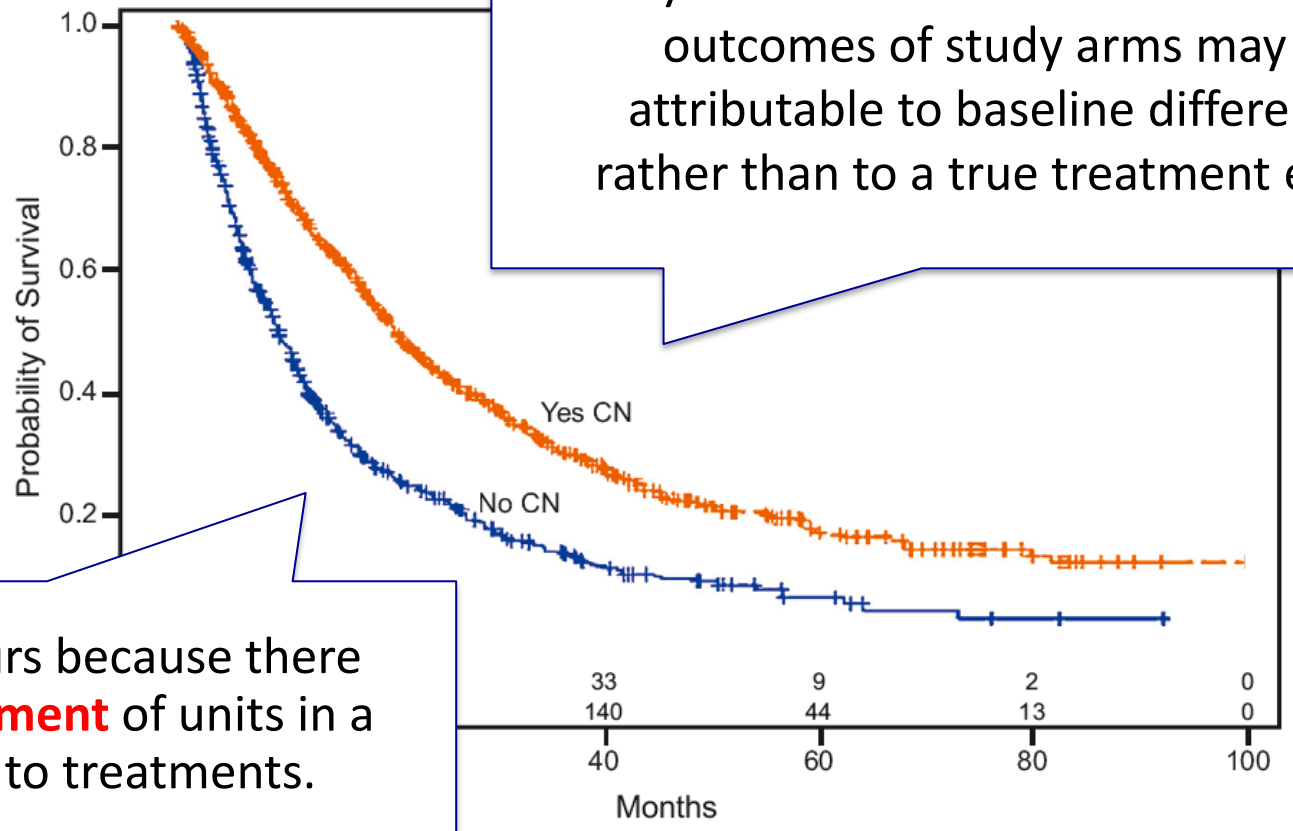
Version 5.1.0

Cytoreductive Nephrectomy in Patients with Synchronous Metastases from Renal Cell Carcinoma: Results from the International Metastatic Renal Cell Carcinoma Database Consortium

Daniel Y.C. Heng[a,,†], J. Connor Wells[a,†], Brian I. Rini[b], Benoit Beuselinck[c], Jae-Lyun Lee[d], Jennifer J. Knox[e], Georg A. Bjarnason[f], Sumanta Kumar Pal[g], Christian K. Kollmannsberger[h], Takeshi Yuasa[i], Sandy Srinivas[j], Frede Donskov[k], Aristotelis Bamias[l], Lori A. Wood[m], D. Scott Ernst[n], Neeraj Agarwal[o], Ulka N. Vaishampayan[p], Sun Young Rha[q], Jenny J. Kim[r], Toni K. Choueiri[s]*

EUROPEAN UROLOGY 66 (2014) 7

Any observed difference between the outcomes of study arms may be attributable to baseline differences rather than to a true treatment effect.



A possible **bias** occurs because there is **no random assignment** of units in a target population to treatments.

Reconciling the Use of Cytoreductive Nephrectomy in the Targeted Therapy Era

Stephen H. Culp *

EUROPEAN UROLOGY 66 (2014) 711–712

Although retrospective, the results of this study are strengthened by the number of patients examined, inclusion of patients from institutions around the world, and lack of patient exclusion based on RCC histology or type of targeted agent.

RANDOMIZATION COMPONENTS

Item	Descriptor
Sequence generation	Method used to generate the random allocation sequence, including details of any restriction (eg, blocking, stratification)
Allocation concealment	Method used to implement the random allocation sequence (eg, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned
Implementation	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups

Phase III Trial of Vinflunine Plus Best Supportive Care Compared With Best Supportive Care Alone After a Platinum-Containing Regimen in Patients With Advanced Transitional Cell Carcinoma of the Urothelial Tract

Joaquim Bellmunt, Christine Théodore, Tomasz Demkov, Boris Komyakov, Lisa Sengelov, Gedske Daugaard, Armelle Caty, Joan Carles, Agnieszka Jagiello-Gruszfeld, Oleg Karyakin, François-Michel Delgado, Patrick Hurteloup, Eric Winquist, Nassim Morsli, Yacine Salhi, Stéphane Culine, and Hans von der Maase

J Clin Oncol 27:4454-4461. © 2009

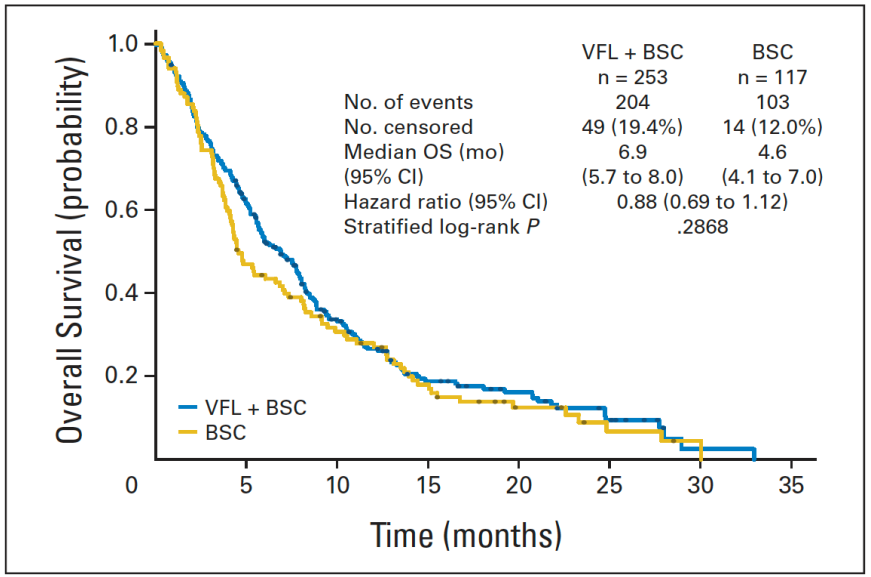


Fig 2. Overall survival (OS) in the intent-to-treat population (n = 370). VFL, vinflunine; BSC, best supportive care.

Although the objective of the median 2-month survival advantage favoring VFL BSC versus BSC was achieved (6.9 v 4.6 months, respectively), this difference was not statistically significant (P .287; Fig 2).

Phase III Trial of Vinflunine Plus Best Supportive Care Compared With Best Supportive Care Alone After a Platinum-Containing Regimen in Patients With Advanced Transitional Cell Carcinoma of the Urothelial Tract

Joaquim Bellmunt, Christine Théodore, Tomasz Demkov, Boris Komyakov, Lisa Sengelov, Gedske Daugaard, Armelle Caty, Joan Carles, Agnieszka Jagiello-Gruszfeld, Oleg Karyakin, François-Michel Delgado, Patrick Hurteloup, Eric Winquist, Nassim Morsli, Yacine Salhi, Stéphane Culine, and Hans von der Maase

J Clin Oncol 27:4454-4461. © 2009

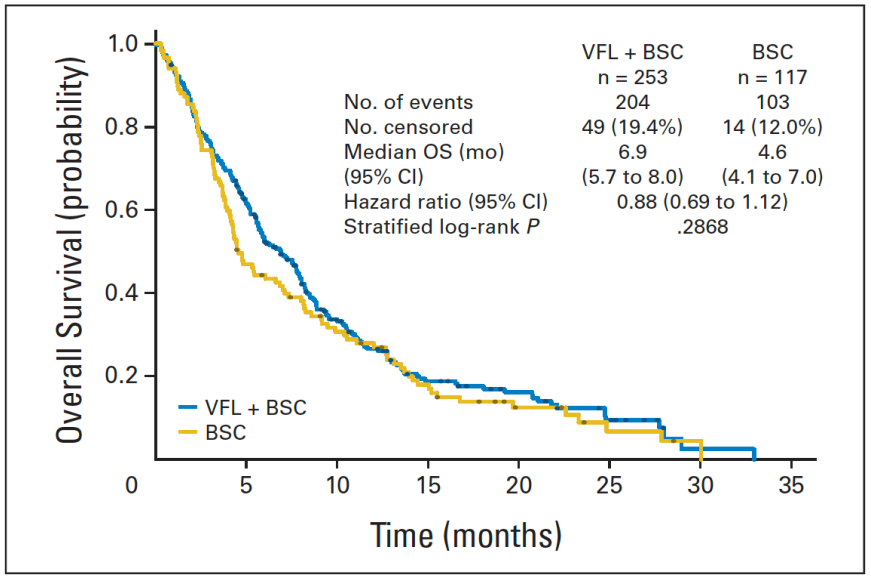


Fig 2. Overall survival (OS) in the intent-to-treat population (n = 370). VFL, vinflunine; BSC, best supportive care.

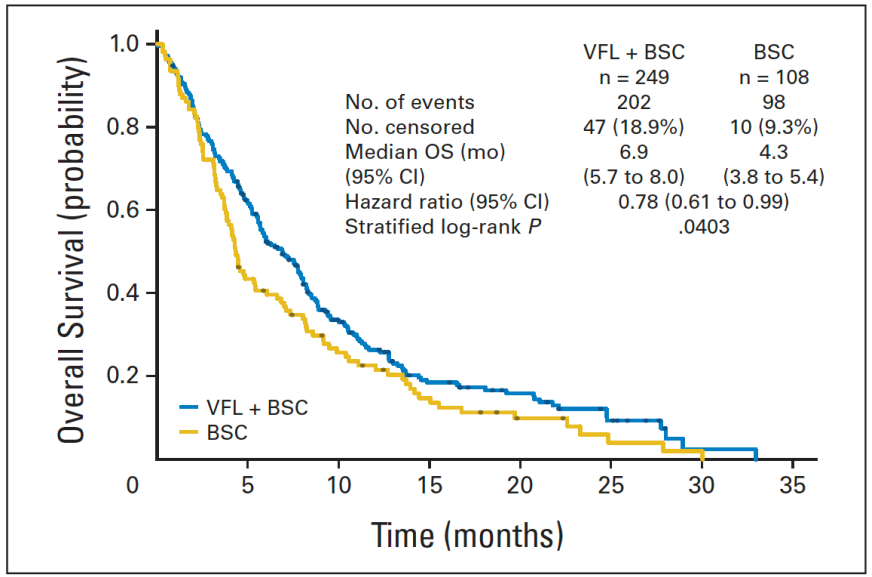
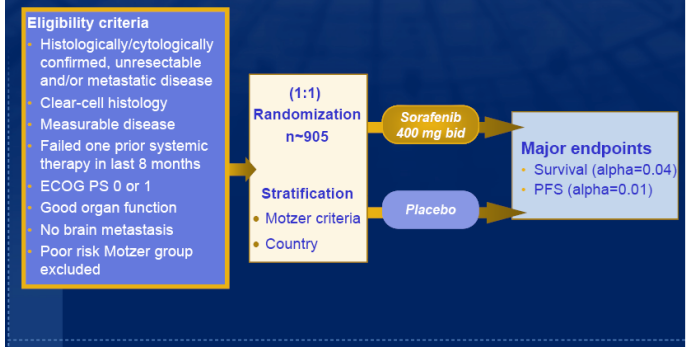


Fig 3. Overall survival (OS) in the eligible population (n = 357; 96.5% of intent-to-treat population). VFL, vinflunine; BSC, best supportive care.

