SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

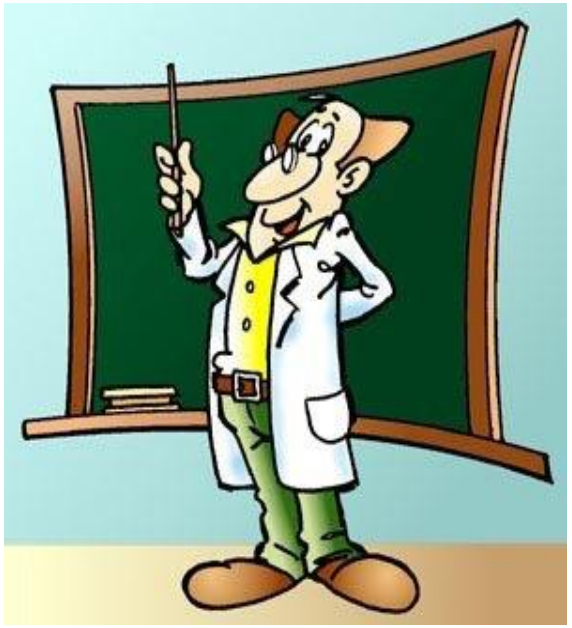
2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"



SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

Programma

Venerdì 26 Gennaio 2024

10.00-10.10 Saluti del Direttore Generale della Ricerca dell'IRCCS "Sacro Cuore - Don Calabria" Negrar di Valpolicella (VR)
Dr. Mario PICCININI

10.10-10.15 Presentazione ed obiettivi del Corso
Stefania GORI
Giovanni L. PAPPAGALLO

SESSIONE I - NOZIONI DI BASE

10.30-10.45 Etica della ricerca
Fabrizio NICOLIS

10.45-11.05 Quesito clinico di riferimento
Giovanni L. PAPPAGALLO

11.05-11.25 Fasi della sperimentazione clinica

11.25-11.45 Disegni di studio osservazionali ed interventistici *(compresi basket & umbrella design)*
Emilio BRIA

11.45-12.00 Discussione

12.00-13.00 Lavoro di gruppo
Summing-Up
(metodo: *What? So What? Now What?*)
Tutor: **Cristina MAZZI**

13.00-14.15 Colazione di lavoro

14.15-15.15 Misure di effetto relativo e assoluto
Giovanni L. PAPPAGALLO

15.15-16.15 Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
Cristina MAZZI

16.15-16.30 Discussione

16.30-16.45 Coffee Break

16.45-17.45 Lavoro di gruppo
Summing-Up
(metodo: *What? So What? Now What?*)
Tutor: **Cristina MAZZI**

18.00 Termine dei lavori della giornata

Sabato 27 Gennaio 2024

SESSIONE II
VALUTAZIONE DELLE EVIDENZE

09.00-09.45 Rilevanza clinica
Vs significatività statistica
Giovanni L. PAPPAGALLO

09.45-10.15 Trasferibilità delle evidenze al quesito clinico
Giovanni L. PAPPAGALLO

10.15-10.30 Discussione

10.30-10.45 Coffee Break

10.45-11.15 Affidabilità delle prove (rischio di bias, imprecisione degli effetti)
Giovanni L. PAPPAGALLO

11.15-12.00 Dialogo tra clinico e metodologo: "Lettura di uno studio clinico"
Ettore D'ARGENTO
Giovanni L. PAPPAGALLO

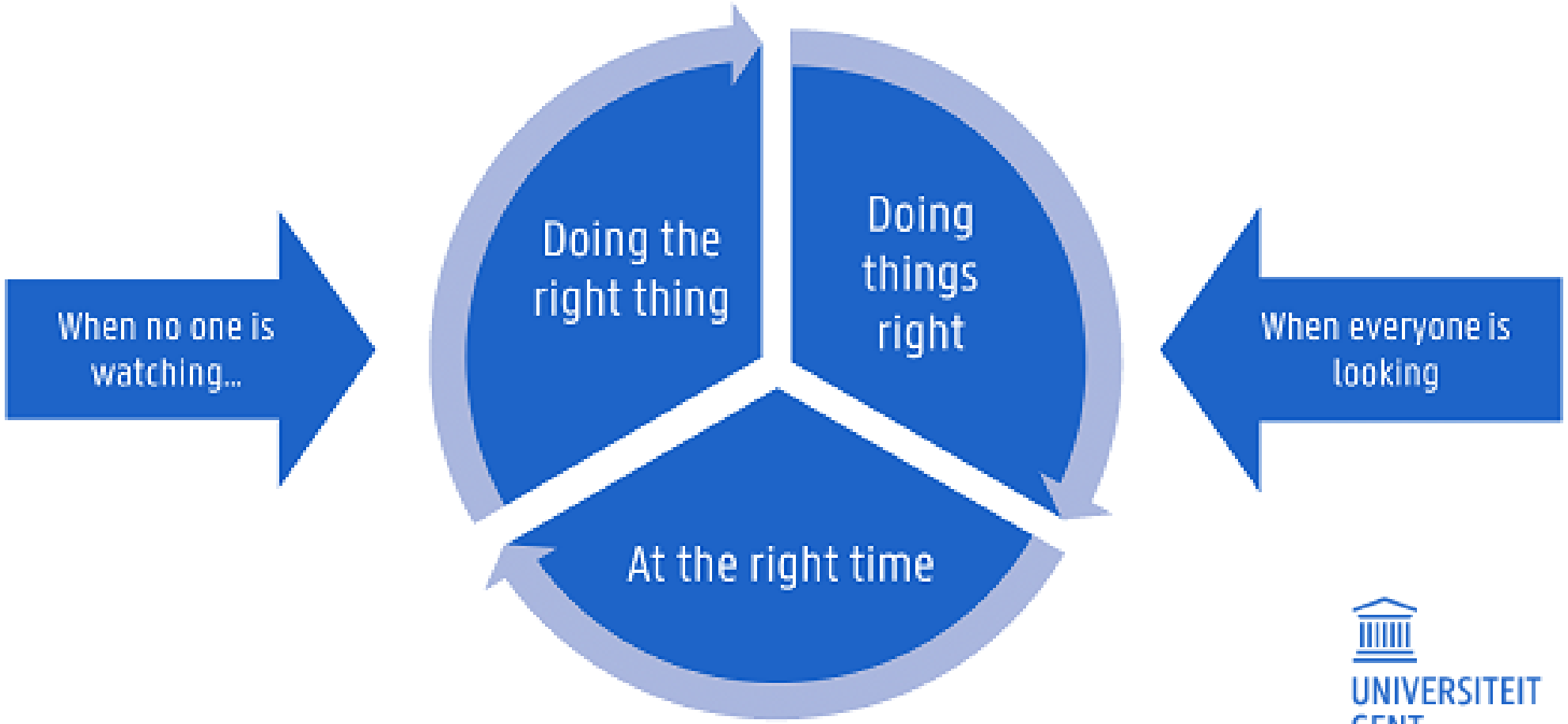
12.00-13.00 Lavoro di gruppo
Summing-Up
(metodo: *What? So What? Now What?*)
Tutor: **Cristina MAZZI**

13.00-13.30 Compilazione questionario ECM

13.30 Conclusione del Corso

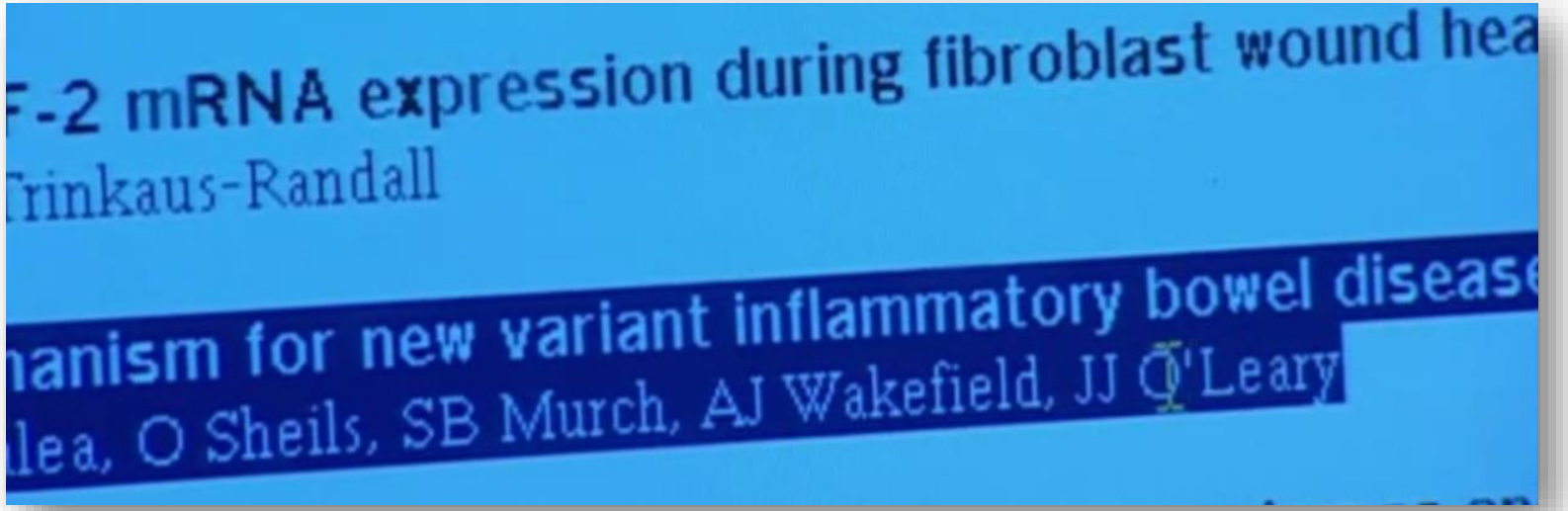
- **Etica e integrità della ricerca**
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

Research integrity: what is it about?

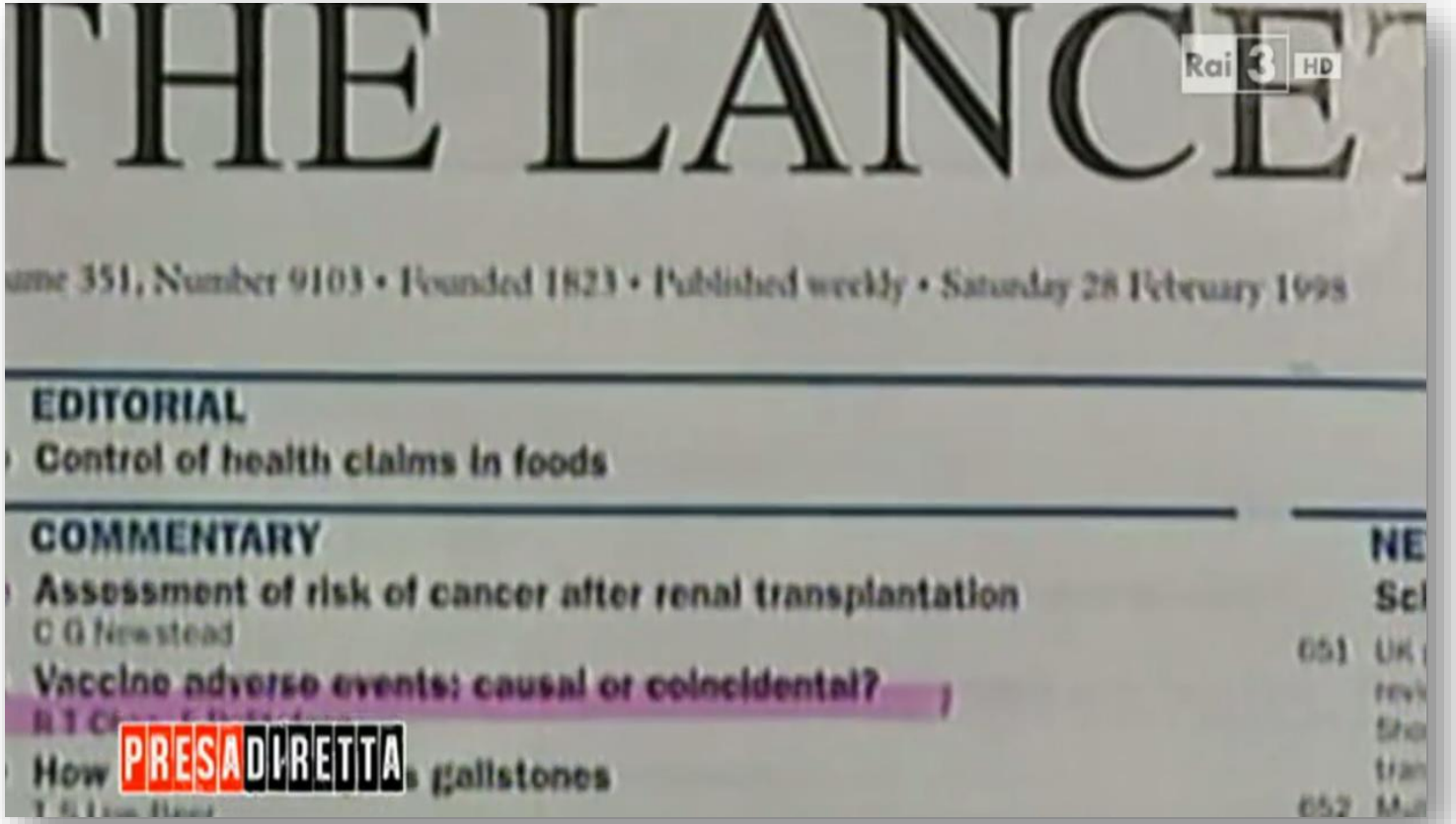


UNIVERSITEIT GENT

La storia di **Andrew Wakefield** inizia nel Royal Free Hospital, importante policlinico universitario pubblico di Londra.



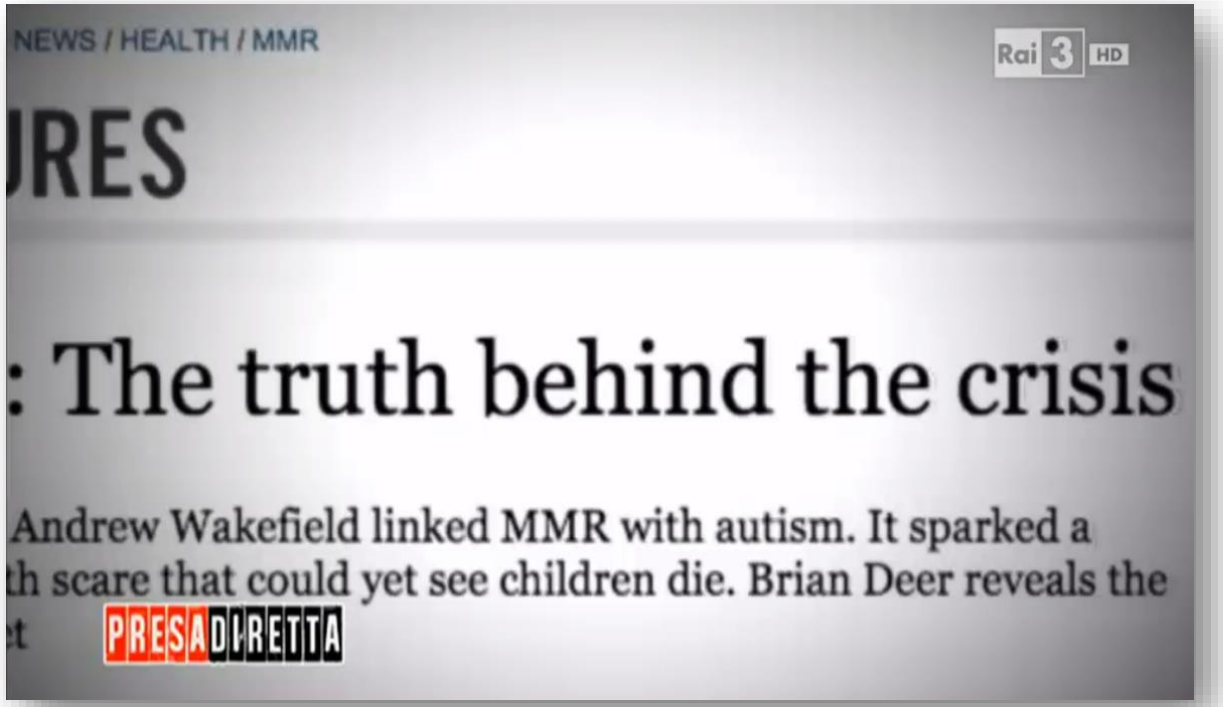
Nel **luglio 1997** Wakefield, gastroenterologo, rende pubblica una ricerca che sta conducendo nell'ospedale in cui si ipotizza il legame tra l'assunzione del triplo vaccino MPR e l'insorgenza dell'autismo e del Morbo di Crohn.



I risultati della ricerca condotta su un **campione** di soli **12 bambini** vengono pubblicati 6 mesi dopo nel **febbraio** del **1998** sulla prestigiosa rivista scientifica The Lancet *(Impact Factor 1998: 11,793 – I.F. 2021: 202,731 – I.F. 2022: 168,9).*

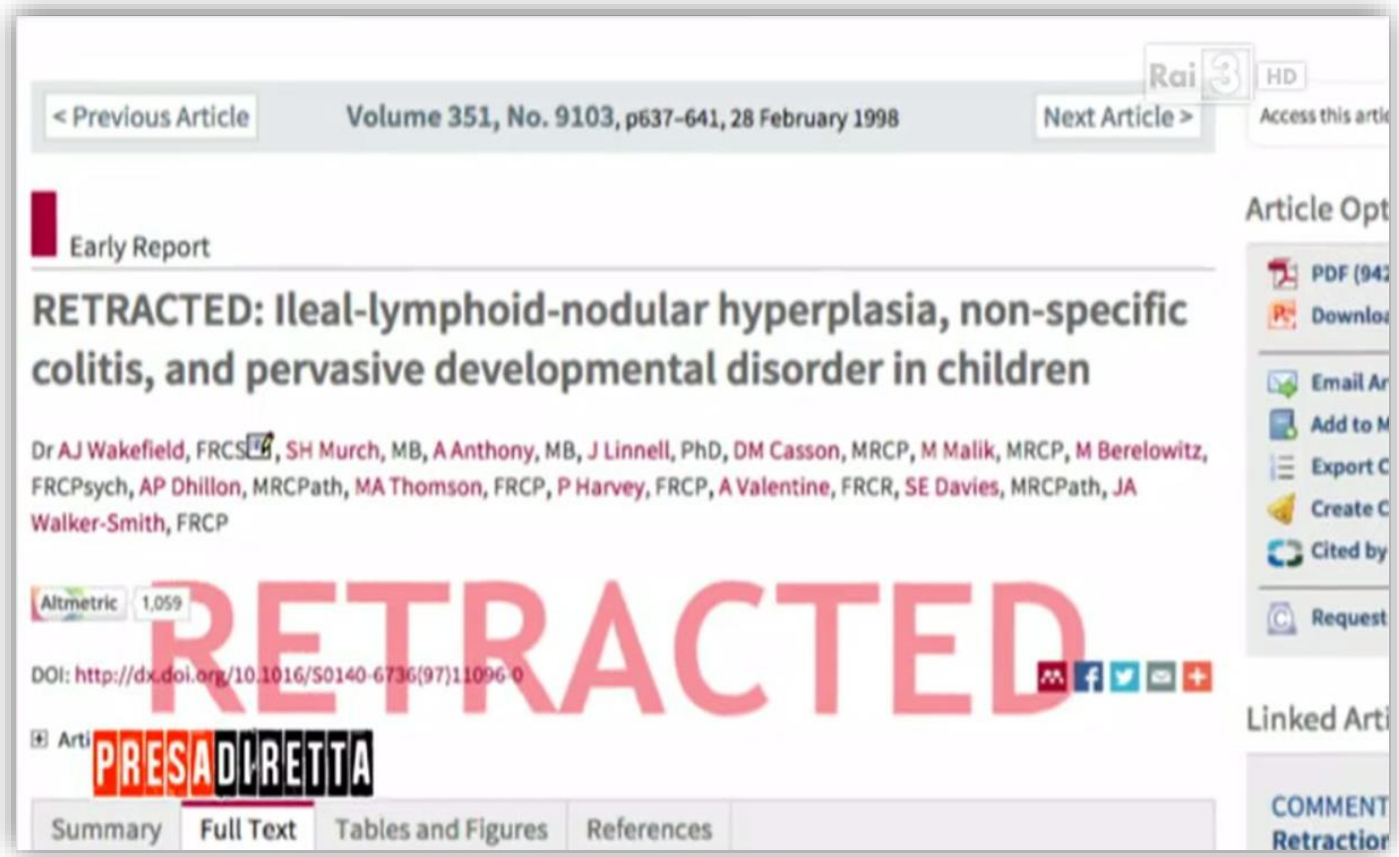
Al primo articolo ne seguono altri su The Lancet, da parte della stessa equipe di ricercatori.

…a Wakefield va tutto bene **fino a dicembre del 2003**



Un' inchiesta sul **Sunday Times** rivela la presenza di **conflitto di interessi**, sia per Wakefield che per il suo team di medici (relazione con avvocati che convincevano i genitori di bambini autistici a fare causa all'azienda farmaceutica produttrice del vaccino MPR; fondazione di una società privata che vendeva kit diagnostici da usare per i bambini autistici che secondo loro si erano ammalati per colpa dei vaccini).

Si è poi scoperto che la ricerca non reggeva dal punto di vista scientifico: era piena di **gravi errori metodologici**: nella **scelta dei campioni** analizzati e nelle **procedure di analisi**



Partial retraction:
2004

Fully retraction:
2010

La rivista The Lancet decide di cancellare per sempre l'articolo di Wakefield considerandolo una vera e propria **frode scientifica**

Wakefield è stato radiato nel 2010 dal Consiglio Generale dei Medici Britannici, che lo ha giudicato “disonesto”, “irresponsabile” e in pieno conflitto di interesse nel proprio lavoro di ricerca.

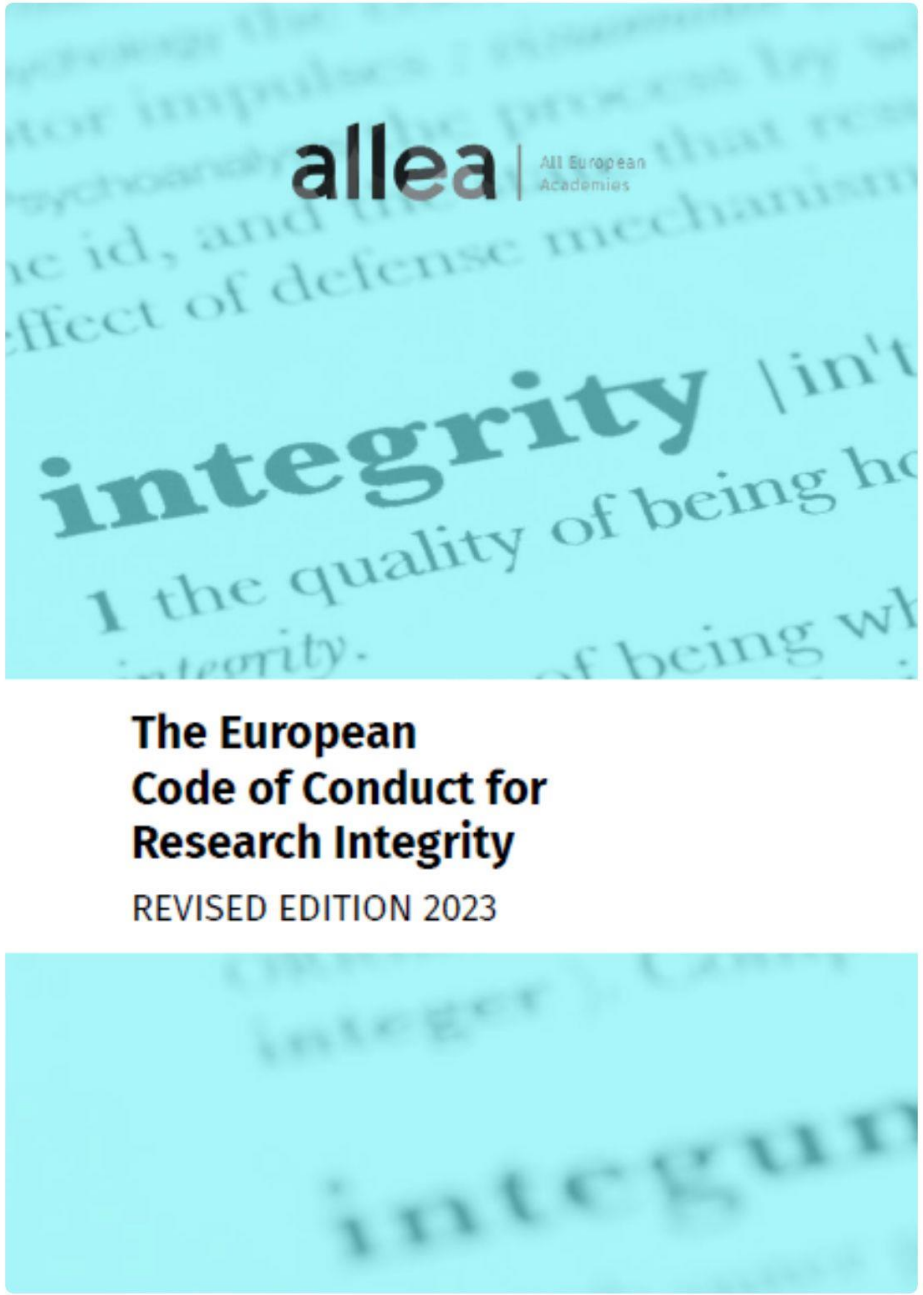
Il Consiglio lo ha ritenuto colpevole di aver agito contro l’interesse dei propri pazienti, tutti bambini disabili, e di aver mentito all’editore di un’importante rivista scientifica, The Lancet.

2007

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT
GLOBAL SCIENCE FORUM

Best Practices for Ensuring Scientific Integrity and Preventing Misconduct

The Organisation for Economic Co-operation and Development (OECD) is an **international** organisation that works to build ***better policies for better lives***. Our goal is to shape policies that foster prosperity, equality, opportunity and well-being for all. We draw on 60 years of experience and insights to better prepare the world of tomorrow.



- Prima pubblicazione: 2011
- Prima Revisione: 2017
- Seconda Revisione: 2023

ALLEA
ALL European
Academies

ALLEA Member Academies

Albania: Akademia e Shkencave e Shqipërisë; **Armenia**: Գիտությունների ազգային ակադեմիա; **Austria**: Österreichische Akademie der Wissenschaften; **Bielorussia**: Нацыянальная акадэмія навук Беларусі; **Belgio**: Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique; Koninklijke Vlaamse Academie van Belgie voor Wetenschappen en Kunsten; Koninklijke Academie voor Nederlandse Taal- en Letterkunde; Académie Royale de langue et de litterature françaises de Belgique; **Bosnia-Erzegovina**: Akademija nauka i umjetnosti Bosne i Hercegovine; **Bulgaria**: Българска академия на науките; **Croazia**: Hrvatska Akademija Znanosti i Umjetnosti; **Repubblica Ceca:** Akademie věd České republiky; Učená společnost České republiky; **Dinamarca**: Kongelige Danske Videnskabernes Selskab; **Estonia**: Eesti Teaduste Akadeemia; **Finlandia**: Tiedeakatemiain neuvottelukunta; **Francia**: Académie des Sciences - Institut de France; Académie des Inscriptions et Belles-Lettres; **Georgia:** საქართველოს მეცნიერებათა ეროვნული აკადემია; **Germania**: Leopoldina - Nationale Akademie der Wissenschaften; Union der deutschen Akademien der Wissenschaften; Akademie der Wissenschaften in Göttingen, Akademie der Wissenschaften und der Literatur Mainz, Bayerische Akademie der Wissenschaften, Berlin-Brandenburgische Akademie der Wissenschaften, Akademie der Wissenschaften in Hamburg, Heidelberger Akademie der Wissenschaften, Nordrhein-Westfälische Akademie der Wissenschaften und der Künste, Sächsische Akademie der Wissenschaften zu Leipzig (membri associati); **Grecia**: Ακαδημία Αθηνών; **Ungheria:** Magyar Tudományos Akadémia; **Irlanda**: The Royal Irish Academy - Acadamh Ríoga na hÉireann; **Israele**: האקדמיה הלאומית הישראלית למדעים; **Italia**: Accademia Nazionale dei Lincei; Istituto Veneto di Scienze, Lettere ed Arti; Accademia delle Scienze di Torino; **Kosovo**: Akademia e Shkencave dhe e Arteve e Kosovës; **Lettonia**: Latvijas Zinātņu akadēmija; **Lituania**: Lietuvos mokslų akademijos; **Macedonia**: Македонска Академија на Науките и Уметностите; **Moldova**: Academia de Ştiinţe a Moldovei; **Montenegro**: Crnogorska akademija nauka i umjetnosti; **Paesi Bassi:** Koninklijke Nederlandse Akademie van Wetenschappen; **Norvegia**: Det Norske Videnskaps-Akademi; Det Kongelige Norske Videnskabers Selskab; **Polonia**: Polska Akademia Umiejętności; Polska Akademia Nauk; **Portogallo**: Academia das Ciências de Lisboa; **Romania**: Academia Română; **Russia**: Российская академия наук (membro associato); **Serbia**: Srpska Akademija Nauka i Umetnosti; **Slovacchia**: Slovenská Akadémia Vied; **Slovenia**: Slovenska akademija znanosti in umetnosti; **Spagna**: Real Academia de Ciencias Exactas, Físicas y Naturales; Reial Acadèmia de Ciències i Arts de Barcelona; Institut d'Estudis Catalans; **Svezia**: Kungl. Vetenskapsakademien; Kungl. Vitterhets Historie och Antikvitets Akademien; **Svizzera**: Akademien der Wissenschaften Schweiz; **Turchia**: Türkiye Bilimler Akademisi; Bilim Akademisi; **Ucraina**: Національна академія наук України; **Regno Unito**: The British Academy; The Learned Society of Wales; The Royal Society; The Royal Society of Edinburgh

ALLEA
ALL European Academies

<u>Italia</u>:

- Accademia Nazionale dei Lincei
- Istituto Veneto di Scienze, Lettere ed Arti
- Accademia delle Scienze di Torino

A) I **PRINCIPI** su cui si fonda l'integrità della ricerca

B) Le **BUONE PRASSI** di ricerca

C) Le **VIOLAZIONI** dell'integrità della ricerca

4 FUNDAMENTAL PRINCIPLES OF RESEARCH INTEGRITY

The European Code of Conduct for Research Integrity, 2017

1

RELIABILITY

in ensuring the quality of research, reflected in the design, the methodology, the analysis, and the use of resources.

HONESTY

in developing, undertaking, reviewing, reporting, and communicating research in a transparent, fair, full, and unbiased way.

2

3

RESPECT

for colleagues, research participants, society, ecosystems, cultural heritage, and the environment.

ACCOUNTABILITY

for research, from idea to publication, for its management and organisation, for training, supervision, and mentoring, and for its wider impact.

4

A) PRINCIPI su cui si fonda l'integrità della ricerca

1- AFFIDABILITÀ

nel garantire la **qualità della ricerca** che si riflette nella progettazione, nella metodologia, nell'analisi e nell'uso delle risorse.

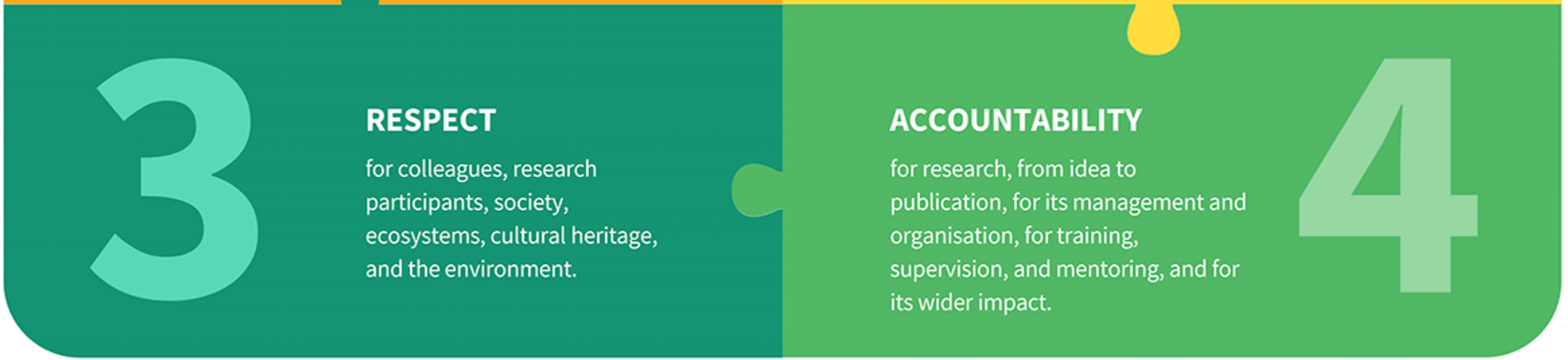
A) PRINCIPI su cui si fonda l'integrità della ricerca

1- AFFIDABILITÀ

nel garantire la **qualità della ricerca** che si riflette nella progettazione, nella metodologia, nell'analisi e nell'uso delle risorse.