Although the objective of the median 2-month survival advantage favoring VFL BSC versus BSC was achieved (6.9 v 4.6 months, respectively), this difference was not statistically significant (P .287; Fig 2).

In the eligible population (Fig 3), the objective of achieving a 2-month survival difference in OS between the VFLBSC and BSC arms was met (6.9 v 4.3 months, respectively), and this difference is statistically significant (P. 040).

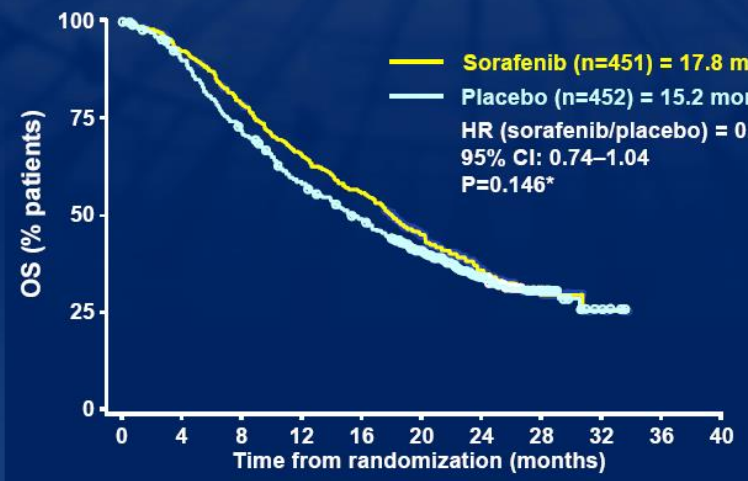
SORAFENIB improves PFS over placebo in 2nd line setting



Escudier et al, NEJM 2007

TARGET: Final OS Analysis

16 Months Post-Crossover: Intent-to-Treat



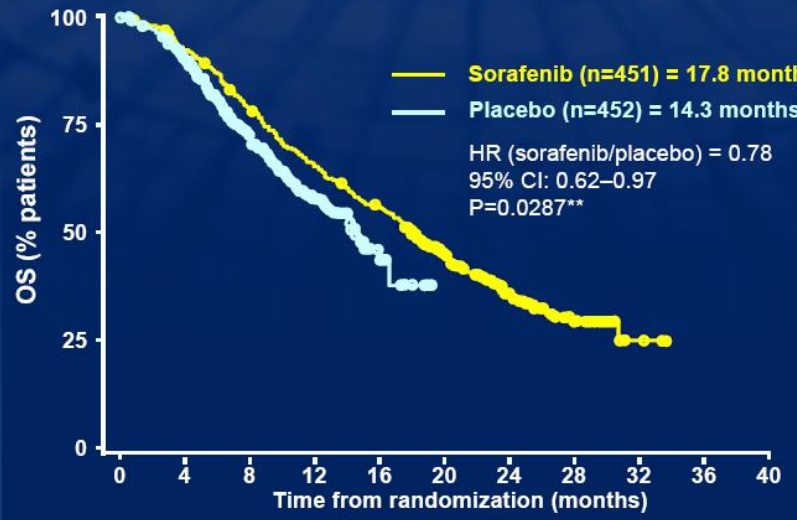
561 events

*Non-significant; O'Brien–Fleming threshold for statistical significance α=0.037

Bukowski et al, ASCO 2007

TARGET: Pre-planned Secondary Analysis

*OS Data for Placebo Patients Censored**



*Censored at 30 June 2005, approx. start of crossover

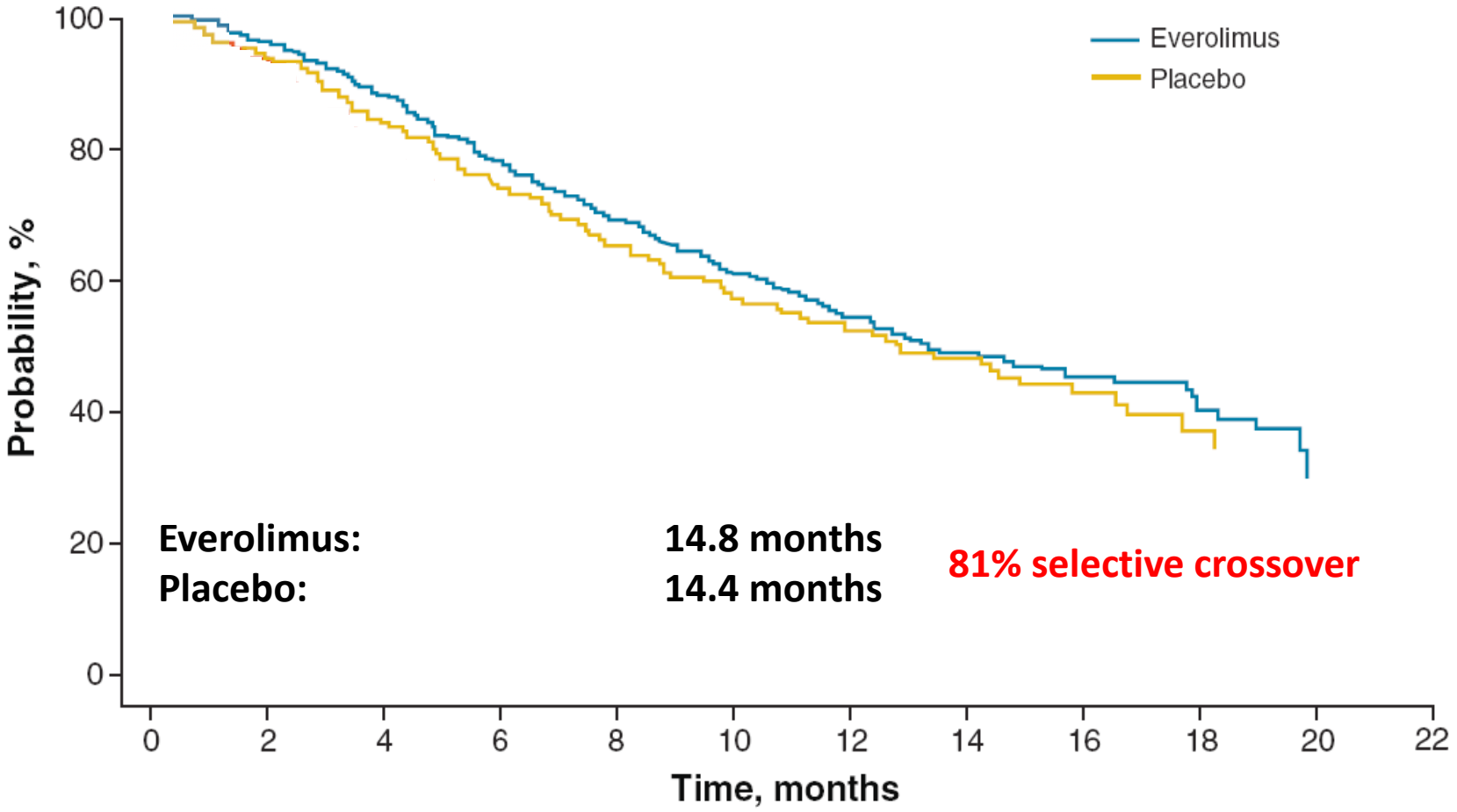
**Statistically significant: O'Brien–Fleming threshold for statistical significance α=0.037

Bukowski et al, ASCO 2007

Overall Survival of Metastatic Renal Cell Carcinoma Patients Corrected for Crossover Using Inverse Probability of Censoring Weights and Rank-Preserving Structural Failure Time Models: Two Analyses From the RECORD-1 trial

Pasi Korhonen,[1] Elisabetta Malangone,[2] Steven Sherman,[2] Roman Casciano,[2] Robert J. Motzer,[3] Jeanfrancois Baladi,[4] Tomas Haas,[5] Emmanuel Zuber,[5] Norbert Hollaender,[5] David E. Lebwohl[4]
[1]StatFinn Oy, Espoo, Finland; [2]Analytica International, New York, NY; [3]Memorial Sloan-Kettering Cancer Center, New York, NY; [4]Novartis Oncology, Florham Park, NJ; [5]Novartis Pharma AG, Basel, Switzerland

ITT analysis



Everolimus: **14.8 months**
Placebo: **14.4 months**

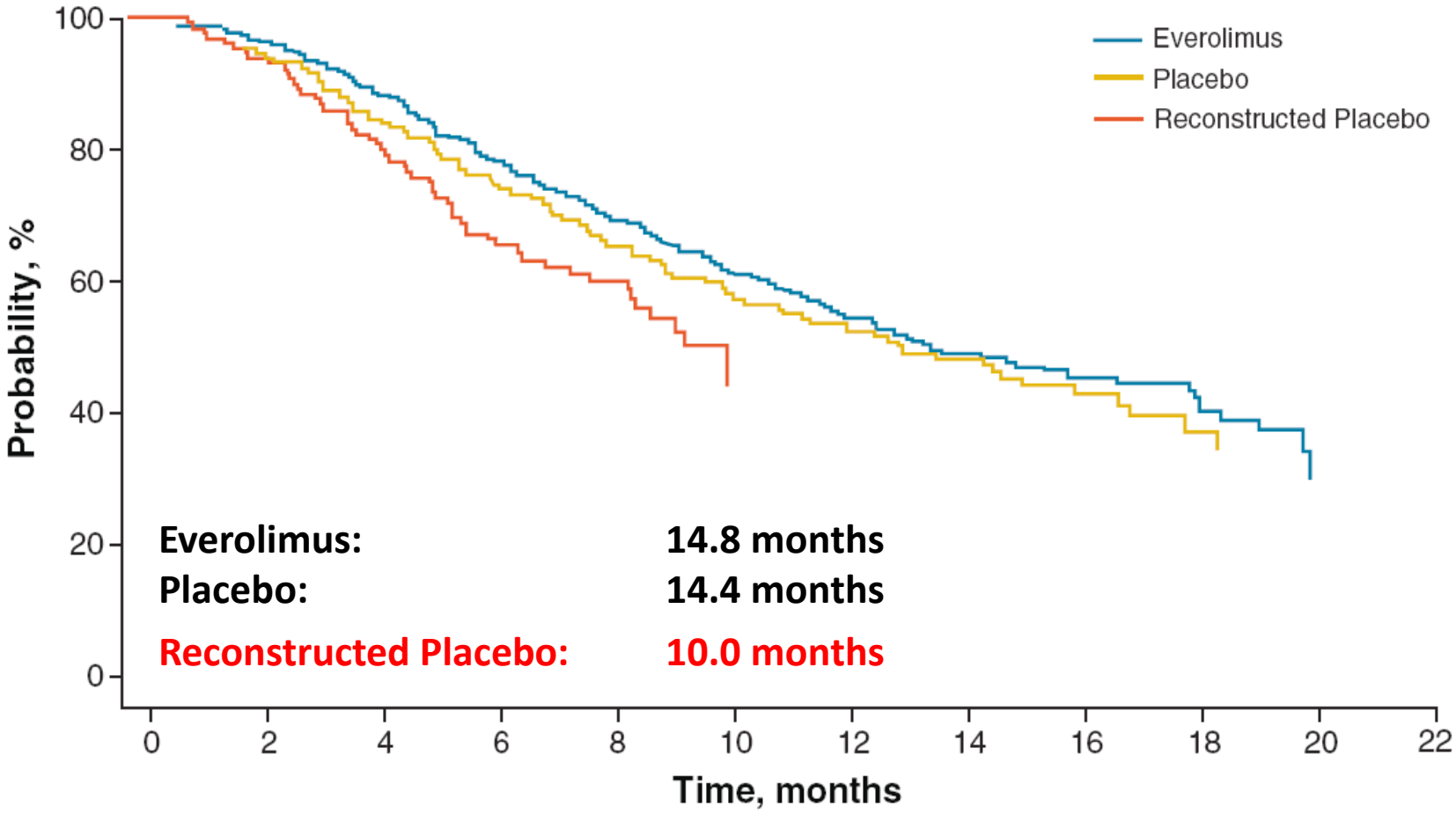
81% selective crossover

Presented at the 2010 ASCO Annual Meeting, June 4–8, 2010, Chicago, IL.