2- ONESTÀ

nello sviluppare, condurre, rivedere, riferire e comunicare la **ricerca** in maniera **trasparente**, **equa**, **completa** e **obiettiva**.



A) PRINCIPI su cui si fonda l'integrità della ricerca

1- AFFIDABILITÀ
nel garantire la **qualità della ricerca** che si riflette nella progettazione, nella metodologia, nell'analisi e nell'uso delle risorse.

2- ONESTÀ
nello sviluppare, condurre, rivedere, riferire e comunicare la **ricerca** in maniera **trasparente**, **equa**, **completa** e **obiettiva**.

3- RISPETTO per:
colleghi, partecipanti alla ricerca, società, ecosistemi, patrimonio culturale e ambiente

ACCOUNTABILITY

for research, from idea to publication, for its management and organisation, for training, supervision, and mentoring, and for its wider impact.

4

A) PRINCIPI su cui si fonda l'integrità della ricerca

1- AFFIDABILITÀ
nel garantire la **qualità della ricerca** che si riflette nella progettazione, nella metodologia, nell'analisi e nell'uso delle risorse.

2- ONESTÀ
nello sviluppare, condurre, rivedere, riferire e comunicare la **ricerca** in maniera **trasparente**, **equa**, **completa** e **obiettiva**.

3- RISPETTO per:
colleghi, partecipanti alla ricerca, società, ecosistemi, patrimonio culturale e ambiente

4- RESPONSABILITÀ
dall'idea iniziale alla pubblicazione, per la gestione e organizzazione della ricerca, per la formazione, la supervisione e il tutoraggio, e infine per i suoi impatti più ampi.

B) BUONE PRASSI di ricerca

Nei diversi contesti:

- Ambiente di ricerca
- Formazione, supervisione e tutoraggio
- Procedure di ricerca
- Salvaguardie (ambiente, rapporti tra ricercatori, etc.)
- Qualità e gestione dei dati
- Collaborazione
- Pubblicazione e diffusione
- Revisione, valutazione ed editing

C) VIOLAZIONI dell'integrità della ricerca

Le situazioni principali in cui si configura una ***frode scientifica***:

- **FABBRICAZIONE**: invenzione di risultati che vengono registrati come se fossero reali
- **FALSIFICAZIONE**: manipolazione di materiali, attrezzature o processi di ricerca, oppure ingiustificata modifica, omissione o soppressione di dati o risultati
- **PLAGIO**: utilizzo dei lavori e delle idee di altre persone senza citare la fonte originaria, violando così i diritti dell'autore o degli autori originari sulla propria produzione intellettuale

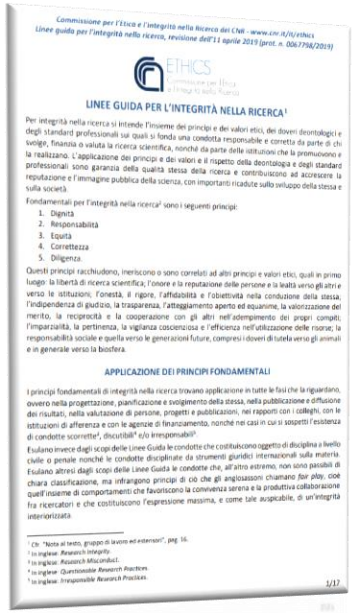
Consiglio Nazionale delle Ricerche

Commissione per l'Etica e l'Integrità nella Ricerca

ETHICS
Commissione per l'Etica
e l'Integrità nella Ricerca

La Commissione per l'Etica e l'Integrità nella Ricerca è presieduta dal Presidente del Cnr. La Commissione è un organismo indipendente con funzioni di consulenza in materia di etica della ricerca, bioetica e biodiritto, inclusi gli aspetti etici, deontologici e giuridici ricompresi nell'ambito della integrità nella ricerca (*Research Integrity*), così come descritta in letteratura scientifica e nelle principali Carte e Convenzioni internazionali nonché nelle "Linee guida per l'integrità nella ricerca" del Cnr, approvate il 10 giugno 2015 e revisionate nel 2019.

Commissione per l'Etica e l'Integrità nella Ricerca del CNR - www.cnr.it/it/ethics
Linee guida per l'integrità nella ricerca, revisione dell'11 aprile 2019 (prot. n. 0067798/2019)


Commissione per l'Etica e l'Integrità nella Ricerca del CNR - www.cnr.it/it/ethics
Linee guida per l'integrità nella ricerca, revisione dell'11 aprile 2019 (prot. n. 0067798/2019)

ETHICS
Commissione per l'Etica e l'Integrità nella Ricerca

LINEE GUIDA PER L'INTEGRITÀ NELLA RICERCA[1]

Per integrità nella ricerca si intende l'insieme dei principi e dei valori etici, dei doveri deontologici e degli standard professionali sui quali si fonda una condotta responsabile e corretta da parte di chi svolge, finanzia o valuta la ricerca scientifica, nonché da parte delle istituzioni che la promuovono e la realizzano. L'applicazione dei principi e dei valori e il rispetto della deontologia e degli standard professionali sono garanzia della qualità stessa della ricerca e contribuiscono ad accrescere la reputazione e l'immagine pubblica della scienza, con importanti ricadute sullo sviluppo della stessa e sulla società.

Fondamentali per l'integrità nella ricerca[2] sono i seguenti principi:

1. Dignità
2. Responsabilità
3. Equità
4. Correttezza
5. Diligenza.

Questi principi racchiudono, ineriscono o sono correlati ad altri principi e valori etici, quali in primo luogo: la libertà di ricerca scientifica; l'onore e la reputazione delle persone e la lealtà verso gli altri e verso le istituzioni; l'onestà, il rigore, l'affidabilità e l'obiettività nella conduzione della stessa; l'indipendenza di giudizio, la trasparenza, l'atteggiamento aperto ed equanime, la valorizzazione del merito, la reciprocità e la cooperazione con gli altri nell'adempimento dei propri compiti; l'imparzialità, la pertinenza, la vigilanza coscienziosa e l'efficienza nell'utilizzazione delle risorse; la responsabilità sociale e quella verso le generazioni future, compresi i doveri di tutela verso gli animali e in generale verso la biosfera.

APPLICAZIONE DEI PRINCIPI FONDAMENTALI

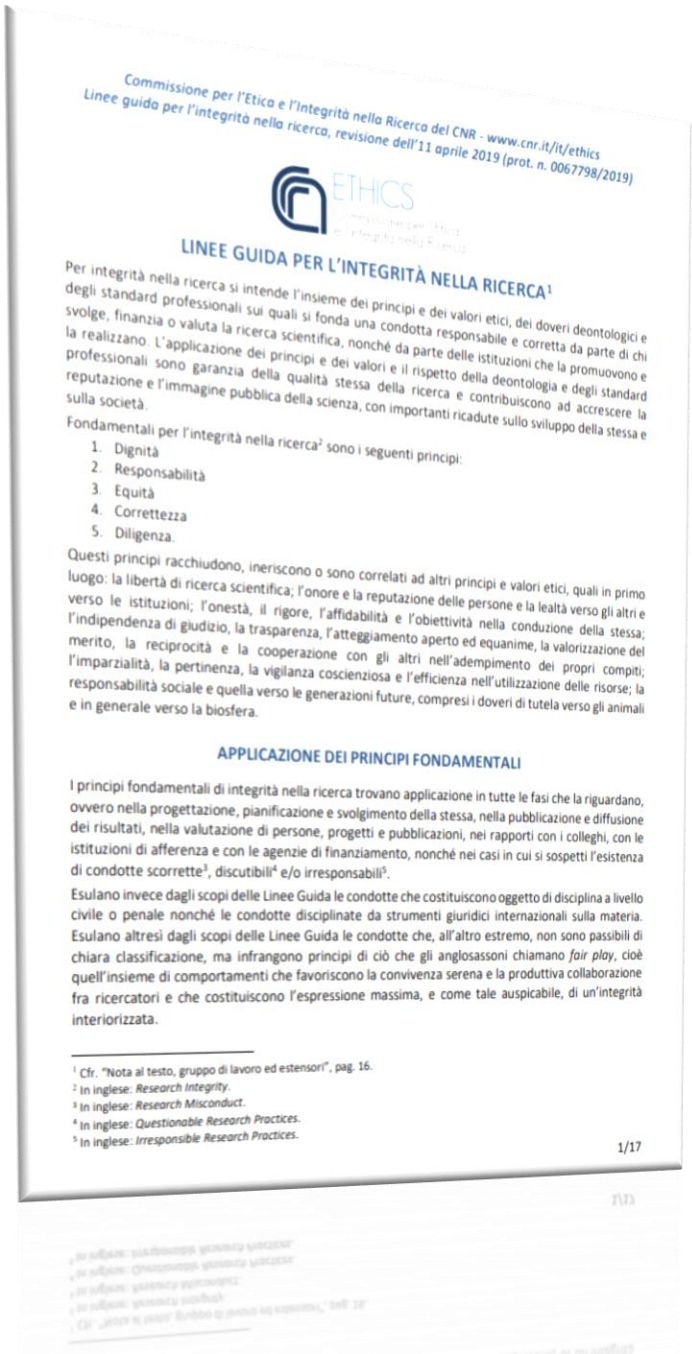
I principi fondamentali di integrità nella ricerca trovano applicazione in tutte le fasi che la riguardano, ovvero nella progettazione, pianificazione e svolgimento della stessa, nella pubblicazione e diffusione dei risultati, nella valutazione di persone, progetti e pubblicazioni, nei rapporti con i colleghi, con le istituzioni di afferenza e con le agenzie di finanziamento, nonché nei casi in cui si sospetti l'esistenza di condotte scorrette[3], discutibili[4] e/o irresponsabili[5].

Esulano invece dagli scopi delle Linee Guida le condotte che costituiscono oggetto di disciplina a livello civile o penale nonché le condotte disciplinate da strumenti giuridici internazionali sulla materia. Esulano altresì dagli scopi delle Linee Guida le condotte che, all'altro estremo, non sono passibili di chiara classificazione, ma infrangono principi di ciò che gli anglosassoni chiamano *fair play*, cioè quell'insieme di comportamenti che favoriscono la convivenza serena e la produttiva collaborazione fra ricercatori e che costituiscono l'espressione massima, e come tale auspicabile, di un'integrità interiorizzata.

[1] Cfr. "Nota al testo, gruppo di lavoro ed estensori", pag. 16.
[2] In inglese: *Research integrity*.
[3] In inglese: *Research Misconduct*.
[4] In inglese: *Questionable Research Practices*.
[5] In inglese: *Irresponsible Research Practices*.

1/17

LINEE GUIDA PER L'INTEGRITÀ NELLA RICERCA[1]



Commissione per l'Etica e l'Integrità nella Ricerca del CNR - www.cnr.it/it/ethics
Linee guida per l'integrità nella ricerca, revisione dell'11 aprile 2019 (prot. n. 0067798/2019)

ETHICS

LINEE GUIDA PER L'INTEGRITÀ NELLA RICERCA[1]

Per integrità nella ricerca si intende l'insieme dei principi e dei valori etici, dei doveri deontologici e degli standard professionali sui quali si fonda una condotta responsabile e corretta da parte di chi svolge, finanzia o valuta la ricerca scientifica, nonché da parte delle istituzioni che la promuovono e la realizzano. L'applicazione dei principi e dei valori e il rispetto della deontologia e degli standard professionali sono garanzia della qualità stessa della ricerca e contribuiscono ad accrescere la reputazione e l'immagine pubblica della scienza, con importanti ricadute sullo sviluppo della stessa e sulla società.

Fondamentali per l'integrità nella ricerca[2] sono i seguenti principi:

1. Dignità
2. Responsabilità
3. Equità
4. Correttezza
5. Diligenza.

Questi principi racchiudono, ineriscono o sono correlati ad altri principi e valori etici, quali in primo luogo: la libertà di ricerca scientifica; l'onore e la reputazione delle persone e la lealtà verso gli altri e verso le istituzioni; l'onestà, il rigore, l'affidabilità e l'obiettività nella conduzione della stessa; l'indipendenza di giudizio, la trasparenza, l'atteggiamento aperto ed equanime, la valorizzazione del merito, la reciprocità e la cooperazione con gli altri nell'adempimento dei propri compiti; l'imparzialità, la pertinenza, la vigilanza coscienziosa e l'efficienza nell'utilizzazione delle risorse; la responsabilità sociale e quella verso le generazioni future, compresi i doveri di tutela verso gli animali e in generale verso la biosfera.

APPLICAZIONE DEI PRINCIPI FONDAMENTALI

I principi fondamentali di integrità nella ricerca trovano applicazione in tutte le fasi che la riguardano, ovvero nella progettazione, pianificazione e svolgimento della stessa, nella pubblicazione e diffusione dei risultati, nella valutazione di persone, progetti e pubblicazioni, nei rapporti con i colleghi, con le istituzioni di afferenza e con le agenzie di finanziamento, nonché nei casi in cui si sospetti l'esistenza di condotte scorrette[3], discutibili[4] e/o irresponsabili[5].

Esulano invece dagli scopi delle Linee Guida le condotte che costituiscono oggetto di disciplina a livello civile o penale nonché le condotte disciplinate da strumenti giuridici internazionali sulla materia. Esulano altresì dagli scopi delle Linee Guida le condotte che, all'altro estremo, non sono passibili di chiara classificazione, ma infrangono principi di ciò che gli anglosassoni chiamano *fair play*, cioè quell'insieme di comportamenti che favoriscono la convivenza serena e la produttiva collaborazione fra ricercatori e che costituiscono l'espressione massima, e come tale auspicabile, di un'integrità interiorizzata.

[1] Cfr. "Nota al testo, gruppo di lavoro ed estensori", pag. 16.
[2] In inglese: *Research Integrity*.
[3] In inglese: *Research Misconduct*.
[4] In inglese: *Questionable Research Practices*.
[5] In inglese: *Irresponsible Research Practices*.

1/17

A) PRINCIPI FONDAMENTALI per l'integrità nella ricerca

B) CONDOTTE che **PROMUOVONO** l'integrità nella ricerca, in tutte le fasi della ricerca (progettazione, svolgimento, pubblicazione dei risultati, etc.)

C) CONDOTTE LESIVE dell'integrità nella ricerca, in tutte le fasi della ricerca

CONDOTTE che **PROMUOVONO** l'integrità nella ricerca

Un esempio

3. Educare all'integrità nella ricerca: le istituzioni scientifiche e chi in esse riveste ruoli di coordinamento o direzione scientifica o amministrativa contribuiscono, nelle forme consentite dal proprio ruolo, a formare i ricercatori riguardo ai principi dell'integrità nella ricerca e in generale alle responsabilità sociali implicate dalle sperimentazioni.

ETHICS
Commissione per l'Etica
e l'Integrità nella Ricerca

LINEE GUIDA PER L'INTEGRITÀ NELLA RICERCA

CONDOTTE che **PROMUOVONO** l'integrità nella ricerca

Un esempio

5. Comunicare con obiettività e responsabilità: nella misura consentita da ciascuna diversa forma e modalità di pubblicazione, i ricercatori forniscono in modo scrupoloso, obiettivo e imparziale la maggior quantità possibile di elementi e informazioni anche su aspetti quali:
    a. la letteratura fondamentale e le conoscenze antecedenti lo studio;
    b. lo scopo originario e i metodi definiti *prima* dello svolgimento della ricerca;
    c. le eventuali modifiche negli obiettivi e nelle metodologie intercorse *dopo* l'avvio della ricerca;
    d. i risultati significativi conseguiti, compresi quelli negativi o nulli;
    e. le possibili interpretazioni, l'ambito di applicabilità e le limitazioni dei risultati conseguiti.