Overall Survival of Metastatic Renal Cell Carcinoma Patients Corrected for Crossover Using Inverse Probability of Censoring Weights and Rank-Preserving Structural Failure Time Models: Two Analyses From the RECORD-1 trial

Pasi Korhonen,[1] Elisabetta Malangone,[2] Steven Sherman,[2] Roman Casciano,[2] Robert J. Motzer,[3] Jeanfrancois Baladi,[4] Tomas Haas,[5] Emmanuel Zuber,[5] Norbert Hollaender,[5] David E. Lebwohl[4]
[1]StatFinn Oy, Espoo, Finland; [2]Analytica International, New York, NY; [3]Memorial Sloan-Kettering Cancer Center, New York, NY; [4]Novartis Oncology, Florham Park, NJ; [5]Novartis Pharma AG, Basel, Switzerland

RPSFT analysis



Presented at the 2010 ASCO Annual Meeting, June 4–8, 2010, Chicago, IL.

SOURCES OF BIAS IN CLINICAL TRIALS

Type of bias	Description	
Selection bias.	Systematic differences between baseline characteristics of the groups that are compared.	• Sequence generation. • Allocation concealment.
Performance bias.	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.	• Blinding of participants and personnel. • Other potential threats to validity.
Detection bias.	Systematic differences between groups in how outcomes are determined.	• Blinding of outcome assessment. • Other potential threats to validity.
Attrition bias.	Systematic differences between groups in withdrawals from a study.	• Incomplete outcome data
Reporting bias.	Systematic differences between reported and unreported findings.	• Selective outcome reporting

Cochrane Handbook for Systematic Reviews of Interventions

Version 5.1.0

If no patient blinding was performed…

… were they **unbiased** when filling the QoL questionnaire?

If no physician blinding was performed…

TREATMENT A

Not at risk of renal impairment

TREATMENT B

Renal impairment as common adverse event

Same frequency of creatinine testing?

SOURCES OF BIAS IN CLINICAL TRIALS

Type of bias	Description	
Selection bias.	Systematic differences between baseline characteristics of the groups that are compared.	• Sequence generation. • Allocation concealment.
Performance bias.	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.	• Blinding of participants and personnel. • Other potential threats to validity.
Detection bias.	Systematic differences between groups in how outcomes are determined.	• Blinding of outcome assessment. • Other potential threats to validity.
Attrition bias.	Systematic differences between groups in withdrawals from a study.	• Incomplete outcome data
Reporting bias.	Systematic differences between reported and unreported findings.	• Selective outcome reporting

Cochrane Handbook for Systematic Reviews of Interventions

Version 5.1.0

If no evaluator blinding was performed…

… was he (totally) **unbiased** when evaluating the scan?

SOURCES OF BIAS IN CLINICAL TRIALS

Type of bias	Description	
Selection bias.	Systematic differences between baseline characteristics of the groups that are compared.	• Sequence generation. • Allocation concealment.
Performance bias.	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.	• Blinding of participants and personnel. • Other potential threats to validity.
Detection bias.	Systematic differences between groups in how outcomes are determined.	• Blinding of outcome assessment. • Other potential threats to validity.
Attrition bias.	Systematic differences between groups in withdrawals from a study.	• Incomplete outcome data
Reporting bias.	Systematic differences between reported and unreported findings.	• Selective outcome reporting

Cochrane Handbook for Systematic Reviews of Interventions

Version 5.1.0

CONSORT

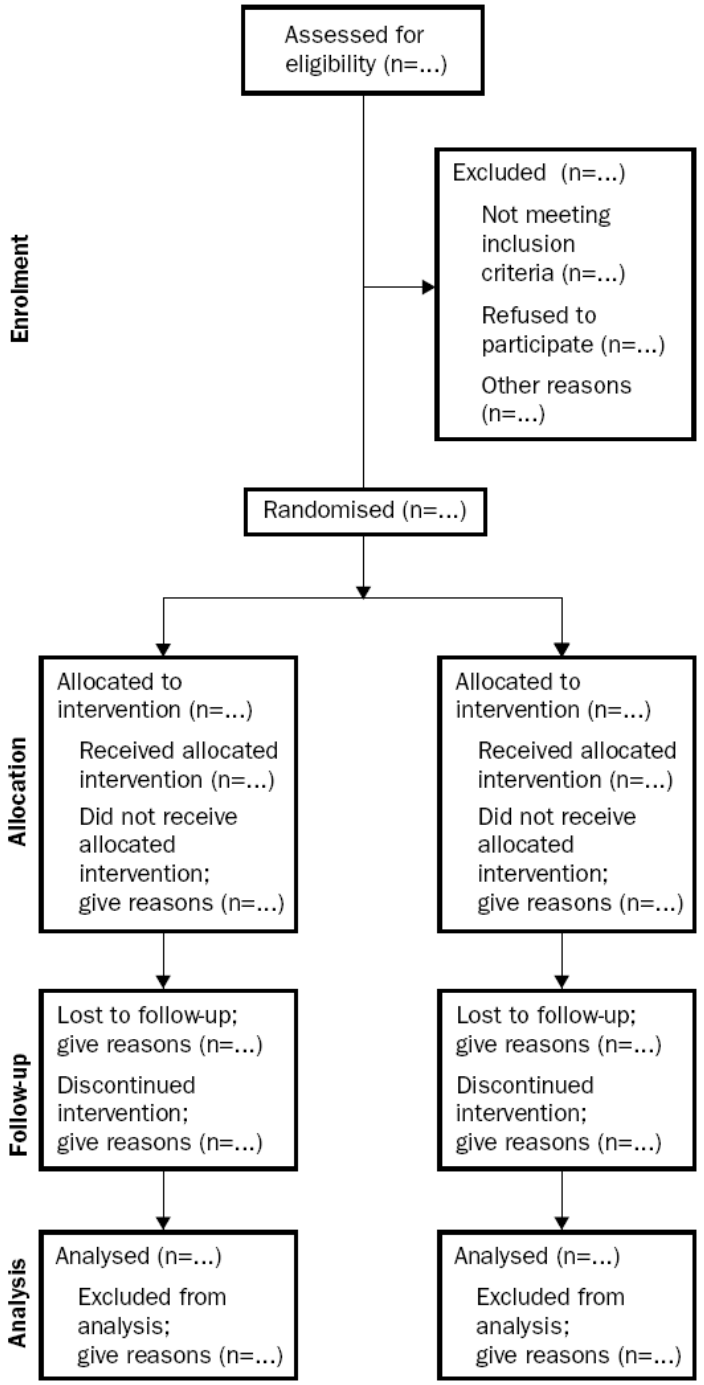
TRANSPARENT REPORTING of TRIALS

The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials

*David Moher, Kenneth F Schulz, Douglas G Altman, for the CONSORT Group**

Lancet 2001; **357:** 1191–94

A ciascuno studio è richiesto di dare conto del flusso di pazienti nelle fasi di arruolamento, assegnazione del trattamento, follow-up e analisi



Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews

Claire L Vale *senior research scientist*, Jayne F Tierney *senior research scientist* , Sarah Burdett *senior research scientist*

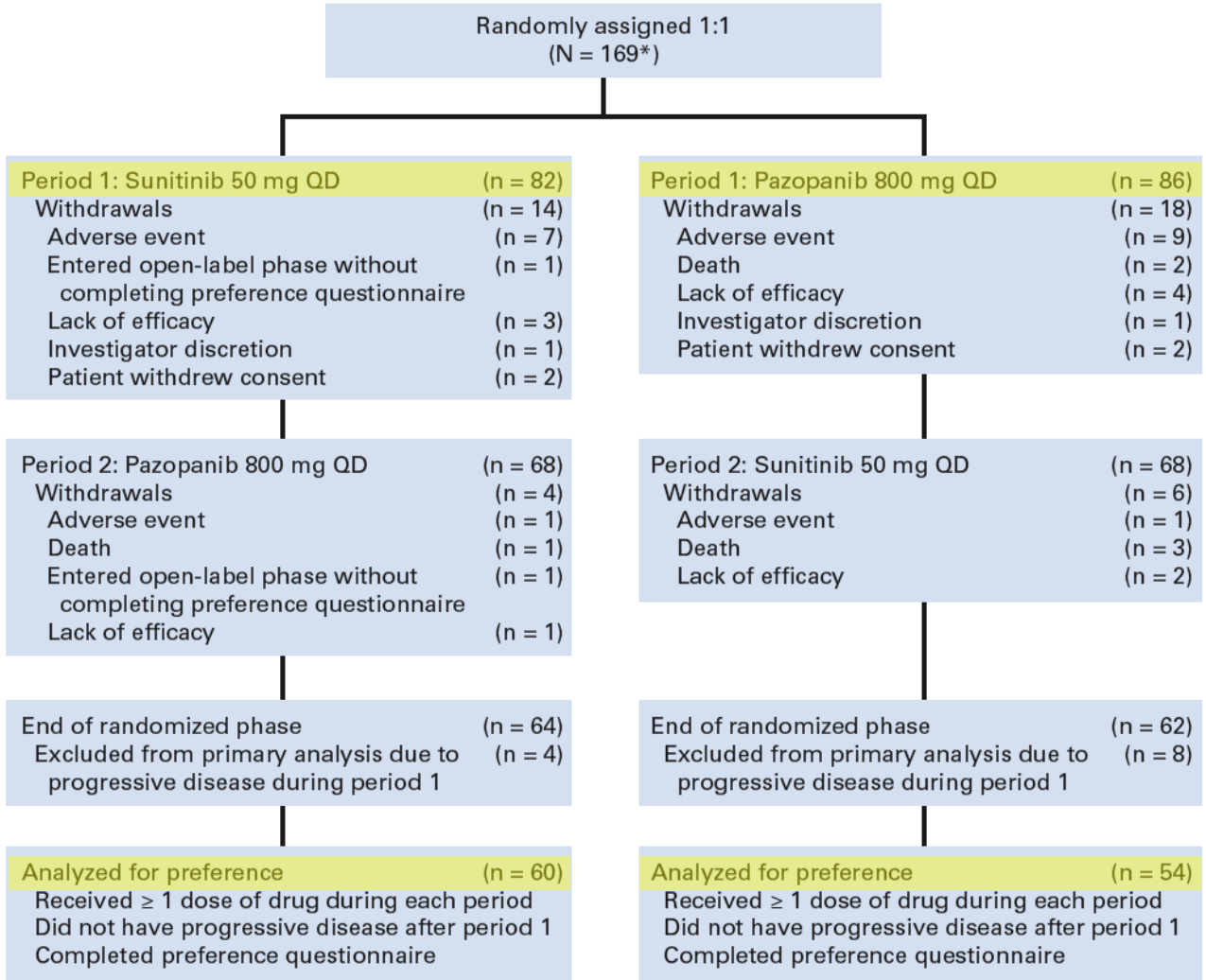
BMJ 2013;346:f1798 doi: 10.1136/bmj.f1798 (Published 22 April 2013)

To evaluate attrition bias, on the basis of whether the outcome data were incomplete or not, the authors had to establish a rule of thumb to ensure consistency between assessments. Trials were assessed as low risk of bias if less than 10% of patients were excluded overall and if similar proportions were excluded from both arms. Trials were judged as high risk of bias if there were considerable imbalances between arms or if more than 10% of randomised patients were excluded from the analysis.