ETHICS
Commissione per l'Etica e l'Integrità nella Ricerca

Research Letter

April 16, 2020

The Use of Not-Negative Conclusions to Describe Results of Formally Negative Trials Presented at Oncology Meetings

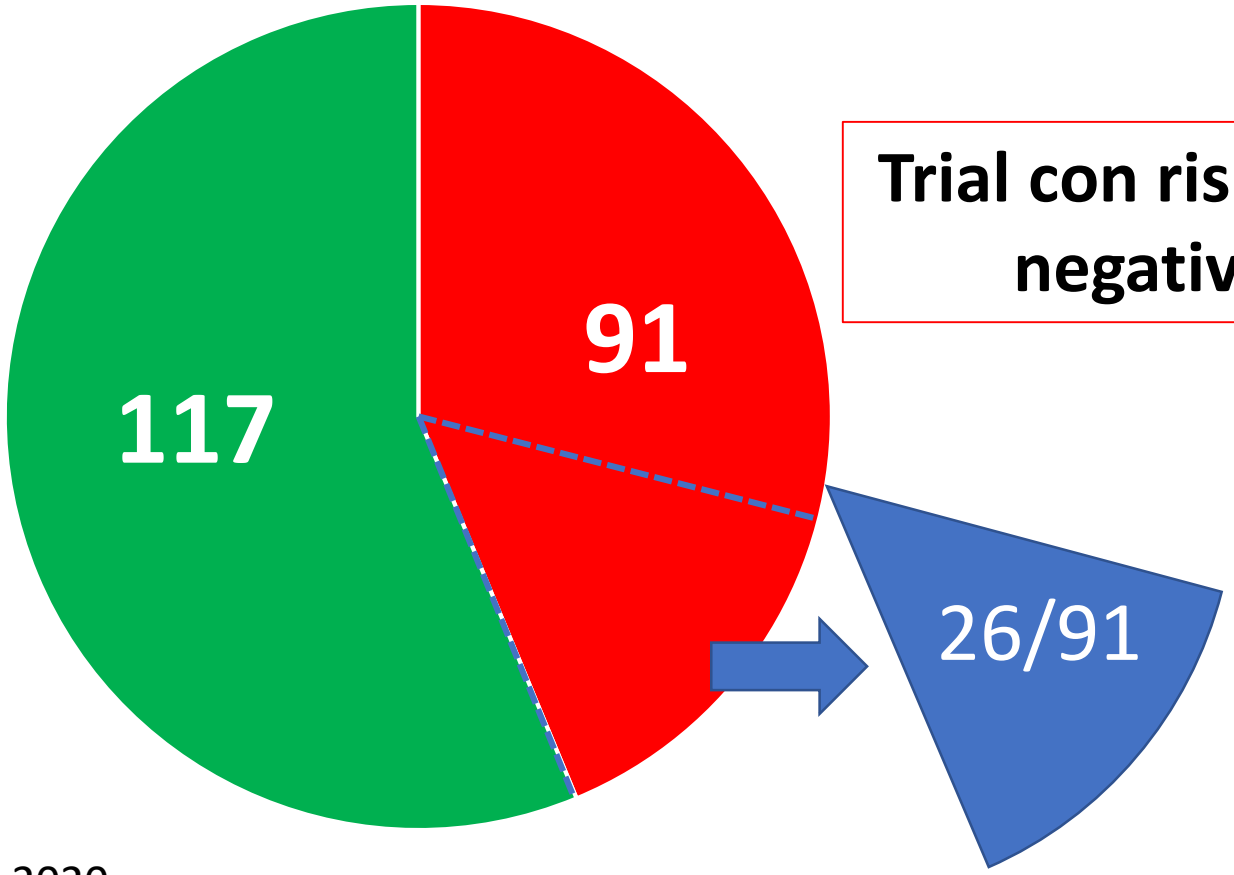
Massimo Di Maio, MD[1,2]; Marco Audisio, MD[1,2]; Claudia Cardone, MD[3,4]; et al

» Author Affiliations | Article Information

JAMA Oncol. 2020;6(6):926-927. doi:10.1001/jamaoncol.2020.0475

Trial Clinici randomizzati di Fase 3
Presentati ai congressi ASCO e ESMO dal 2017 al 2019 (tot=208)

Trial con risultati positivi



Trial con risultati negativi

Trial presentati con CONCLUSIONI NON NEGATIVE: il 29% dei trial con risultati negativi

Di Maio M et al., JAMA Oncol. 2020

I Trial con risultati negativi MA presentati con CONCLUSIONI NON NEGATIVE:

- Sono aumentati nel corso degli anni
- Sono risultati essere maggiori ai congressi ASCO vs ESMO
- Sono superiori tra i trial no-profit vs profit

Di Maio M et al., JAMA Oncol. 2020

NOT-NEGATIVE CONCLUSIONS:

(1) numerically better outcome in the experimental arm, despite a nonsignificant P value,

(2) emphasis on positive subgroup(s),

(3) emphasis on positive secondary end point (s),

(4) noninferiority interpretation of a negative superiority trial

Di Maio M et al., JAMA Oncol. 2020

In conclusion, we believe that more attention should be paid to the statements included in the conclusions of oral presentations at meetings, and the discussants' role is crucial. When the primary end point is not met, the word *negative* should be explicitly used.

Di Maio M et al., JAMA Oncol. 2020

CONDOTTE LESIVE dell'integrità nella ricerca

Un esempio

PRATICHE DISCUTIBILI E/O IRRESPONSABILI:

- ostacolare, rallentare, o sabotare indirettamente e involontariamente il lavoro dei colleghi attraverso la non-condivisione protratta oltre i limiti professionalmente e scientificamente giustificabili, di dati, metodi, risultati negativi di esperimenti, informazioni su errori metodologici o di altro tipo;

World Health Organization

Joint statement on public disclosure of results from clinical trials

18 May 2017 | Departmental news | Reading time: 8 min (2142 words)

Some of the world's largest funders of medical research and international non-governmental organizations agreed on new standards that will require all clinical trials they fund or support to be registered and the results disclosed publicly. Currently, about 50% of clinical trials go unreported, often because the results are negative. These unreported trial results leave an incomplete and potentially misleading picture of the risks and benefits of vaccines, drugs and medical devices, and can lead to use of suboptimal or even harmful products

DECRETO LEGISLATIVO 23 dicembre 2022, n. **200**.

Riordino della disciplina degli Istituti di ricovero e cura a carattere scientifico.

G.U. Serie generale - n. 304, 30-12-2022

Art. 4. Modifiche all'articolo 8 del decreto legislativo 16 ottobre 2003, n. 288

5-bis. Gli Istituti, nel rispetto della legge 31 maggio 2022, n. 62, garantiscono che l'attività di ricerca e cura si conformi ai **principi** della **correttezza**, **trasparenza**, **equità**, **responsabilità**, **affidabilità** e **completezza** riconosciuti a livello internazionale. Essi pubblicano tutti i dati e le fonti della ricerca in modo veritiero e oggettivo, al fine di consentire la verifica e la riproducibilità, con specifico riferimento al mantenimento dei dati utilizzati. A tal fine, per garantire la valutazione dell'attività scientifica, anche con riguardo agli effetti di quest'ultima sulla salute della popolazione, utilizzano indicatori di efficacia ed efficienza della qualità dell'attività di ricerca riconosciuti a livello internazionale. **Gli Istituti adottano e aggiornano periodicamente un codice di condotta per l'integrità della ricerca**. Il personale in servizio presso gli IRCCS è tenuto ad aderire ad un codice di condotta che disciplina prescrizioni comportamentali volte al corretto utilizzo delle risorse e al rispetto di regole di *fair competition (concorrenza leale).*

TESTO BASE PER L'ADOZIONE DI UN CODICE DI CONDOTTA PER L'INTEGRITÀ DELLA RICERCA

APPENDICE
Estensori del documento

Alla redazione del presente documento hanno collaborato i seguenti componenti, elencati in ordine alfabetico, del gruppo di lavoro appositamente costituito su iniziativa del Ministero della Salute, Direzione generale della ricerca e dell'innovazione in sanità:

Roberto Buccione
IRCCS Ospedale San Raffaele

Francesca Colazzo
Centro Cardiologico Monzino

Rosalba Miceli
Istituto Nazionale dei Tumori

Massimo Monturano
Istituto Europeo di Oncologia

Paola Perego
Istituto Nazionale dei Tumori

Per il Ministero della Salute:

Maria Chiara Siracusa
Direzione Generale della ricerca e dell'innovazione in sanità
Ufficio 5

SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

- Etica e integrità della ricerca
- **Quesito clinico di riferimento**
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

Una situazione (ahimé) comune…

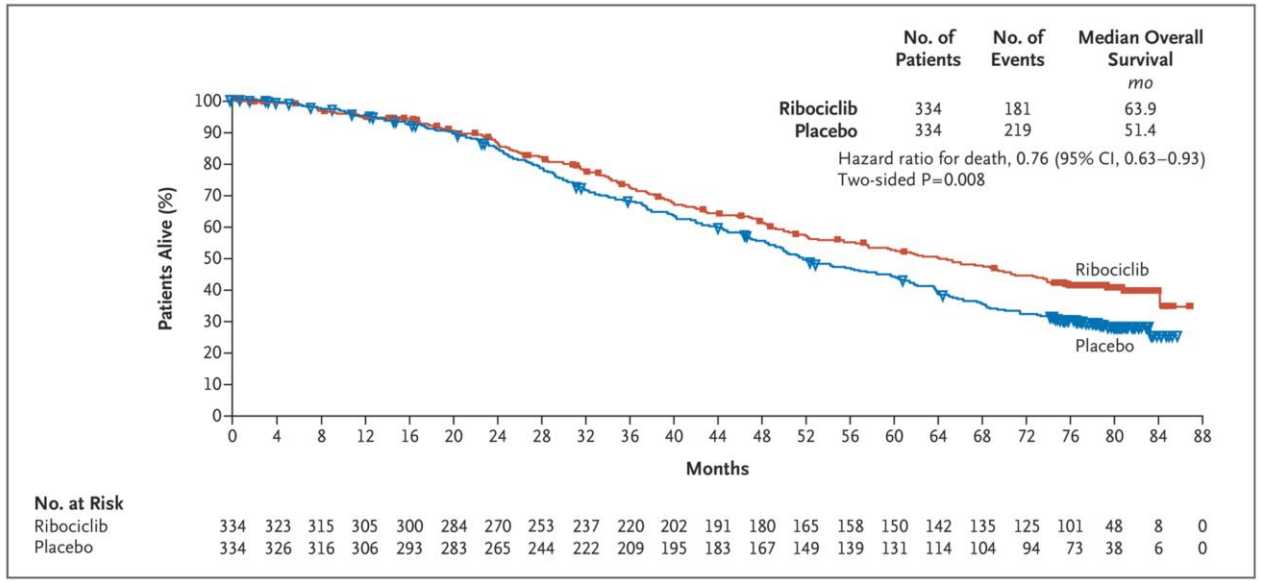
Gigi, che ne pensi di PALOMA-2, MONALEESA-2 e MONARCH-3?

In che senso, scusa?

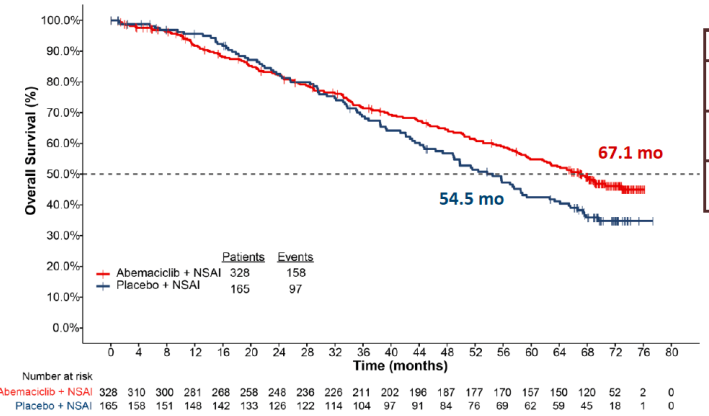
Ma a che tipologia di pazienti ti riferisci?

… qual è il "migliore" tra questi tre studi?

ER+, HER2-, MBC ovviamente…



OS IA2 for the ITT Population



	abemaciclib + NSAI	placebo + NSAI
Median OS, (months)	67.1	54.5
HR (95% CI; *P* value)	0.754 (0.584-0.974) p-value 0.0301*	
Pre-planned OS IA2 Analysis Data cut: 02 Jul 2021		

*p-value did not reach threshold for statistical significance at this interim

31.5% of patients in the control arm and 10.1% in the abemaciclib arm received a subsequent CDK4 & 6 inhibitor

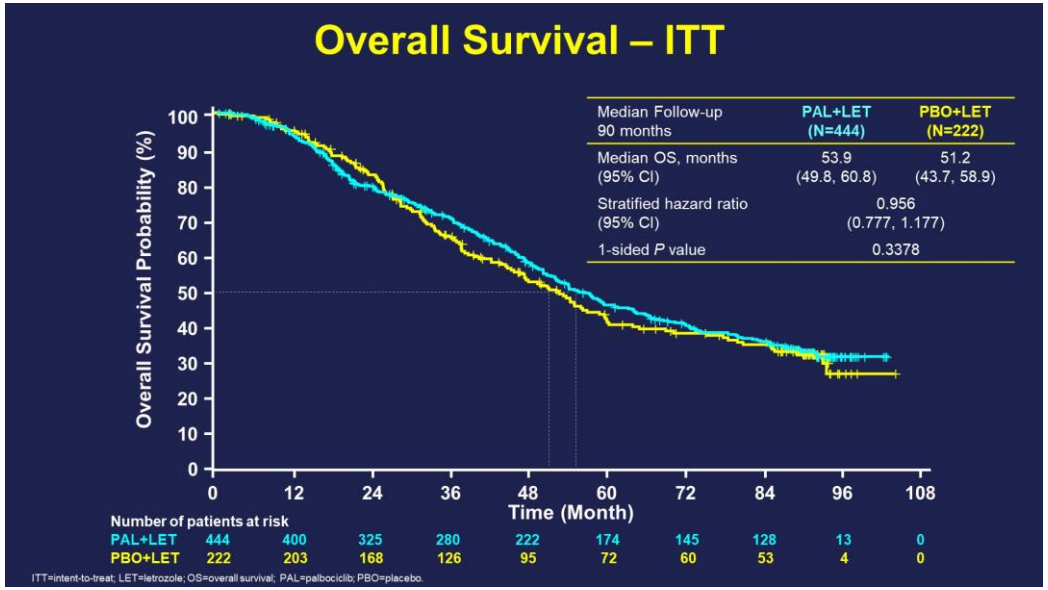
At this interim analysis, statistical significance was not reached but data are maturing favorably (HR 0.754, 95% CI: 0.584-0.974) and follow up continues. The observed difference in median OS was 12.6 months.

PARIS 2022 ESMO congress

Matthew P. Goetz, MD

CDK; cyclin dependent kinase; CI, confidence interval; HR, hazard ratio; IA2, interim analysis 2; mo, months; ITT, intent to treat; NSAI, nonsteroidal aromatase inhibitor; OS, overall survival.

Overall Survival – ITT



ITT=intent-to-treat; LET=letrozole; OS=overall survival; PAL=palbociclib; PBO=placebo.

Una situazione (ahimé) comune…

Gigi, che ne pensi di PALOMA-2, MONALEESA-2 e MONARCH-3?

… qual è il "migliore" tra questi tre studi?

ER+, HER2-, MBC ovviamente…

In che senso, scusa?

Ma a che tipologia di pazienti ti riferisci?

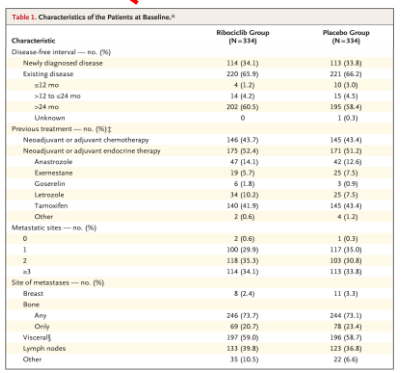
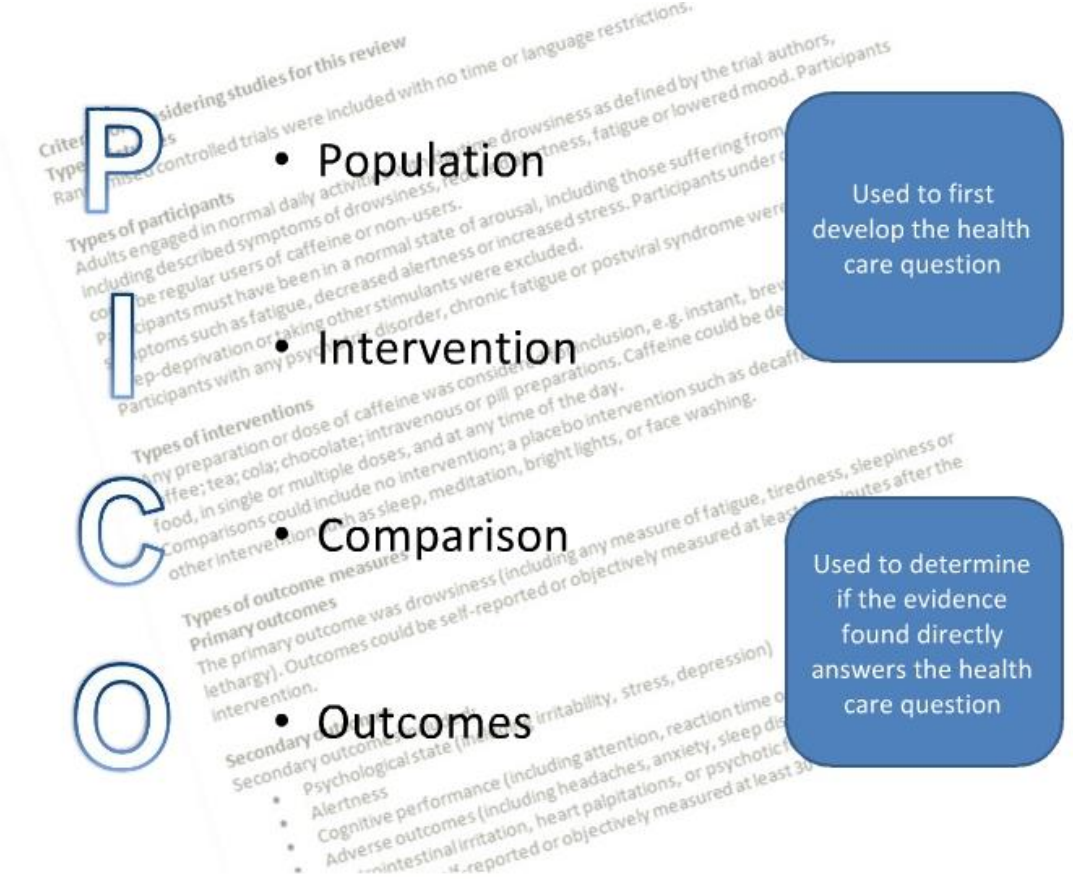


Table 1. Characteristics of the Patients at Baseline.*

Characteristic	Ribociclib Group (N = 334)	Placebo Group (N = 334)
Disease-free interval — no. (%)		
Newly diagnosed disease	114 (34.1)	113 (33.8)
Existing disease	220 (65.9)	221 (66.2)
≤12 mo	4 (1.2)	10 (3.0)
>12 to ≤24 mo	14 (4.2)	15 (4.5)
>24 mo	202 (60.5)	195 (58.4)
Unknown	0	1 (0.3)
Previous treatment — no. (%)‡		
Neoadjuvant or adjuvant chemotherapy	146 (43.7)	145 (43.4)
Neoadjuvant or adjuvant endocrine therapy	175 (52.4)	171 (51.2)
Anastrozole	47 (14.1)	42 (12.6)
Exemestane	19 (5.7)	25 (7.5)
Goserelin	6 (1.8)	3 (0.9)
Letrozole	34 (10.2)	25 (7.5)
Tamoxifen	140 (41.9)	145 (43.4)
Other	2 (0.6)	4 (1.2)
Metastatic sites — no. (%)		
0	2 (0.6)	1 (0.3)
1	100 (29.9)	117 (35.0)
2	118 (35.3)	103 (30.8)
≥3	114 (34.1)	113 (33.8)
Site of metastases — no. (%)		
Breast	8 (2.4)	11 (3.3)
Bone		
Any	246 (73.7)	244 (73.1)
Only	69 (20.7)	78 (23.4)
Visceral§	197 (59.0)	196 (58.7)
Lymph nodes	133 (39.8)	123 (36.8)
Other	35 (10.5)	22 (6.6)

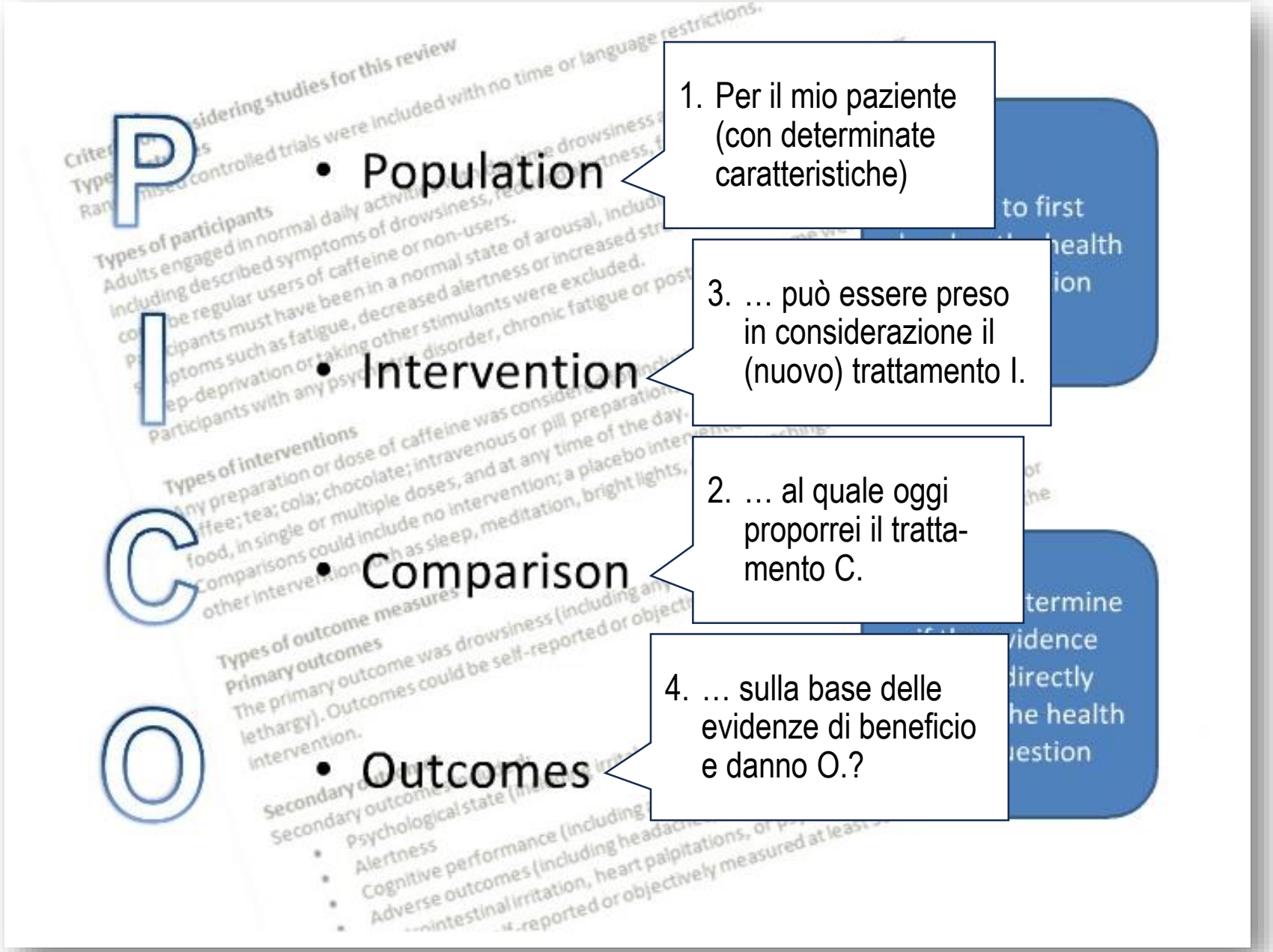
PREMESSA

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



Important Questions

Should be
from practice
NOT
evidence driven


P
• Population
I
• Intervention
C
• Comparison
O
• Outcomes
1. Per il mio paziente (con determinate caratteristiche)
3. … può essere preso in considerazione il (nuovo) trattamento I.
2. … al quale oggi proporrei il tratta-mento C.
4. … sulla base delle evidenze di beneficio e danno O.?

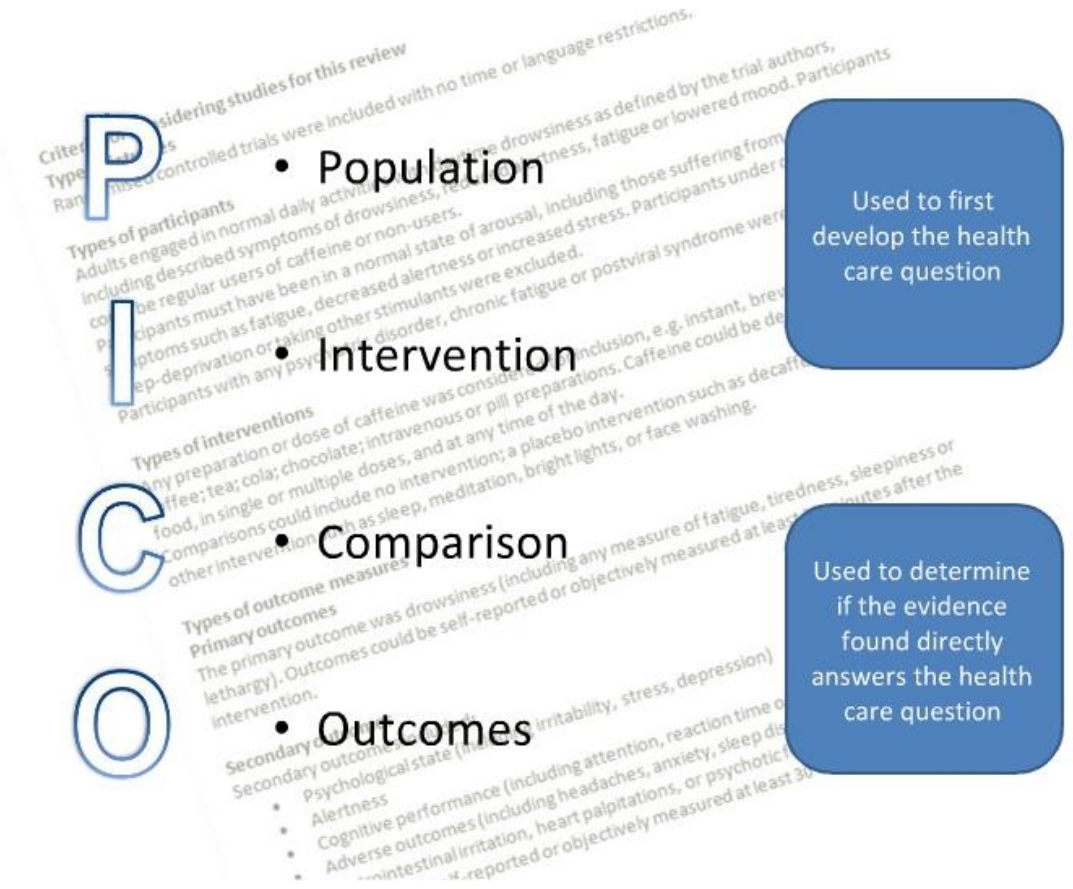
Si chiede quindi alla Letteratura l'evidenza quanto più attinente al quesito clinico:

ER+, HER2-, *recurrent*, MBC

CDK 4/6 inhibitor + AI

AI alone

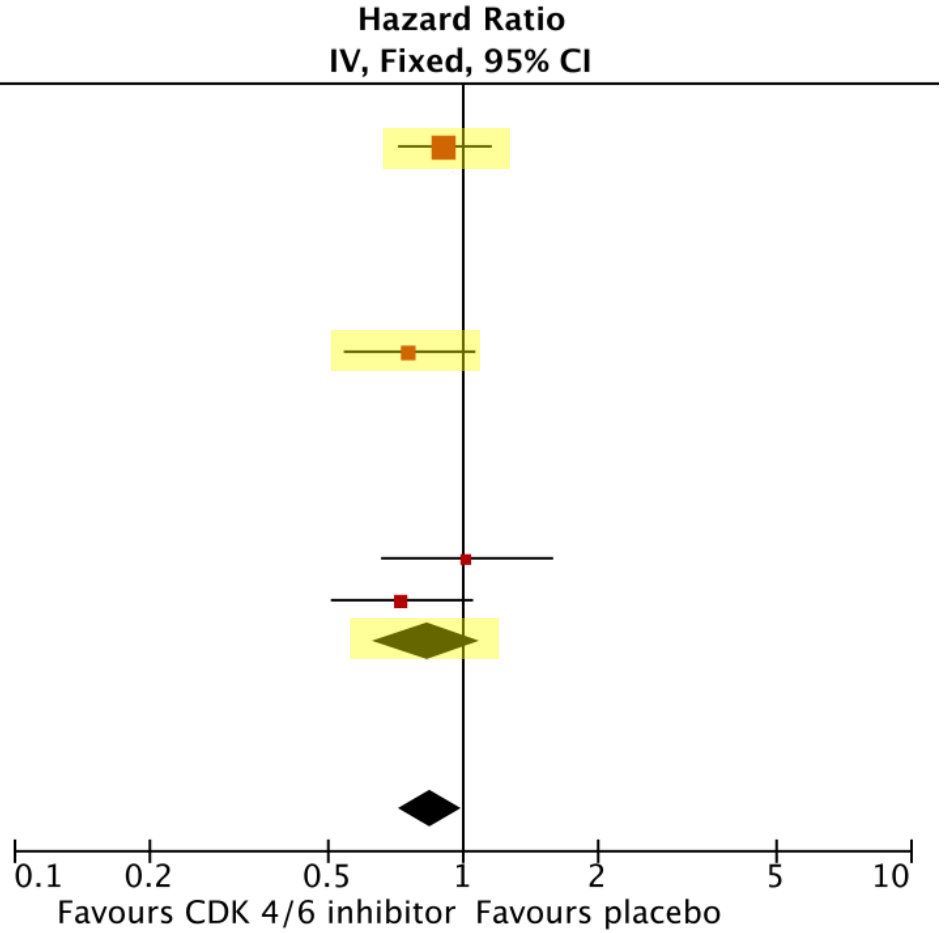
Overall Survival



CDK4/6 inhibitors as 1st-line therapy for HR-positive, advanced breast cancer

Overall Survival – recurrent metastatic disease

Study or Subgroup	log[Hazard Ratio]	SE	Weight	Hazard Ratio IV, Fixed, 95% CI
1.1.1 Monaleesa 2				
Monaleesa 2 (recurrent metastatic disease)	-0.0943	0.1195	**45.1%**	**0.91 [0.72, 1.15]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 0.79 (P = 0.43)				
1.1.2 Monarch 3				
Monarch 3 (recurrent metastatic disease)	-0.2758	0.169	**22.6%**	**0.76 [0.54, 1.06]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 1.63 (P = 0.10)				
1.1.3 Paloma 2				
Paloma 2 (DFI ≤12 months)	0.0198	0.2221	13.1%	1.02 [0.66, 1.58]
Paloma 2 (DFI >12 months)	-0.3147	0.183	19.2%	0.73 [0.51, 1.04]
Subtotal (95% CI)			**32.3%**	**0.84 [0.63, 1.10]**
Heterogeneity: $Chi^2 = 1.35$, df = 1 (P = 0.25); $I^2 = 26\%$				
Test for overall effect: Z = 1.27 (P = 0.20)				
Total (95% CI)			**100.0%**	**0.85 [0.73, 0.99]**
Heterogeneity: $Chi^2 = 2.14$, df = 3 (P = 0.54); $I^2 = 0\%$				
Test for overall effect: Z = 2.03 (P = 0.04)				
Test for subgroup differences: $Chi^2 = 0.79$, df = 2 (P = 0.67), $I^2 = 0\%$				



Hortobagyi GN, et al. N Engl J Med 2022;386:942-50 - Goetz MP, et al. Annals of Oncology (2022) 33 (suppl_7): S808-S869. 10.1016/annonc/annonc1089 - Finn RS, et al. Journal of Clinical Oncology 40, no. 17_suppl (June 10, 2022) LBA1003-LBA1003.