Randomized, Controlled, Double-Blind, Cross-Over Trial Assessing Treatment Preference for Pazopanib Versus Sunitinib in Patients With Metastatic Renal Cell Carcinoma: PISCES Study

Bernard Escudier, Camillo Porta, Petri Bono, Thomas Powles, Tim Eisen, Cora N. Sternberg, Jürgen E. Gschwend, Ugo De Giorgi, Omi Parikh, Robert Hawkins, Emmanuel Sevin, Sylvie Négrier, Sadya Khan, Jose Diaz, Suman Redhu, Faisal Mehmud, and David Cella

J Clin Oncol 32.



Randomized, Controlled, Double-Blind, Cross-Over Trial Assessing Treatment Preference for Pazopanib Versus Sunitinib in Patients With Metastatic Renal Cell Carcinoma: PISCES Study

Bernard Escudier, Camillo Porta, Petri Bono, Thomas Powles, Tim Eisen, Cora N. Sternberg, Jürgen E. Gschwend, Ugo De Giorgi, Omi Parikh, Robert Hawkins, Emmanuel Sevin, Sylvie Négrier, Sadya Khan, Jose Diaz, Suman Redhu, Faisal Mehmud, and David Cella

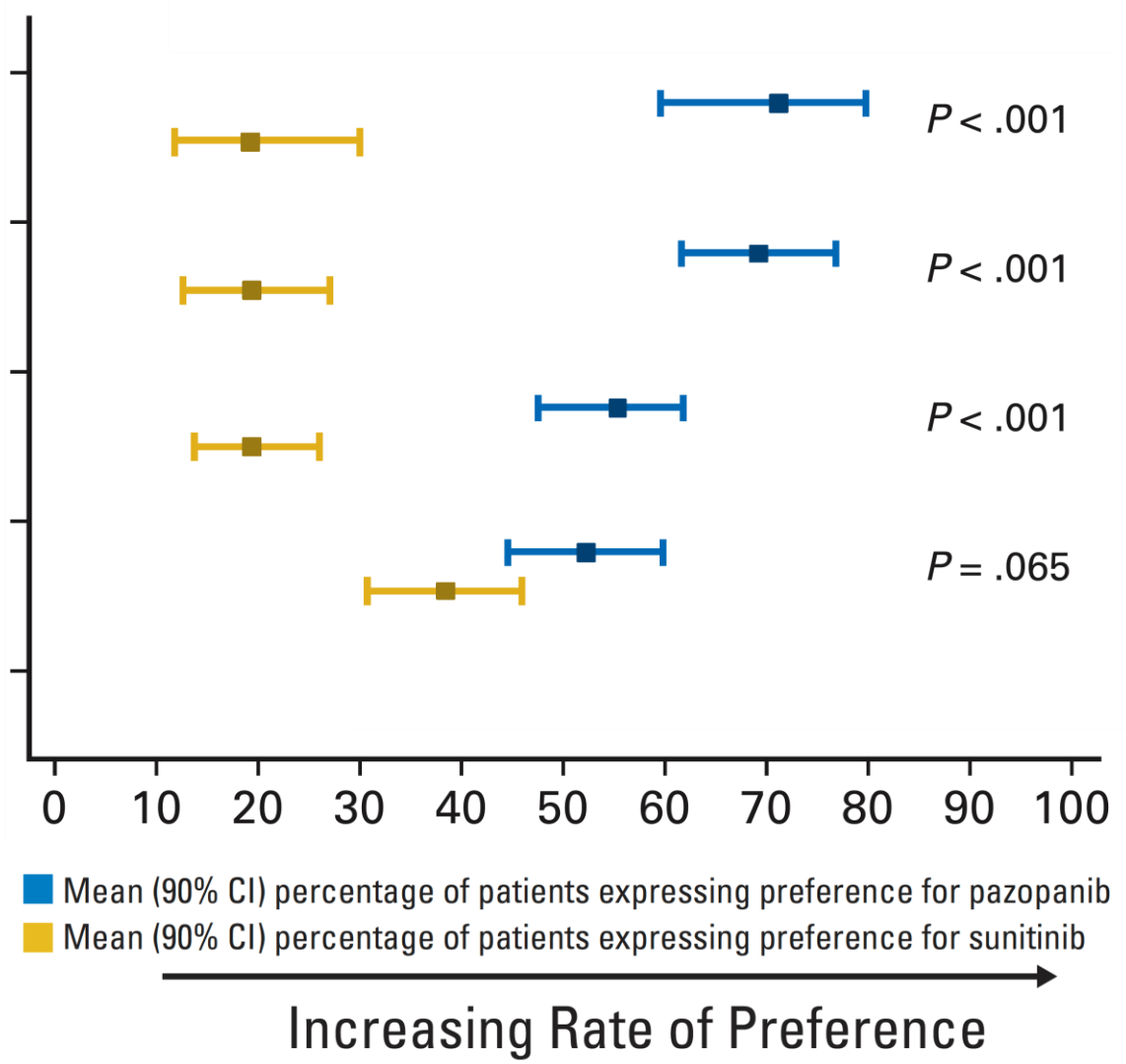
J Clin Oncol 32.

Primary end point (n = 114)

Patients who completed full study treatment (n = 80)

All questionnaires (n = 126)

Worst case: imputation of sunitinib preference for study withdrawals (n = 166)



SOURCES OF BIAS IN CLINICAL TRIALS

Type of bias	Description	
Selection bias.	Systematic differences between baseline characteristics of the groups that are compared.	• Sequence generation. • Allocation concealment.
Performance bias.	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.	• Blinding of participants and personnel. • Other potential threats to validity.
Detection bias.	Systematic differences between groups in how outcomes are determined.	• Blinding of outcome assessment. • Other potential threats to validity.
Attrition bias.	Systematic differences between groups in withdrawals from a study.	• Incomplete outcome data
Reporting bias.	Systematic differences between reported and unreported findings.	• Selective outcome reporting

Cochrane Handbook for Systematic Reviews of Interventions

Version 5.1.0

Instances of SRB

* Selective omission of outcomes from reports
* Selective choice of data for an outcome
* Selective reporting of analyses using the same data
* Selective reporting of subsets of the data
* Selective under-reporting of data

The Cochrane Handbook, 2011

A METHODS GROUP OF

THE COCHRANE COLLABORATION®

Quality of Life in Patients With Metastatic Renal Cell Carcinoma Treated With Sunitinib or Interferon Alfa: Results From a Phase III Randomized Trial

David Cella, Jim Z. Li, Joseph C. Cappelleri, Andrew Bushmakin, Claudie Charbonneau, Sindy T. Kim, Isan Chen, and Robert J. Motzer

J Clin Oncol 26:3763-3769. © 2008 by American Society of Clinical Oncology

Instruments	Difference in Least Squares Means		
FKSI-DRS (primary endpoint)	1.98*	MID = 2 points	✗
FKSI-15	3.27*	MID = 3 points	✓
FACT-G	5.58*	MID = 5 points	✓
PWB	1.42*	MID = 2 points	✗
SFWB	1.20*	MID = 2 points	✗
EWB	0.787*	MID = 2 points	✗
FWB	1.98*	MID = 2 points	✗
EQ-5D Index	0.0364*	MID = 0.09	✗
EQ-VAS	4.74*	MID = 8	✗

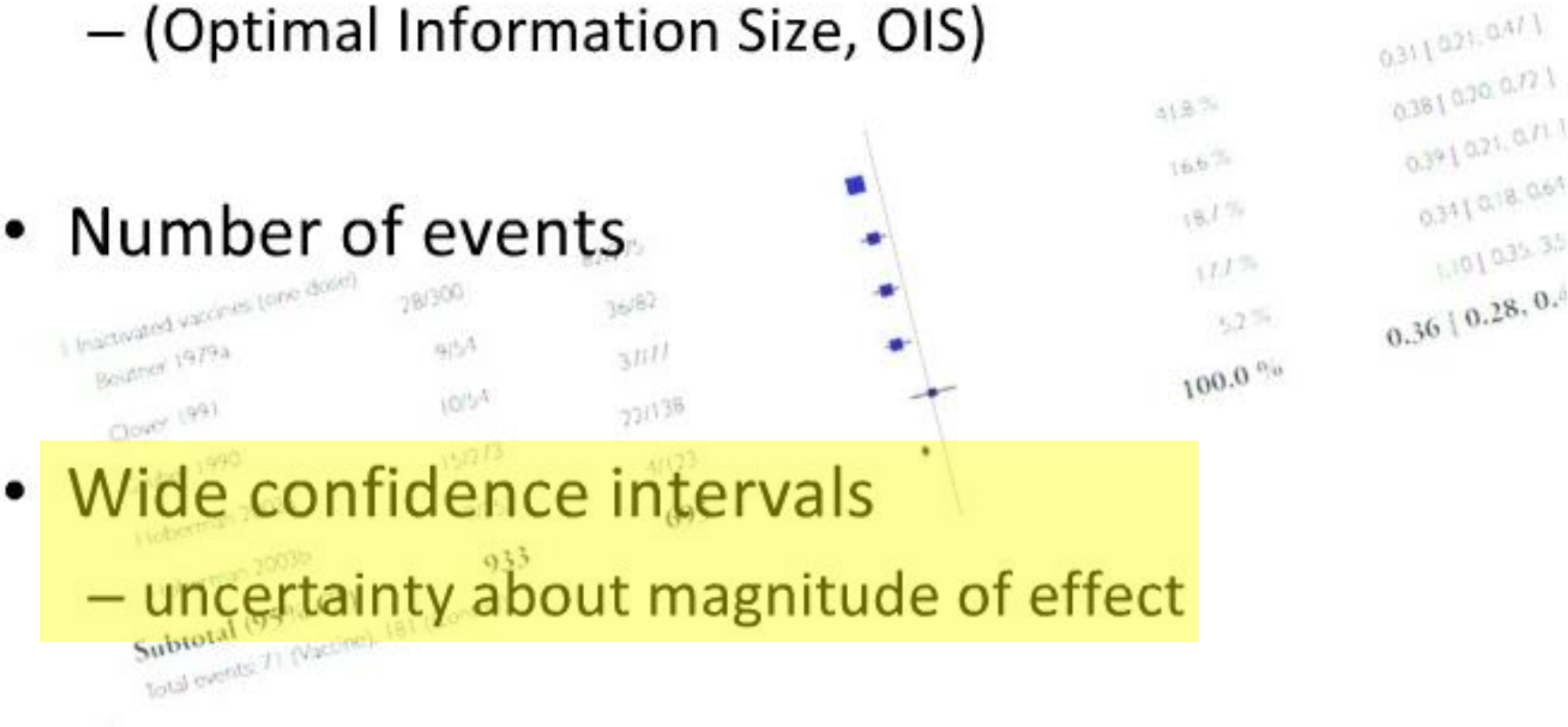
*Results were statistically significant at critical $P = .05$

Patients receiving sunitinib reported higher FKSI-15 and FACT-G scores at each cycle than those receiving IFN-α, per pre-established clinically meaningful thresholds.

When are results precise enough?

Consider

- Small sample size
  - (Optimal Information Size, OIS)
- Number of events
- Wide confidence intervals
  - uncertainty about magnitude of effect



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.

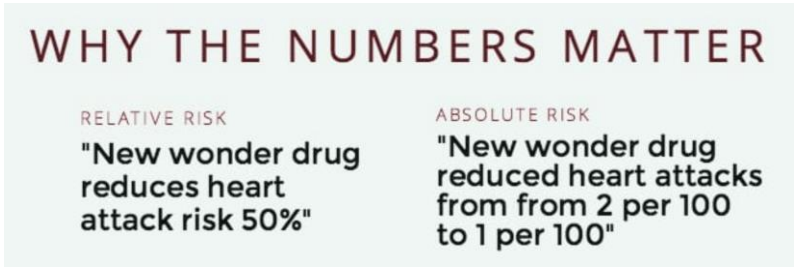
Journal of Clinical Epidemiology 64 (2011) 1283–1293

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.