G.L. Pappagallo, 2023

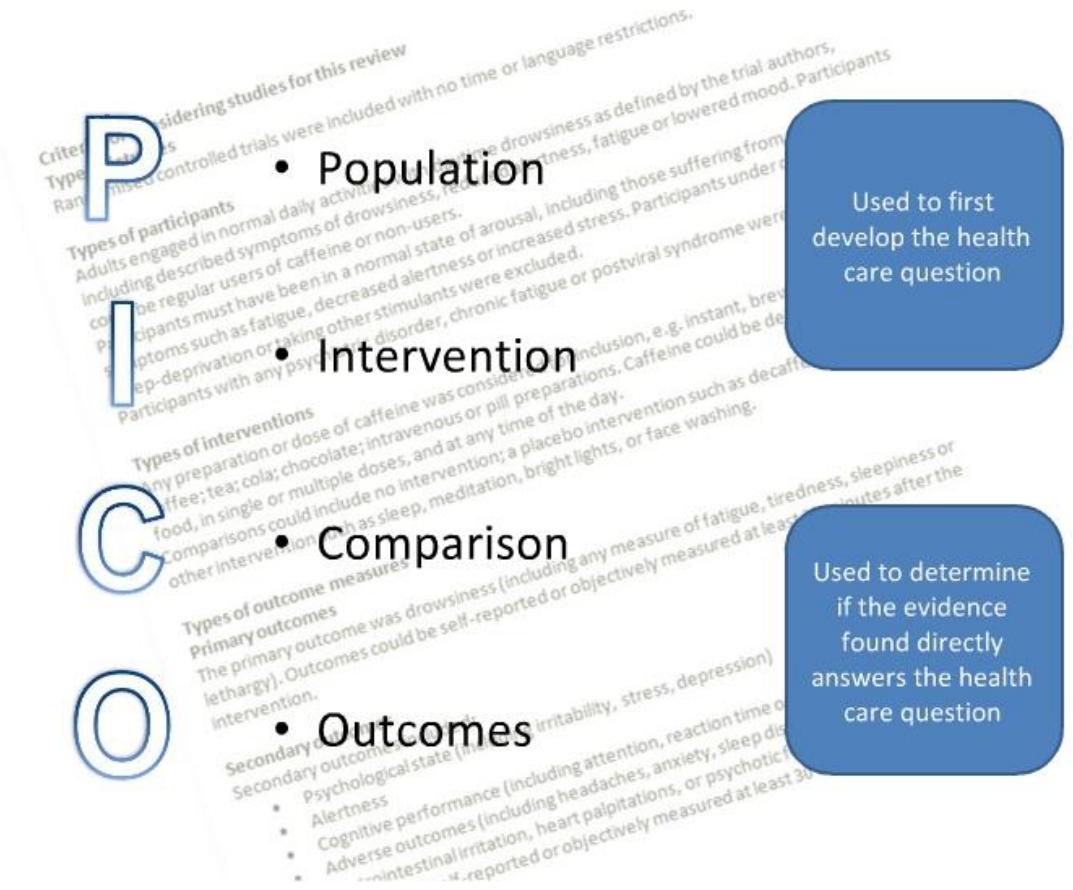
Si chiede quindi alla Letteratura l'evidenza quanto più attinente al quesito clinico:

ER+, HER2-, *de novo*, MBC

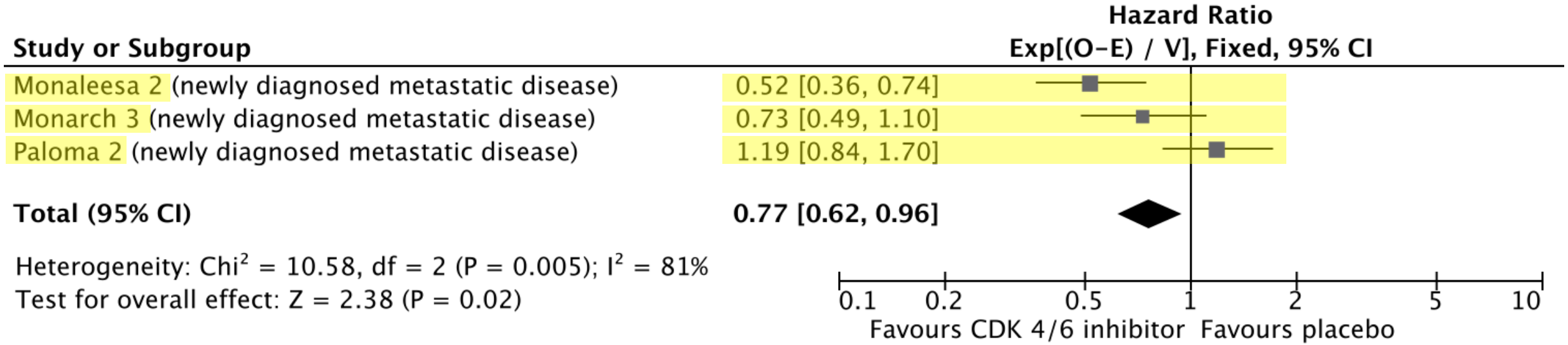
CDK 4/6 inhibitor + AI

AI alone

Overall Survival



Overall Survival – *de novo* metastatic disease



Hortobagyi GN, et al. N Engl J Med 2022;386:942-50 - Goetz MP, et al. Annals of Oncology (2022) 33 (suppl_7): S808-S869. 10.1016/annonc/annonc1089 - Finn RS, et al. Journal of Clinical Oncology 40, no. 17_suppl (June 10, 2022) LBA1003-LBA1003.

Outcomes

Should be
importance driven
NOT
evidence driven

SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

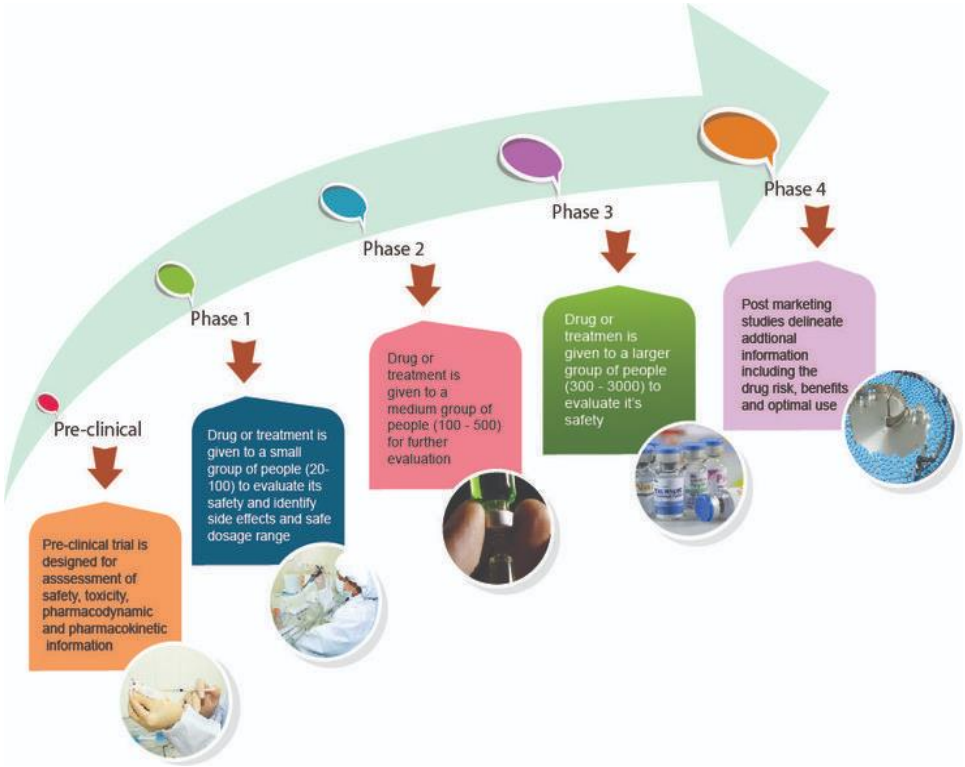
VENERDÌ 26 - SABATO 27 GENNAIO 2024

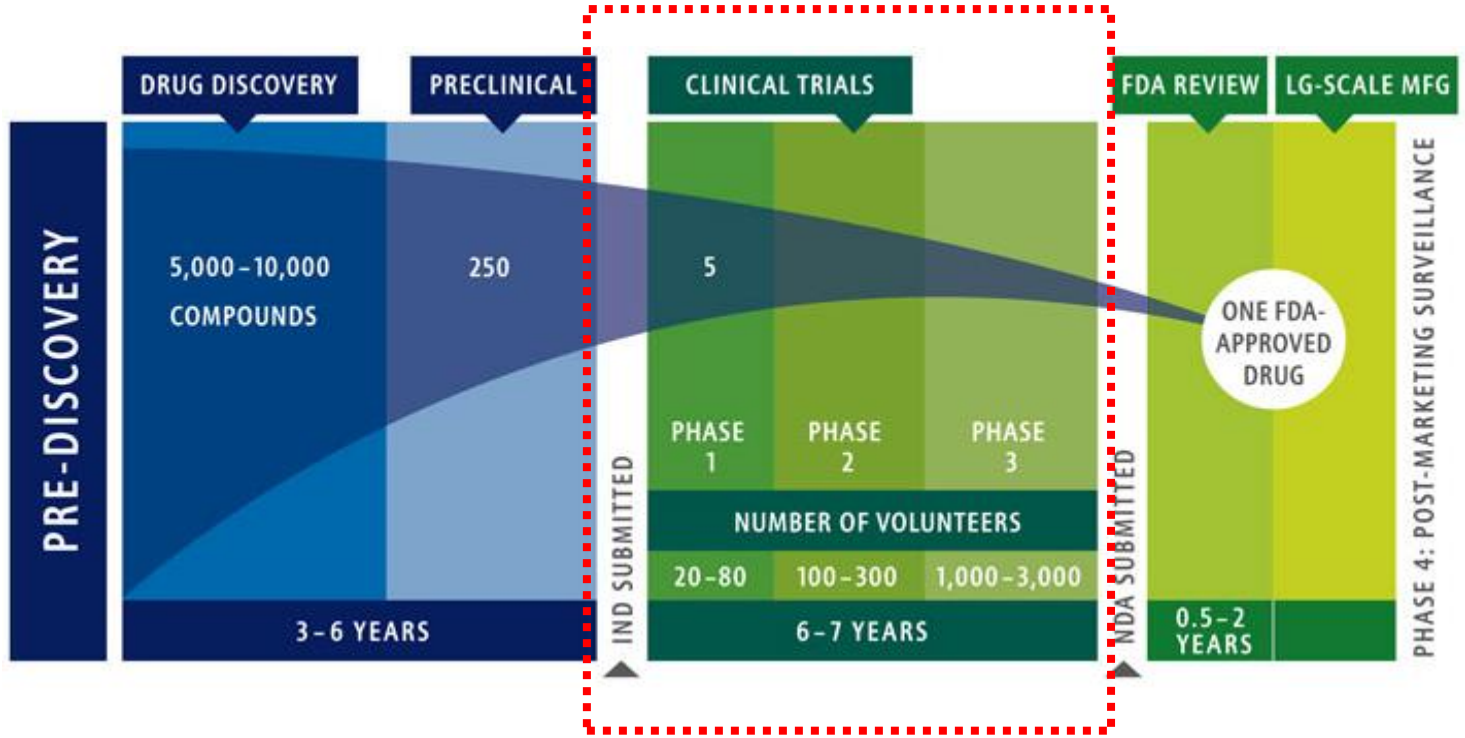
NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

- Etica e integrità della ricerca
- Quesito clinico di riferimento
- **Fasi della sperimentazione clinica**
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

The Long (Risky) Road of Drug Discovery & Development





Source: www.pinterest.it/excaliburhealth

Source: Pharmaceutical Research and Manufacturers of America

Experimental Studies: Phases & Characteristics

Phase	Objectives and Aims	Subjects	Focused on
I	Pharmacokinetics Safety	Healty Subjects Patients	DRUG
II	Activity Safety	Patient	DISEASE
III	Efficacy	Patients	INDIVIDUAL PATIENT

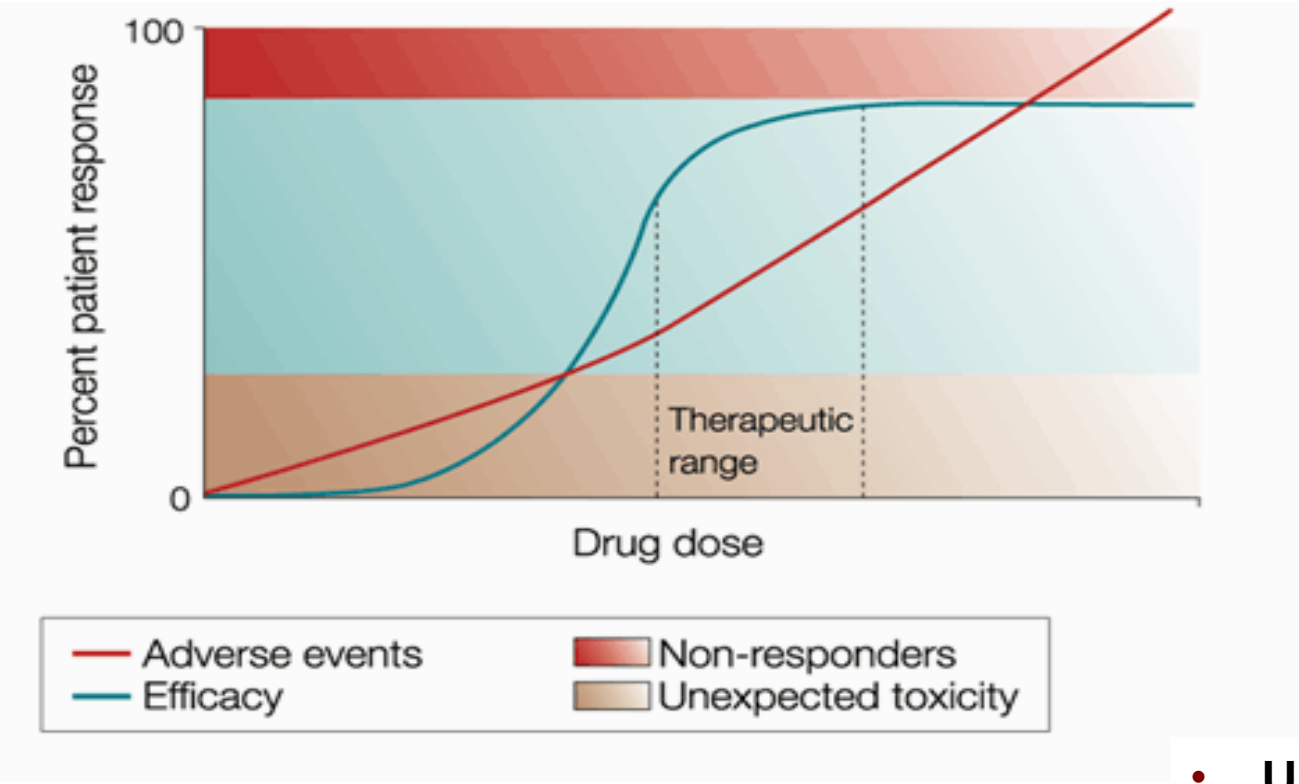
Phase I: *Definition & Endpoints*

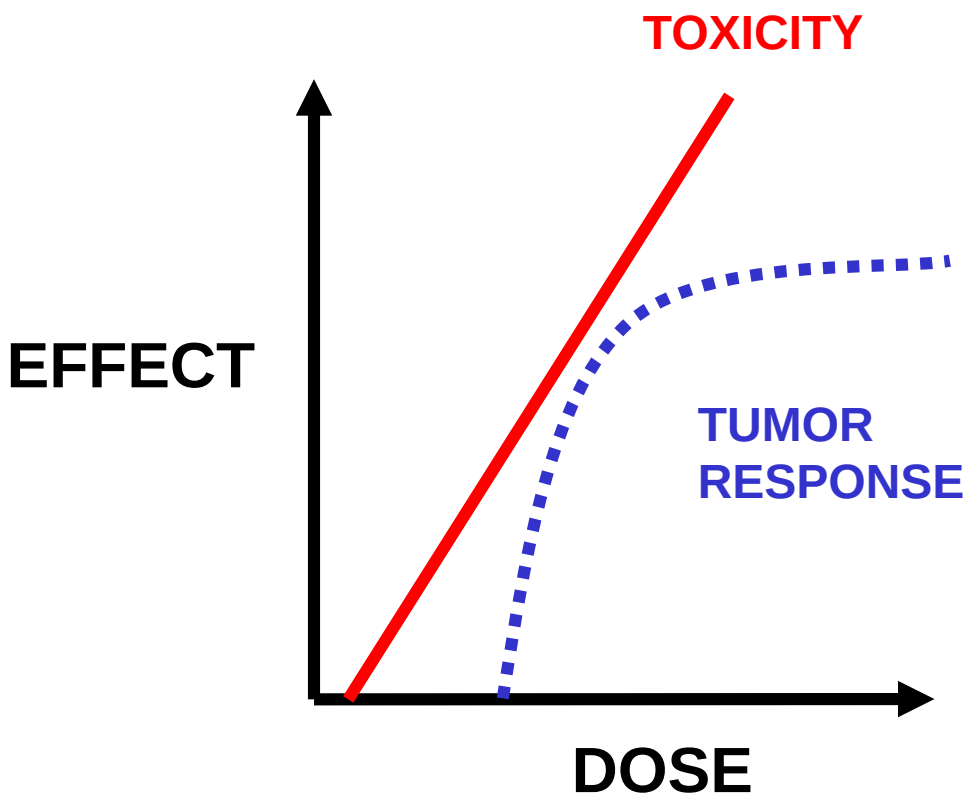
- **1st evaluation of a new cancer therapy in humans**
  - *Dose-escalation studies*
  - First-in-human single agent study
  - Combination of novel (or approved) agents
  - Combination of novel agent and radiation therapy
- **Which agents deserve to enter Phase I?**
  - It is *biologically plausible* that the agent may have activity in cancer (target seems valid and agent affects it)
  - Preclinical or other evidence of efficacy
  - Reasonable expectation of safety (toxicology)
  - Sufficient data on which to base a starting dose