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]

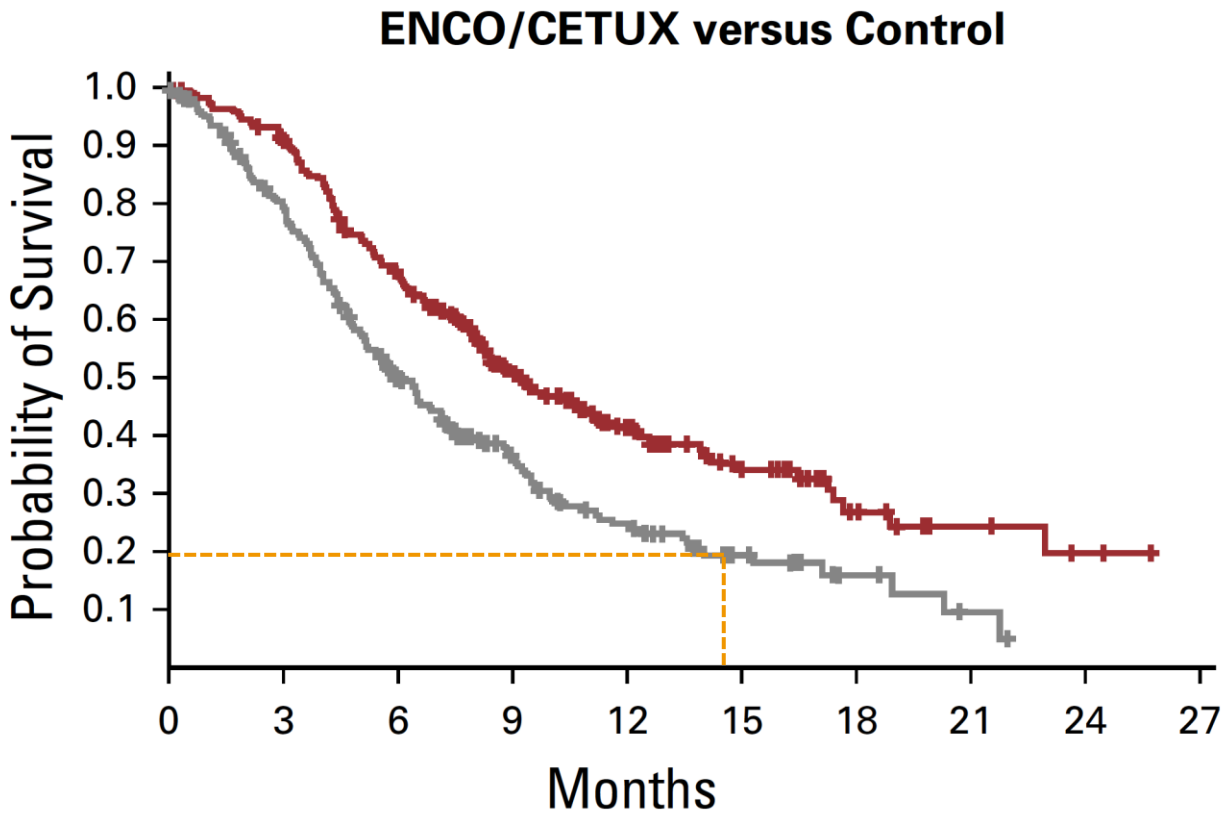
Key Points

- 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.

Encorafenib Plus Cetuximab as a New Standard of Care for Previously Treated *BRAF* V600E–Mutant Metastatic Colorectal Cancer: Updated Survival Results and Subgroup Analyses from the BEACON Study

Josep Tabernero, MD, PhD[1]; Axel Grothey, MD[2]; Eric Van Cutsem, MD, PhD[3]; Rona Yaeger, MD[4]; Harpreet Wasan, MD[5]; Takayuki Yoshino, MD, PhD[6]; Jayesh Desai, MBBS[7]; Fortunato Ciardiello, MD, PhD[8]; Fotios Loupakis, MD, PhD[9]; Yong Sang Hong, MD, PhD[10]; Neeltje Steeghs, MD, PhD[11]; Tormod Kyrre Guren, MD, PhD[12]; Hendrik-Tobias Arkenau, MD, PhD[13]; Pilar García-Alfonso, MD[14]; Elena Elez, MD, PhD[1]; Ashwin Gollerkeri, MD[15]; Kati Maharry, PhD[15]; Janna Christy-Bittel, MSN[15]; and Scott Kopetz, MD, PhD[16]

J Clin Oncol 39:273-284.



- Median f.u.: 14.7 months
- Baseline risk at median f.u.: 80%
- **Risk Difference**: 17 events lower / 100 pts (**95%CI: 26 lower to 9 lower**)
- NNT ~6

LC95% consistenti sia con una rilevante efficacia (↓26 eventi decesso / 100 paz.) sia con un beneficio marginale (↓9 eventi decesso / 100 paz.)

NON seria imprecisione

NON Imprecisione Clinica anche in presenza di valore di P ≥0.05

M.I.D. EORTC QLQ-C30 GHS: **10 punti**

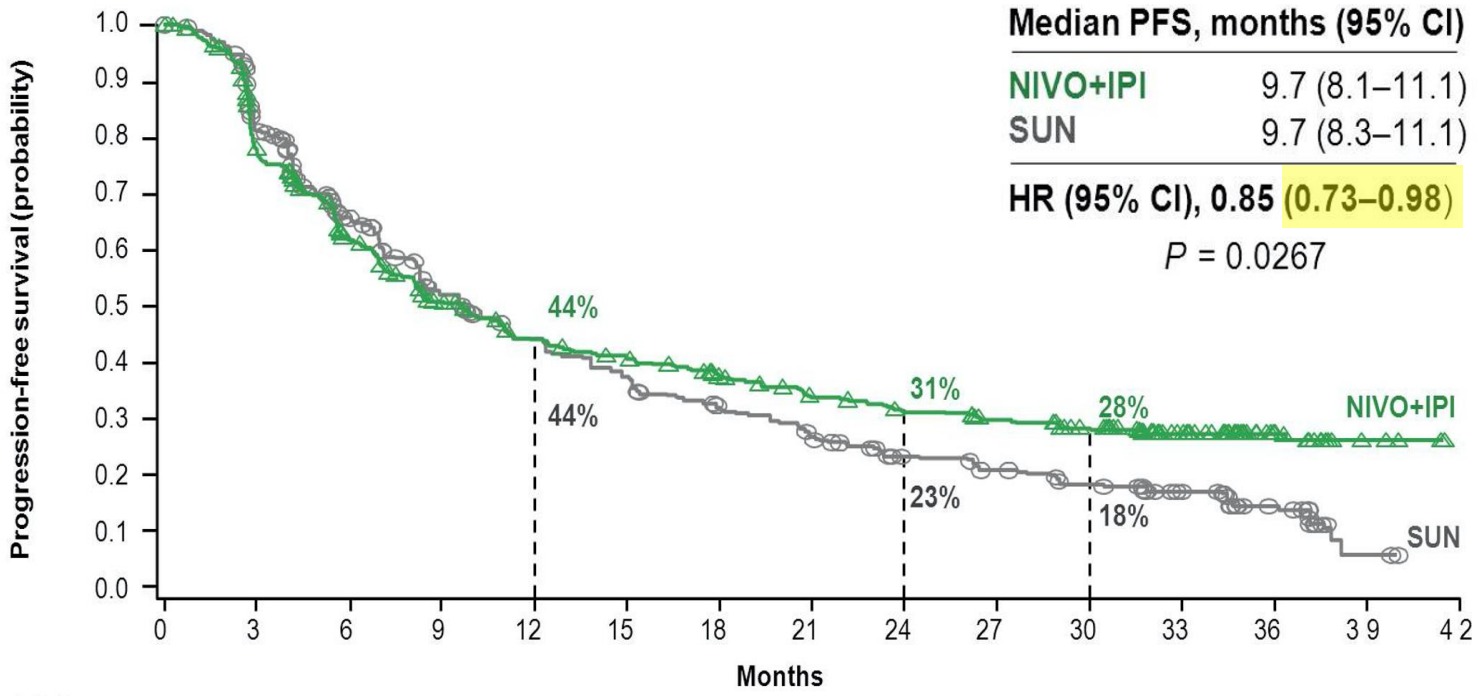
Scale	Cabazitaxel	Lu-PSMA	Diff.
	Pred. Mean (SE) {95% CI}	Pred. Mean (SE) {95% CI}	Pred. Mean (SE) {95% CI} [p-value]
Global health status / QoL	60.4 (1.8) {56.9 to 63.9}	63.4 (1.6) {60.3 to 66.5}	3.0 (2.3) {-1.6 to 7.5} [0.202]

IC95% dell'effetto assoluto compreso nel *range* di non rilevanza clinica (<10 punti)

NON Imprecisione

Lancet 2021; 397: 797–804

Investigator-Assessed Progression-Free Survival per RECIST v1.1: ITT Patients



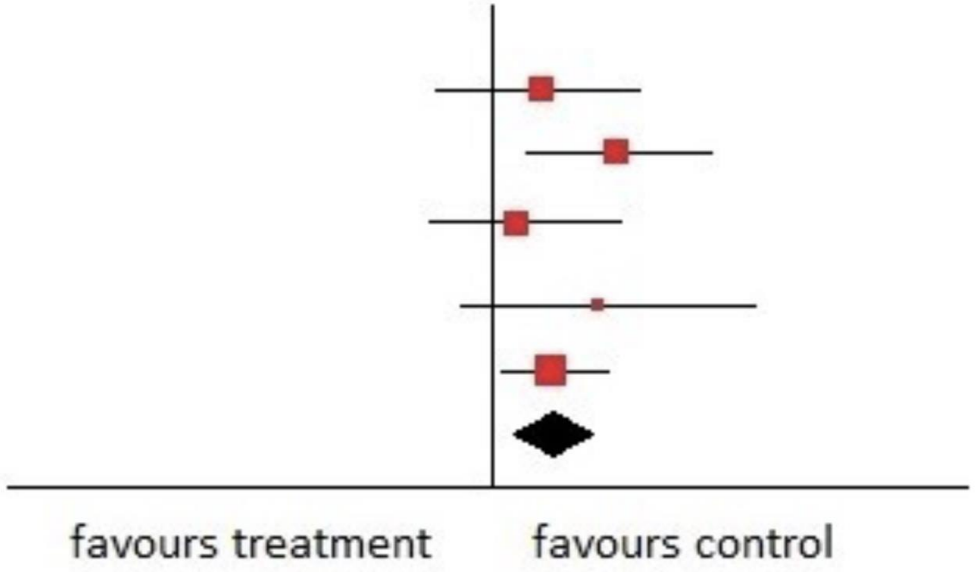
Presented By Nizar Tannir at 2019 Genitourinary Cancers Symposium

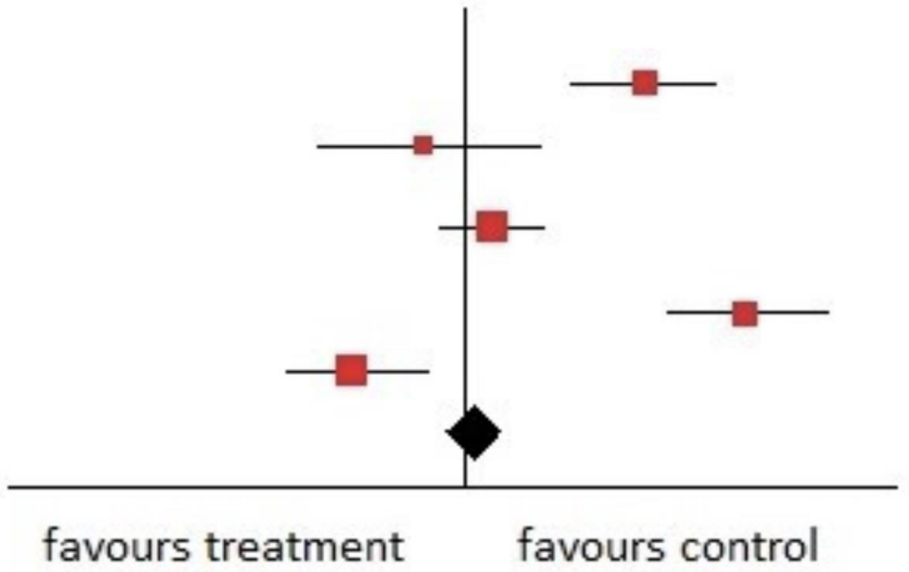
Risk Difference:
5 events lower / 100 pts
(**95%CI: 11 lower to 1 lower**)

LC95% consistenti con una moderata efficacia (↓11 eventi progressione / 100 paz.) e con assenza di beneficio (↓1 evento progressione / 100 paz.)

SERIA imprecisione

Eterogeneità





Eterogeneità delle stime di effetto tra gli studi che *non trova spiegazione logica*

(diversità nel tipo di intervento o nella composizione delle popolazioni studiate)

What is Heterogeneity?

- **Any kind of variability among studies in a systematic review may be termed heterogeneity.**
- **I-squared (I^2)**
  - ✓ **describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance)**.
  - ✓ thresholds for the interpretation of I^2:
    - 0% to 40%: might not be important;
    - 30% to 60%: may represent moderate heterogeneity;
    - **50% to 90%: may represent substantial heterogeneity;**
    - **75% to 100%: considerable heterogeneity.**

[Punti conquistati] for [campionati 1995/96 - 2005/06]

08-Apr-2008

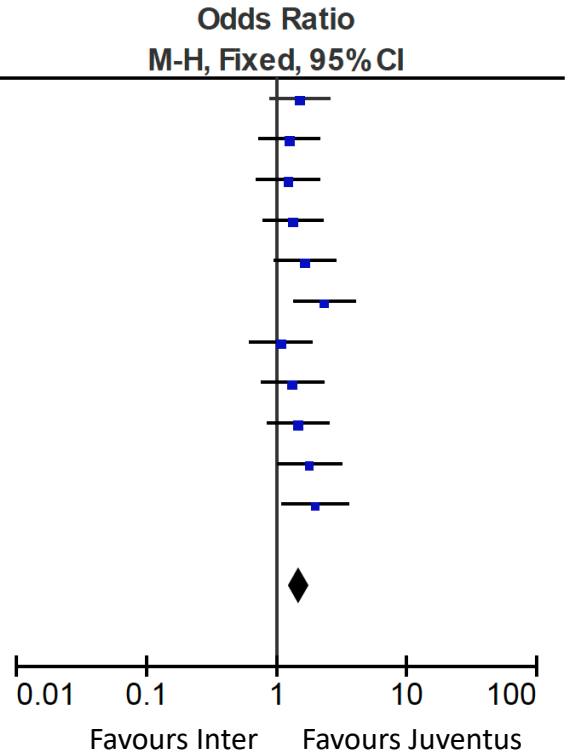
1 Juventus vs Inter

1.1 Punti conquistati / Punti teorici

	Juventus		Inter			Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI
1995/96	65	108	54	108	9.5%	1.51 [0.88, 2.59]
1996/97	65	108	59	108	10.4%	1.26 [0.73, 2.15]
1997/98	74	108	69	108	9.6%	1.23 [0.70, 2.16]
1998/99	54	108	46	108	10.2%	1.35 [0.79, 2.30]
1999/00	71	108	58	108	8.8%	1.65 [0.96, 2.86]
2000/01	73	108	51	108	7.3%	2.33 [1.34, 4.05]
2001/02	71	108	69	108	10.5%	1.08 [0.62, 1.90]
2002/03	72	108	65	108	9.6%	1.32 [0.76, 2.31]
2003/04	69	108	59	108	9.4%	1.47 [0.85, 2.54]
2004/05	86	114	72	114	7.8%	1.79 [1.01, 3.17]
2005/06	91	114	76	114	6.8%	1.98 [1.08, 3.61]
Total (95% CI)		**1200**		**1200**	**100.0%**	**1.50 [1.27, 1.77]**
Total events	791		678			

Heterogeneity: Chi² = 6.30, df = 10 (P = 0.79); I² = 0%

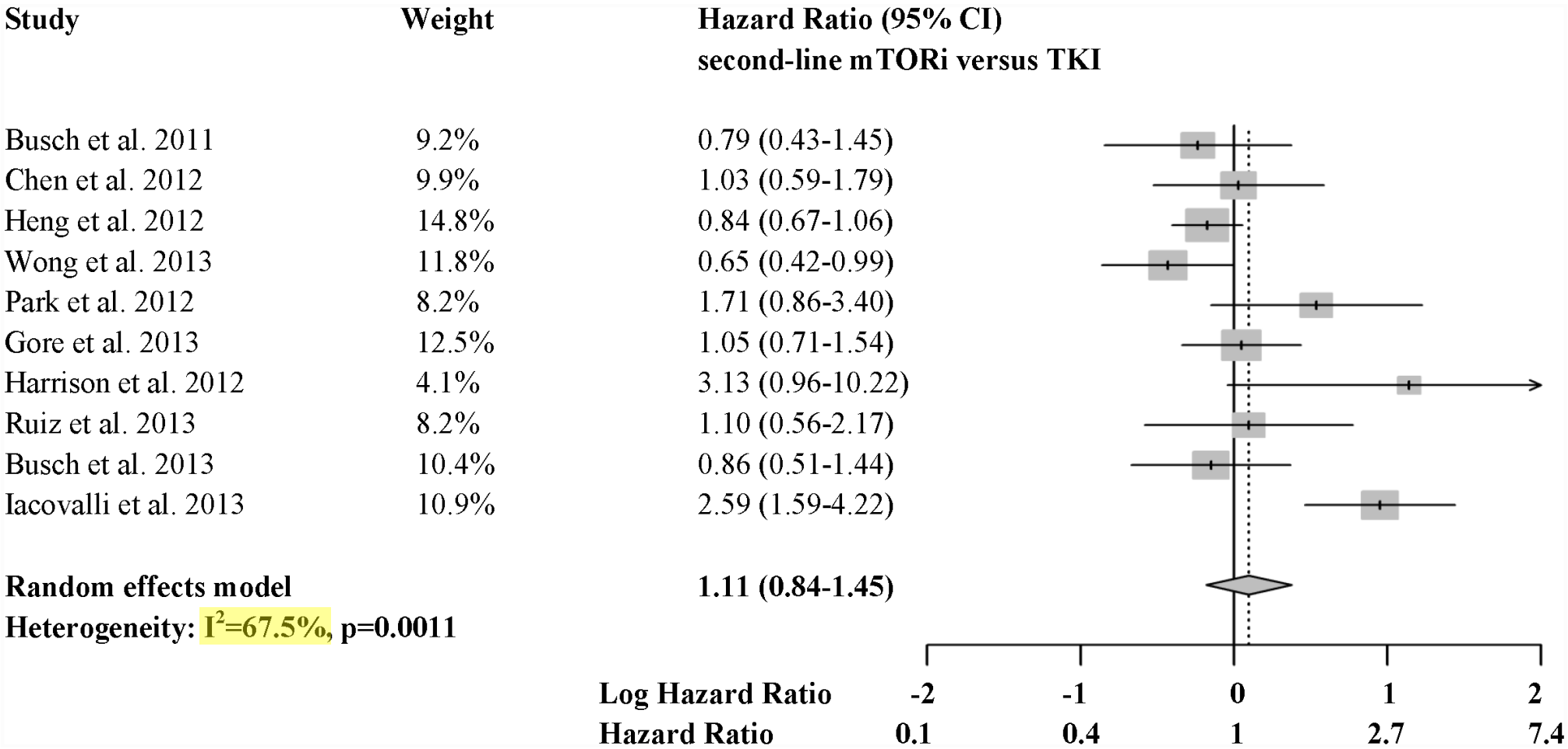
Test for overall effect: Z = 4.77 (P < 0.00001)



Comparative Effectiveness of Second-Line Targeted Therapies for Metastatic Renal Cell Carcinoma: A Systematic Review and Meta-Analysis of Real-World Observational Studies

Daniel Y. Heng[1], James Signorovitch[2], Elyse Swallow[2], Nanxin Li[2], Yichen Zhong[2], Paige Qin[2], Daisy Y. Zhuo[2], Xufang Wang[3], Jinhee Park[3], Sotirios Stergiopoulos[3], Christian Kollmannsberger[4]*

PLOS ONE | DOI:10.1371/journal.pone.0114264 December 10, 2014



ANALISI PER SOTTOGRUPPI

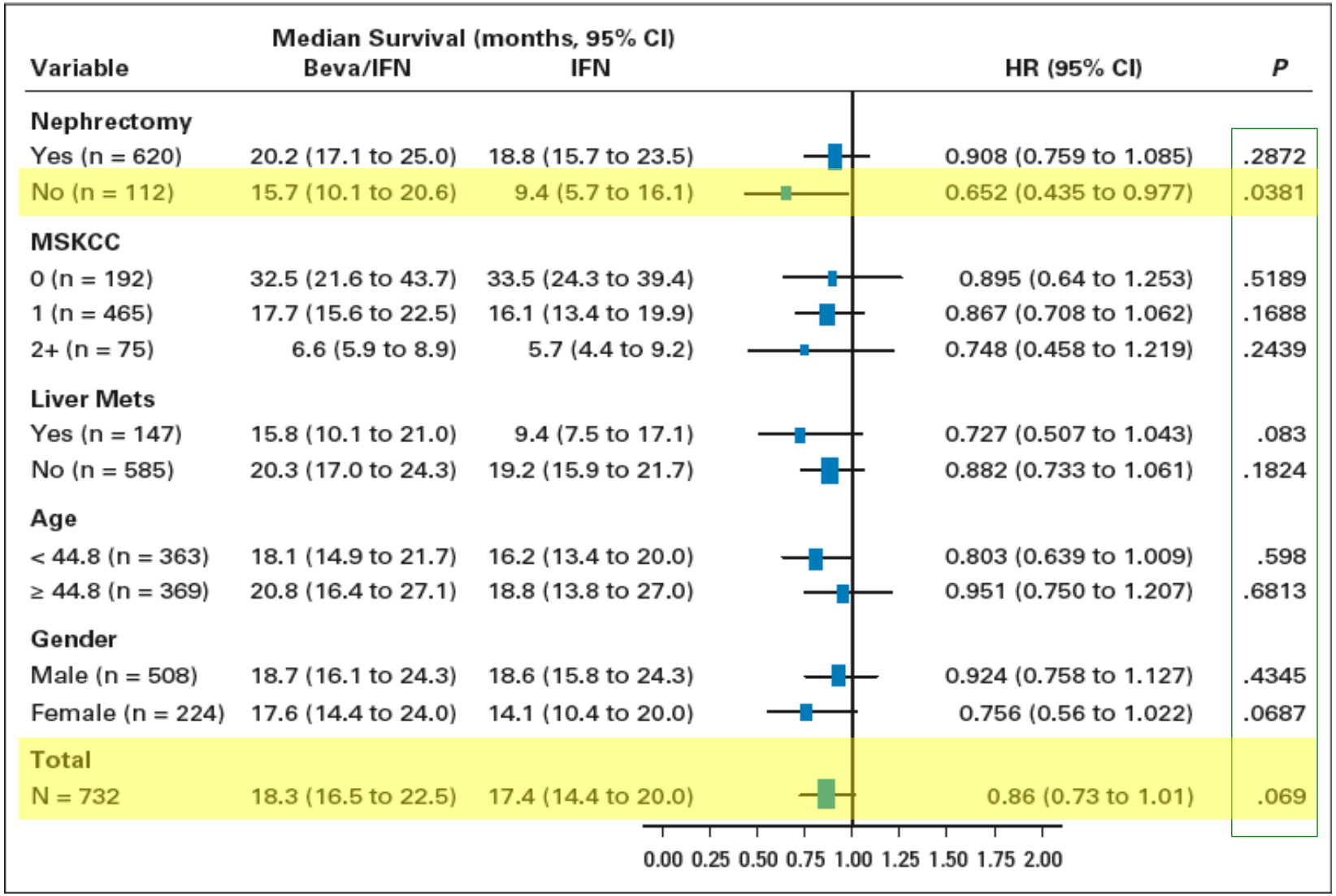
- Il campione originale viene suddiviso in vari strati (**sottogruppi**) contraddistinti da caratteristiche peculiari.
- Tipologia di analisi:
  - ***Post hoc*** (analisi retrospettiva)
  - ***Pre-specified (preannunciata)*** (prevista dal protocollo di studio; criterio di stratificazione?)
  - ***Pre-planned (prepianificata)*** (prevista dal protocollo di studio con piano di analisi specifico)