- **Primary**
  - **Identify the maximally tolerated dose (MTD) and recommended Phase II dose (RP2D)**
    - ***MTD = level @ DLT (in Europe or Japan)***
    - ***MTD = level below DLT (in US)***
  - **Identify dose-limiting toxicities (DLTs)**
    - ***Unacceptable toxicities for severity/duration, which limit further dose escalation***
- **Secondary**
  - **Pharmacokinetics (PK)**
  - **Pharmacodynamics (PD)**

Chemotherapy: Dose-Effect model is Reliable

Dose, Efficacy and Toxicity of Chemotherapy





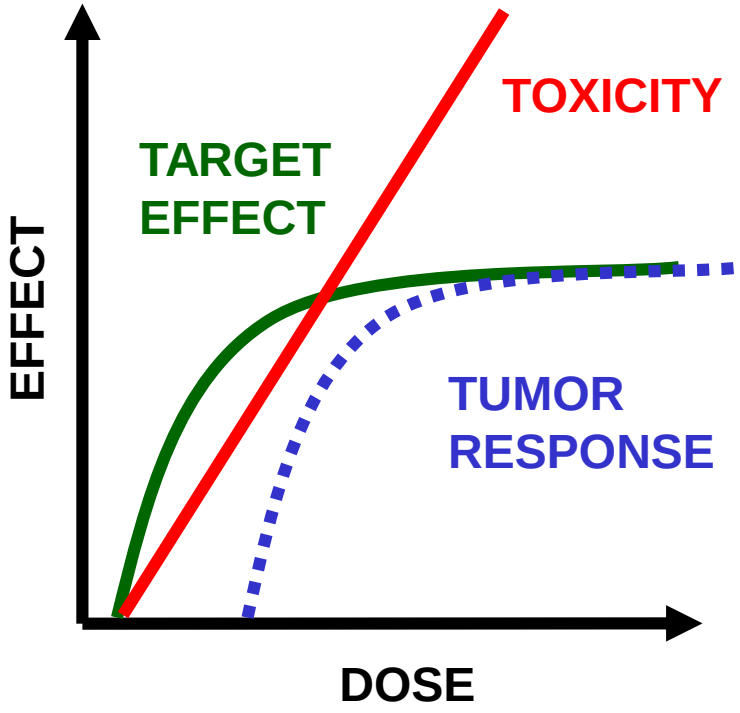
Low therapeutic index (benefits / side effects)

- **Usually have *a direct correlation* between dose and toxicity**
- **Over a threshold, *toxicity limits activity***
- **Phase I aim to find out MTD (maximum tolerated dose) as conducive to the highest efficacy**

Relling and Dervieux, Nature Reviews, 2001

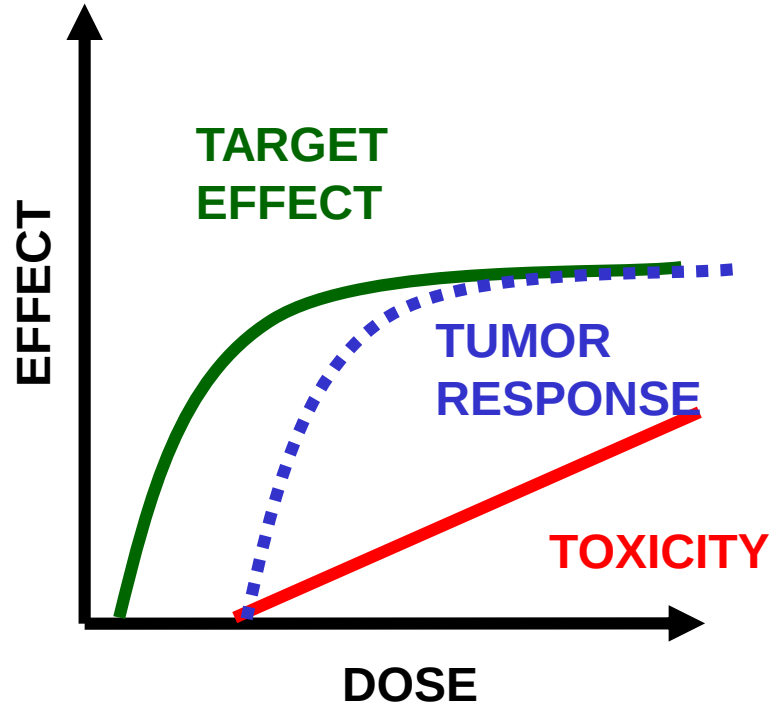
Targeted Agents: Is Dose-Effect model Reliable?

<u>Aim of Phase I</u>: MTD or MTID (minimum target-inhibiting dose)?



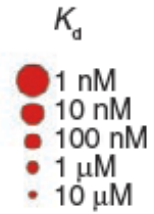
Specificity: **<u>LOW</u>**

Clinical Indication: **<u>No</u>**



<u>HIGH</u>

<u>Yes</u>

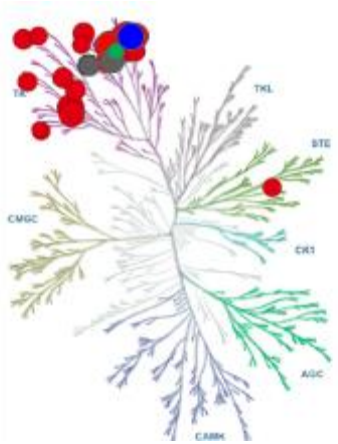


Modified Kinomes from: Karaman MW et al, Nature Biotechnology 2008

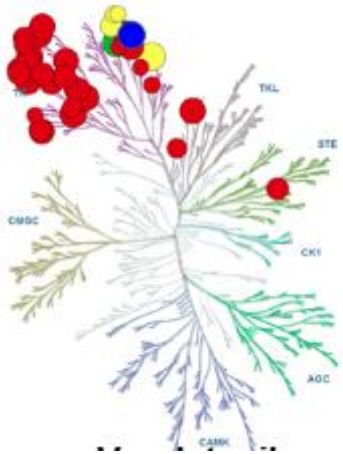
More Selective *RET* Multikinase Inhibitors for *RET*-Rearranged NSCLC

- RET
- KDR/VEGFR2
- FGFR1-3/EGFR
- MET/ALK/ROS
- Other kinases

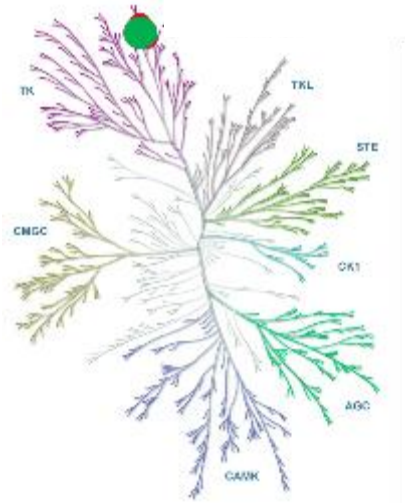
Cabozantinib



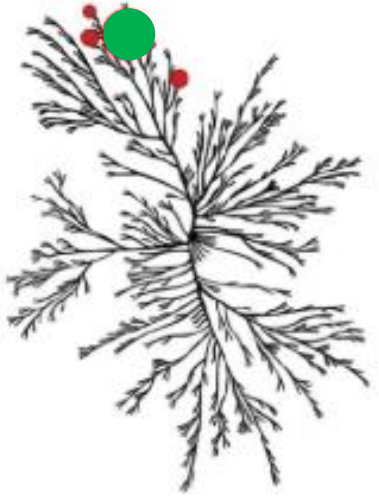
Vandetanib



Selpercatinib (LOXO-292)



Pralsetinib (BLU-667)



Agent	Cabozantinib	Vandetanib	Selpercatinib (LOXO-292)	Pralsetinib (BLU-667)
IC_{50} RET, nM*	11	4	3	0.4
ORR, %	37	18	68	58
▪ CR	5	0	2	1

*Cell free.

Velcheti. WCLC 2017. Abstr OA 12.07. Gautschi. JCO. 2017;35:1403. Drilon. WCLC 2019. Abstr PL02.08. Gainor. ASCO 2019. Abstr 9008. Rahal. AACR 2017. Abstr B151. Solomon. J Thorac Oncol. 2020;15:541.

Phase I Study Basic Principles & Dose Escalation Methods

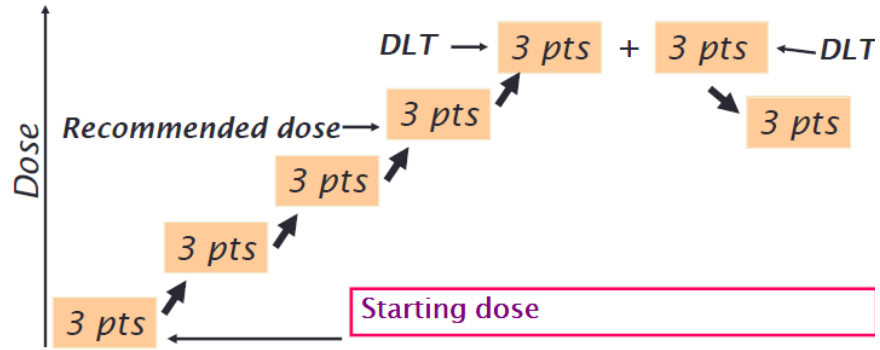
- Define a recommended dose (***dose-escalation context***):
  - **SAFELY** (minimum # of serious toxicities)
  - **EFFICIENTLY** (smallest possible # of pts)
  - **RELIABLY** (high statistical confidence)
  - Start with a safe dose
    - *1/10th of the LD10 in rodents, or 1/3rd of the minimal toxic dose in large animals expressed as mg/m^2*
  - Minimize # of pts treated at sub-toxic/therapeutic doses
  - Escalate dose rapidly in the absence of toxicity
  - Escalate dose slowly in the presence of toxicity
  - Expand patient cohort at recommended phase II dose

- **Rule-Based Designs**
  - **Traditional 3+3 Design**
  - **Accelerated Titration Design**
  - **Pharmacologically Guided Dose Escalation**
- **Model-Based Designs**
  - **Modified Continual Assessment Method**

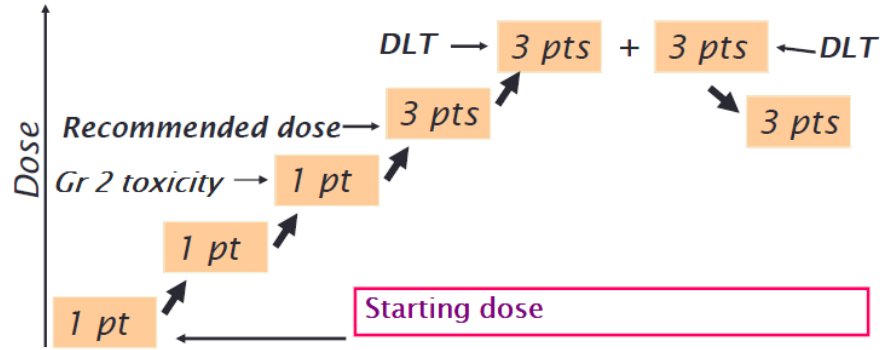
"Escalation in decreasing steps",
L. Fibonacci, 13th century