- L'approccio alla lettura
- Il quesito clinico
- L'obiettivo dello studio
- Il disegno dello studio: studi sperimentali
- Randomizzazione e mascheramento
- Il disegno dello studio: studi osservazionali
- Indicatori di effetto
- Endpoints per le diverse Fasi della sperimentazione
- Il Forest Plot
- Interpretazione statistica dei risultati dello studio
- Interpretazione clinica dei risultati dello studio
- Trasferibilità dei risultati dello studio
- Affidabilità dei risultati dello studio
- **Analisi per sottogruppi**

Phase III Trial of Bevacizumab Plus Interferon Alfa Versus Interferon Alfa Monotherapy in Patients With Metastatic Renal Cell Carcinoma: Final Results of CALGB 90206

Brian I. Rini, Susan Halabi, Jonathan E. Rosenberg, Walter M. Stadler, Daniel A. Vaena, Laura Archer, James N. Atkins, Joel Picus, Piotr Czaykowski, Janice Dutcher, and Eric J. Small

J Clin Oncol 28:2137-2143. © 2010 by American Society of Clinical Oncology



ANALISI PER SOTTOGRUPPI

- Il campione originale viene suddiviso in vari strati (**sottogruppi**) contraddistinti da caratteristiche peculiari.
- Tipologia di analisi:
  - ***Post hoc*** (analisi retrospettiva)
  - ***Pre-specified (preannunciata)*** (prevista dal protocollo di studio; criterio di stratificazione?)
  - ***Pre-planned (prepianificata)*** (prevista dal protocollo di studio con piano di analisi specifico)

ASCO Daily News

Prior results of JAVELIN Renal 101 that compared avelumab plus axitinib versus sunitinib showed that the combination significantly improved median PFS and doubled the ORR.

In an update of these data, preplanned subgroup analyses revealed that the PFS and ORR advantages observed with avelumab and axitinib in the overall population also apply to a wide variety of patient subgroups focused on IMDC and Memorial Sloan Kettering Cancer Center prognostic risk groups, PD-L1 status, prior nephrectomy, smoking status, and body mass index (Abstract 544).

Journal of Clinical Oncology®
An American Society of Clinical Oncology Journal

544

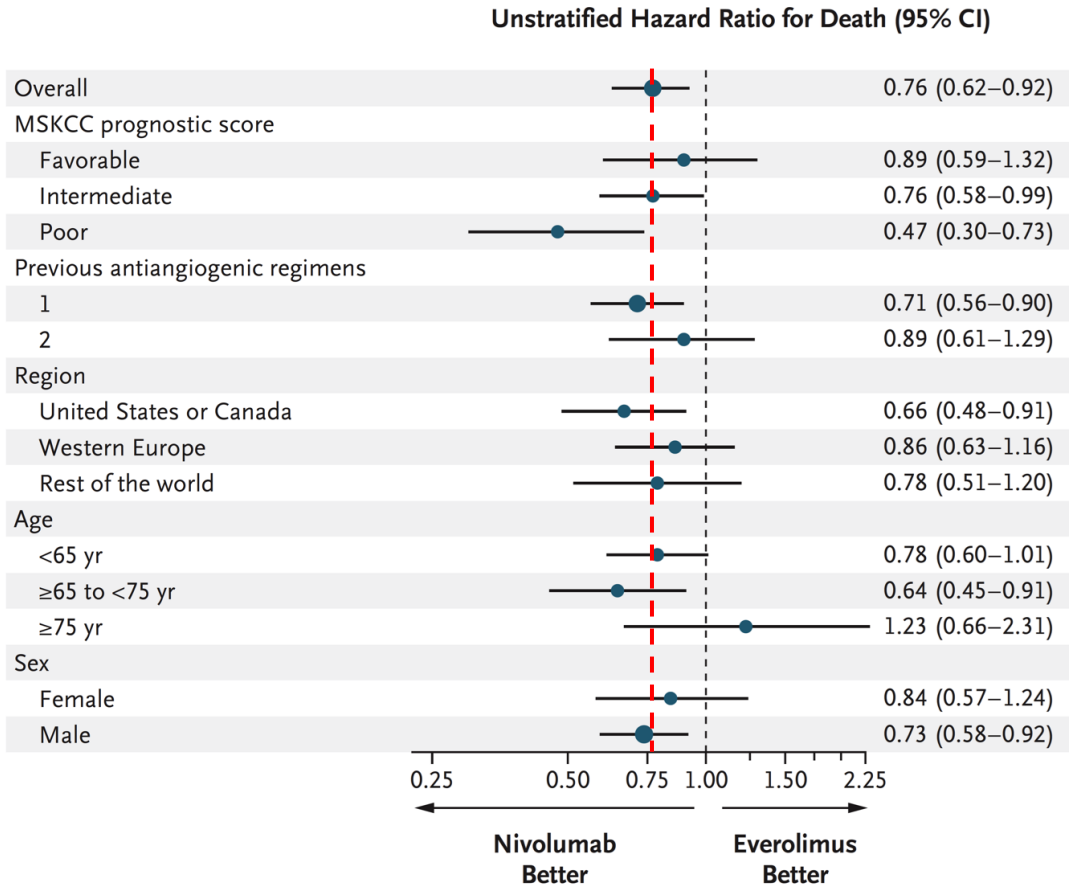
Background: In the ongoing phase 3 JAVELIN Renal 101 trial, progression-free survival (PFS) was longer (median, 13.8 vs 8.4 mo; hazard ratio, 0.69; p= 0.0001) and the objective response rate (ORR) was higher (51% vs 26%) with A + Ax vs S in patients with previously untreated aRCC. Here we report outcomes from an analysis of several prespecified subgroups.

Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma

R.J. Motzer, B. Escudier, D.F. McDermott, S. George, H.J. Hammers, S. Srinivas, S.S. Tykodi, J.A. Sosman, G. Procopio, E.R. Plimack, D. Castellano, T.K. Choueiri, H. Gurney, F. Donskov, P. Bono, J. Wagstaff, T.C. Gauler, T. Ueda, Y. Tomita, F.A. Schutz, C. Kollmannsberger, J. Larkin, A. Ravaud, J.S. Simon, L.-A. Xu, I.M. Waxman, and P. Sharma, for the CheckMate 025 Investigators*

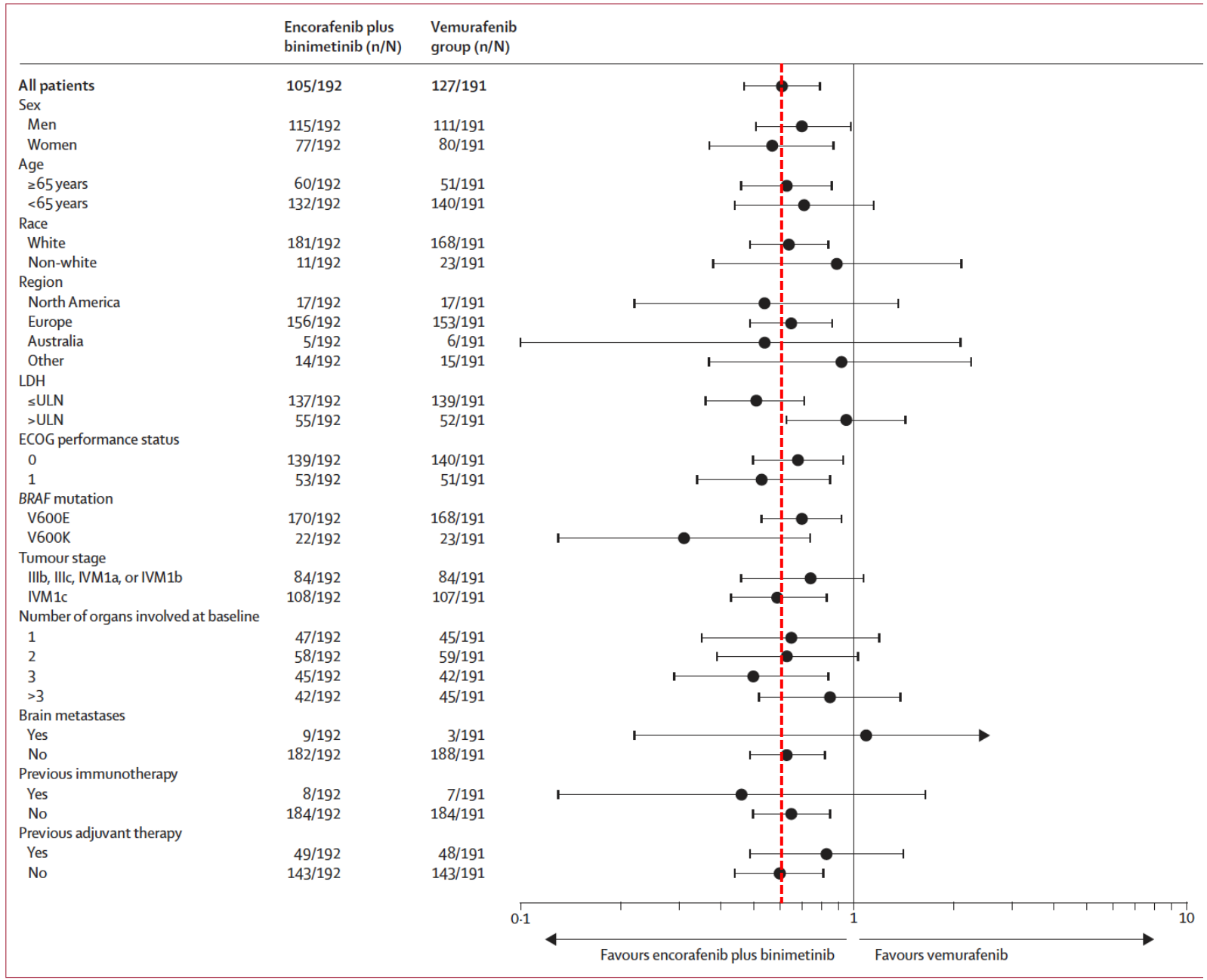
N Engl J Med 2015;373:1803-13.

Subgroup Analyses of Overall Survival

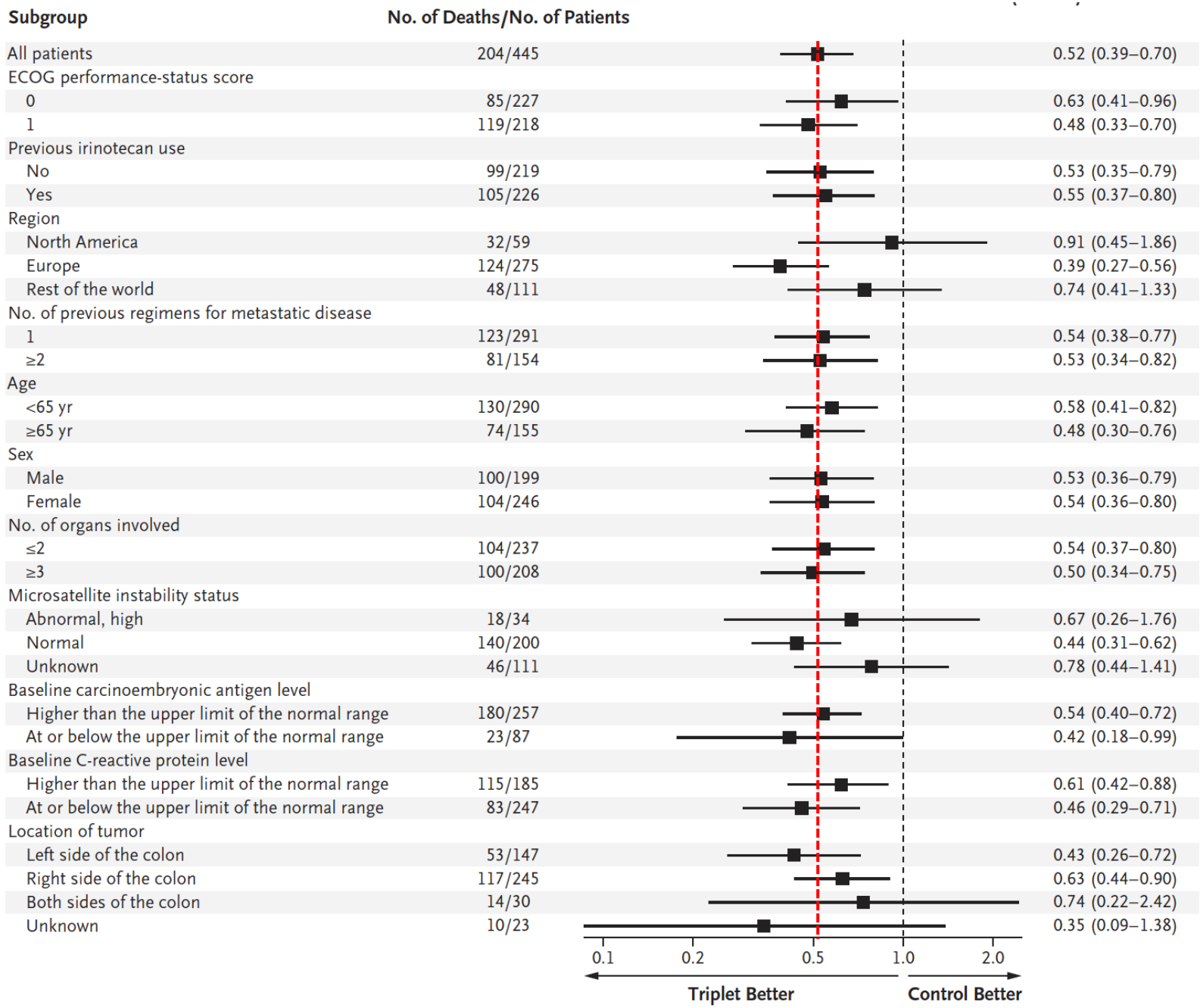


Il riferimento non è (più) la linea di non-effetto, bensì la linea tracciata in corrispondenza dell'effetto osservato sull'intero campione…

… un trattamento non dovrebbe essere limitato / escluso in una specifica sottopopolazione (solo) sulla base di un LC95% al di là (al di qua) della linea di non-effetto!



Lancet Oncol 2018; 19: 1315–27



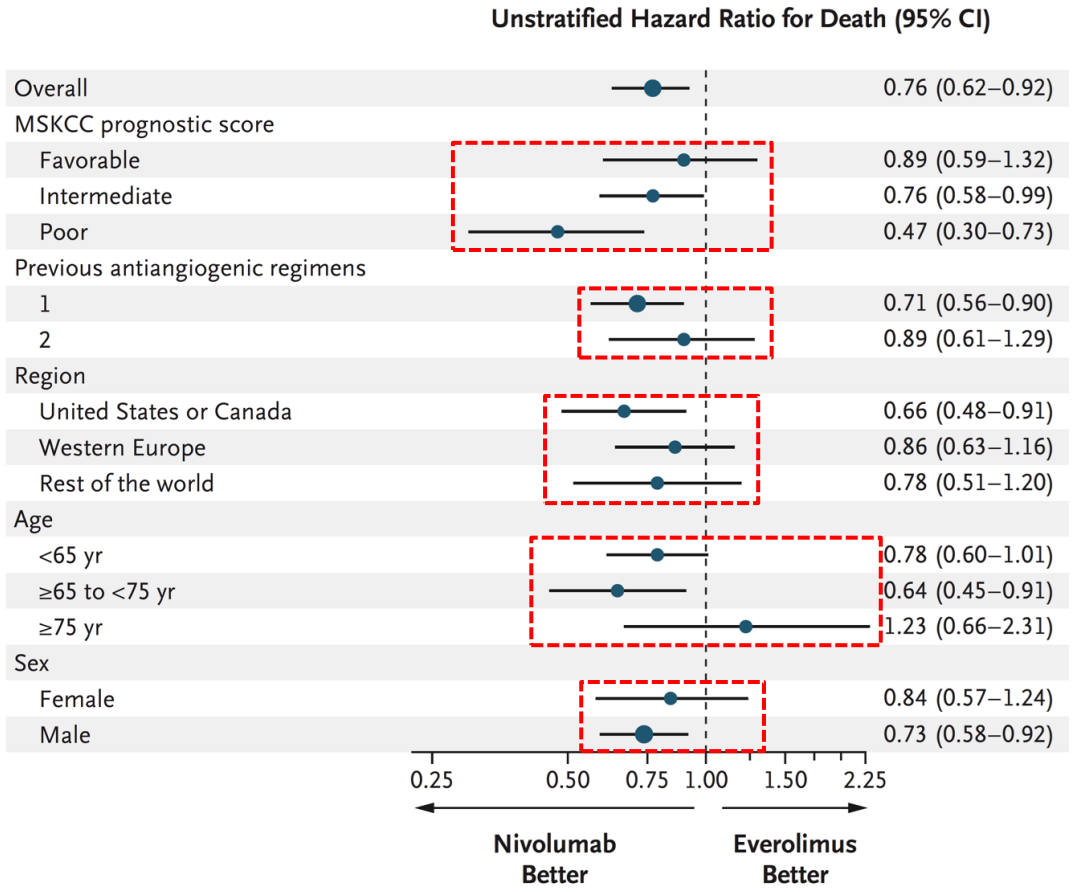
N Engl J Med 2019;381:1632-43.

Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma

R.J. Motzer, B. Escudier, D.F. McDermott, S. George, H.J. Hammers, S. Srinivas, S.S. Tykodi, J.A. Sosman, G. Procopio, E.R. Plimack, D. Castellano, T.K. Choueiri, H. Gurney, F. Donskov, P. Bono, J. Wagstaff, T.C. Gauler, T. Ueda, Y. Tomita, F.A. Schutz, C. Kollmannsberger, J. Larkin, A. Ravaud, J.S. Simon, L.-A. Xu, I.M. Waxman, and P. Sharma, for the CheckMate 025 Investigators*

N Engl J Med 2015;373:1803-13.

Subgroup Analyses of Overall Survival



The heterogeneity of the treatment effect within each subgroup shown in Figure 2A was tested with the use of an interaction test in a Cox proportional-hazards model with treatment, subgroup, and treatment-by-subgroup interaction as covariates. None of the interaction terms were significant at the 0.05 level.

Il test di interazione …

- Se significativo, indica che "*l'efficacia relativa del trattamento è significativamente* ***diversa*** *tra i sottogruppi considerati*"
- **Non indica** se quanto osservato in ciascuno dei sottogruppi è statisticamente significativo
- E' solo la maniera migliore per dimostrare che c'è una **ipotesi** che vale la pena di verificare prospetticamente

F. Perrone, AIOM 2008

Potenziali problemi…

- **High Risk of Bias** (sottogruppo non oggetto di stratificazione)
- **Multiplicity** (*type I error inflation* dovuta a confronti ripetuti)
- **Imprecision** (LC95% compatibili con interpretazioni cliniche di segno opposto)

Only one thing is worse than doing subgroup analyses…

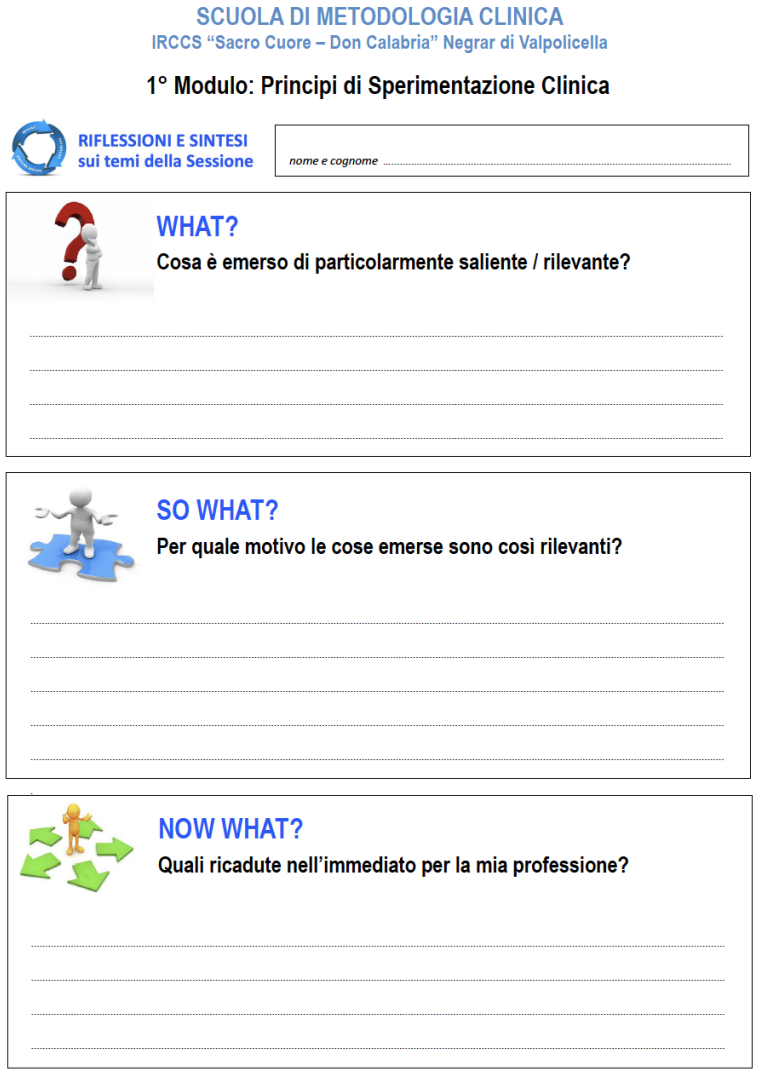
believing the results!

R. Peto

ANALISI PER SOTTOGRUPPI

- Il campione originale viene suddiviso in vari strati (**sottogruppi**) contraddistinti da caratteristiche peculiari.
- Tipologia di analisi:
  - ***Post hoc*** (analisi retrospettiva)
  - ***Pre-specified (preannunciata)*** (prevista dal protocollo di studio; criterio di stratificazione?)
  - ***Pre-planned (prepianificata)*** (prevista dal protocollo di studio con piano di analisi specifico)

1. Riflettete da soli per 10 min. e compilate il form →
2. Confrontatevi con i Colleghi del Vostro tavolo per 15 min., declinate un W^3 condiviso e delegate un portavoce
3. Riportate sulla lavagna il Vostro W^3 condiviso su almeno due aspetti ritenuti rilevanti e impattanti sulla professione (in 5 min.)
4. Presentate ai Colleghi degli altri tavoli il Vostro W^3 condiviso



SCUOLA DI METODOLOGIA CLINICA

IRCCS "Sacro Cuore – Don Calabria" Negrar di Valpolicella

1° Modulo: Principi di Sperimentazione Clinica

RIFLESSIONI E SINTESI sui temi della Sessione

nome e cognome .....

WHAT?

Cosa è emerso di particolarmente saliente / rilevante?

SO WHAT?

Per quale motivo le cose emerse sono così rilevanti?

NOW WHAT?

Quali ricadute nell'immediato per la mia professione?