Le Tourneau C, et al. JNCI 2009

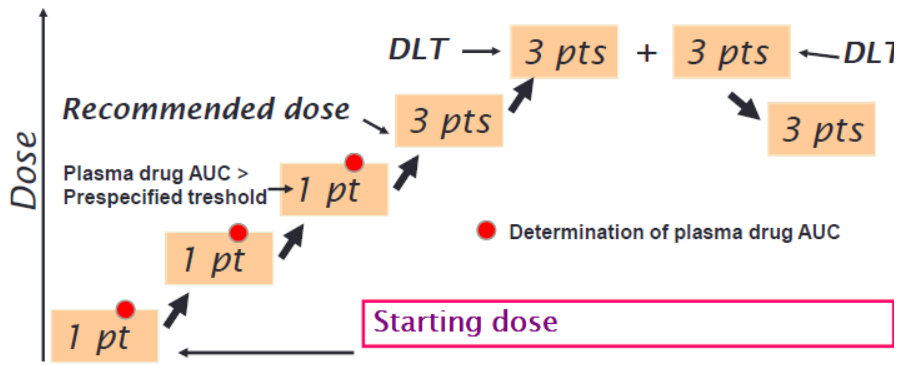
Phase I "Standard" 3 + 3 Design



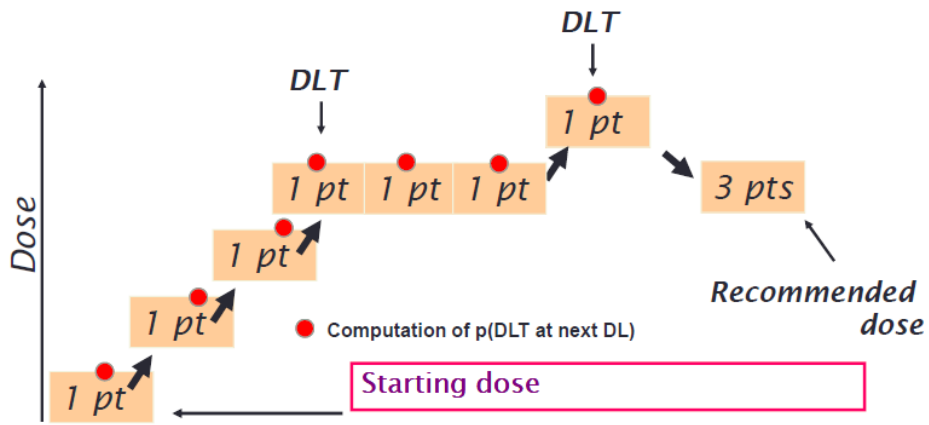
Accelerated Titrated Design



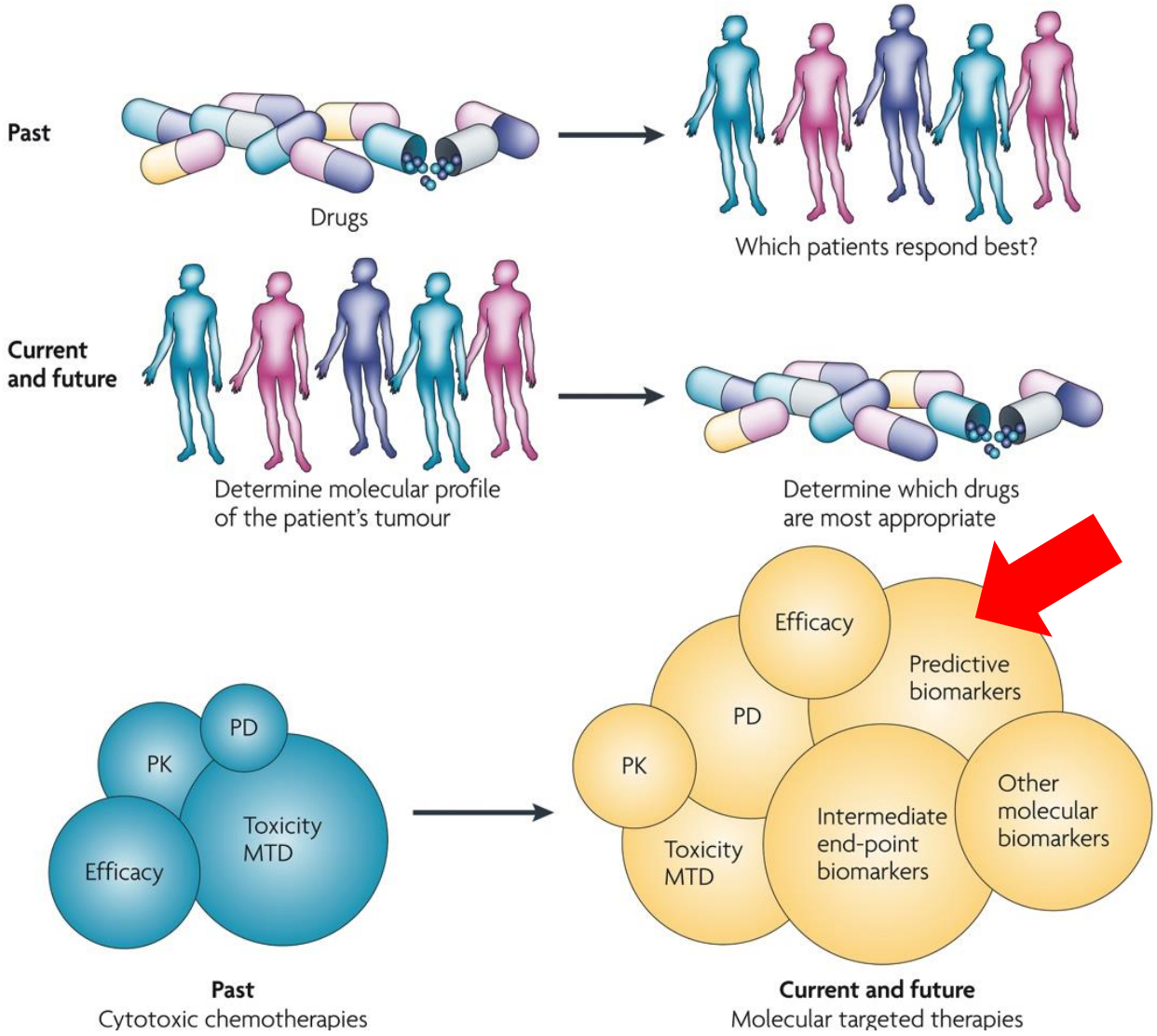
Pharmacologically Guided Dose Escalation



Model-Based Phase I Design: Modified Continual Assessment Method



Phase Is: A Changing Paradigm Overtime



Yap TA, et al. Nat Rev Cancer 2010

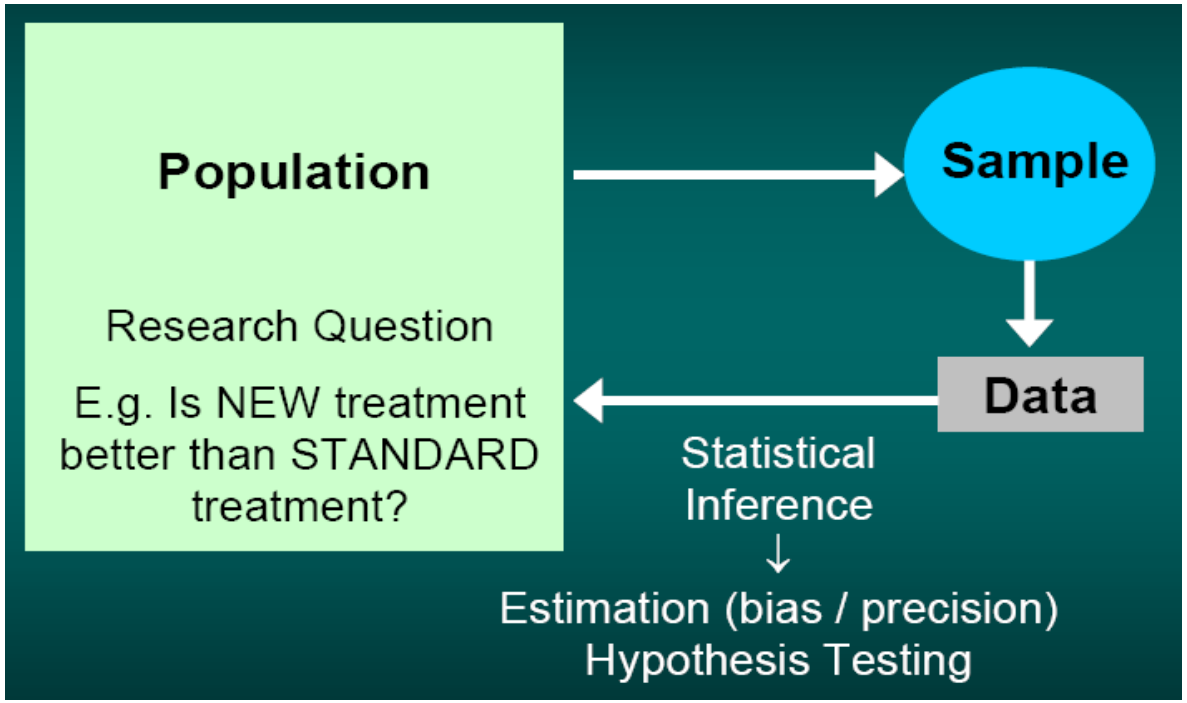
Predictive Biomarkers

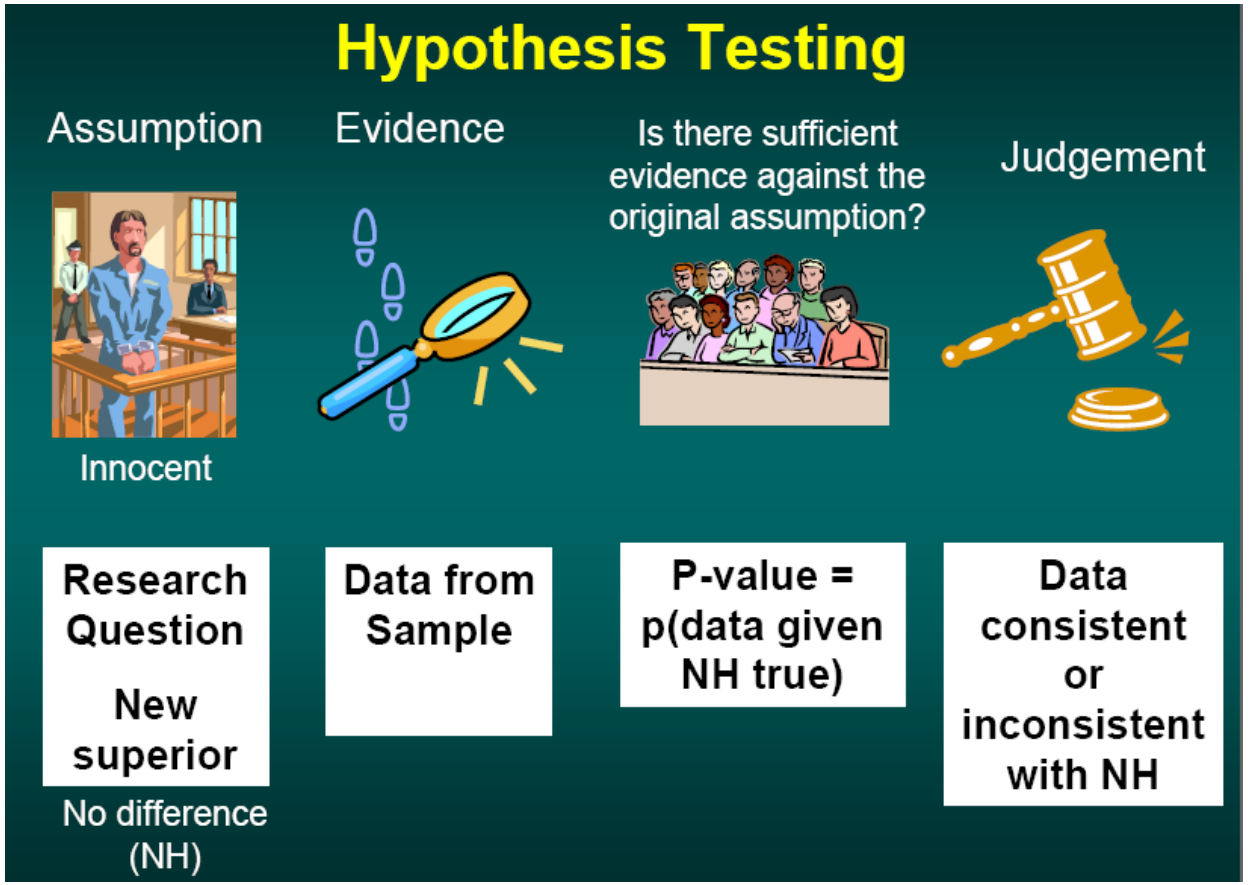
- **Together with corresponding PK data should confirm target modulation and help to identify the *biologically active dose range***
- **Advantages**
  - **Decrease the number of patients receiving ineffective treatments**
  - **Minimize the need for retrospective subgroup analysis in later phase trials**
- **Drawbacks**
  - **Might not be applicable to broad-spectrum inhibitors**
  - **Regulatory issues**
  - **Difficulties in recruiting**
  - **Potential benefit in unselected population could be missed - prevalence**

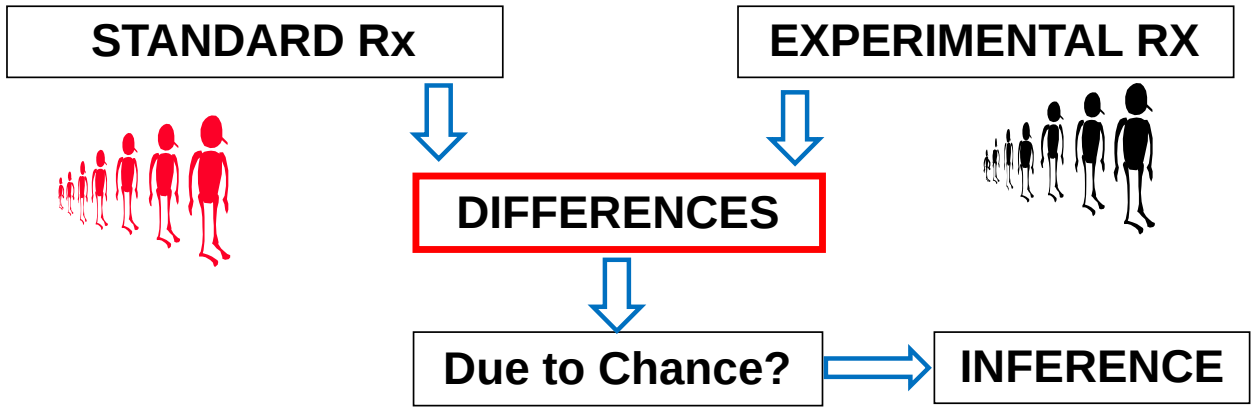
Experimental Studies: Phases & Characteristics

Phase	Objectives and Aims	Subjects	Focused on
I	Pharmacokinetics Safety	Healty Subjects Patients	DRUG
II	**Activity** **Safety**	**Patients**	**DISEASE**
III	Efficacy	Patients	INDIVIDUAL PATIENT

Why Do We Need Statistics?







Aim of Statistical Analysis: Search for the Truth!

Courtesy of Billingham C, 2008

What do we assess in clinical trials?

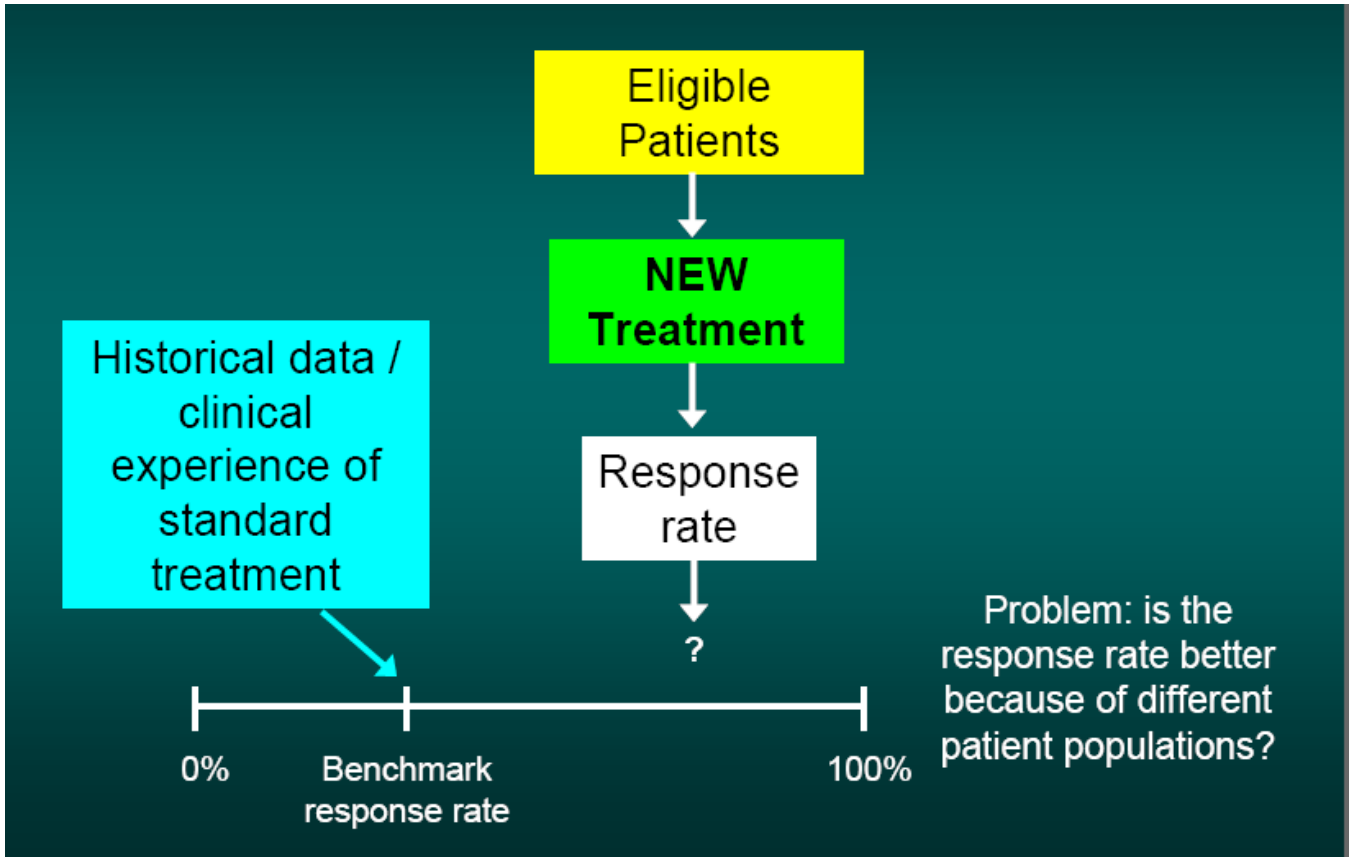
- **Activity:**
  - ability of the treatment to induce modifications of the disease thanks to which it is assumed that the patient may have a benefit **[Phase II]**
- **Efficacy:**
  - ability of the treatment to induce a clinical benefit in patients who are administered ***in an experimental context*** **[Phase III]**
- **Effectiveness:**
  - ability of a treatment to be effective in a ***non-experimental, concrete and coincident with the clinical practice*** **[Phase IV?]**

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

Courtesy of Pappagallo G.

Single-Arm Phase II: Pros and Cons

Single-Arm Phase II Study



Setting the Bar in Phase II Studies

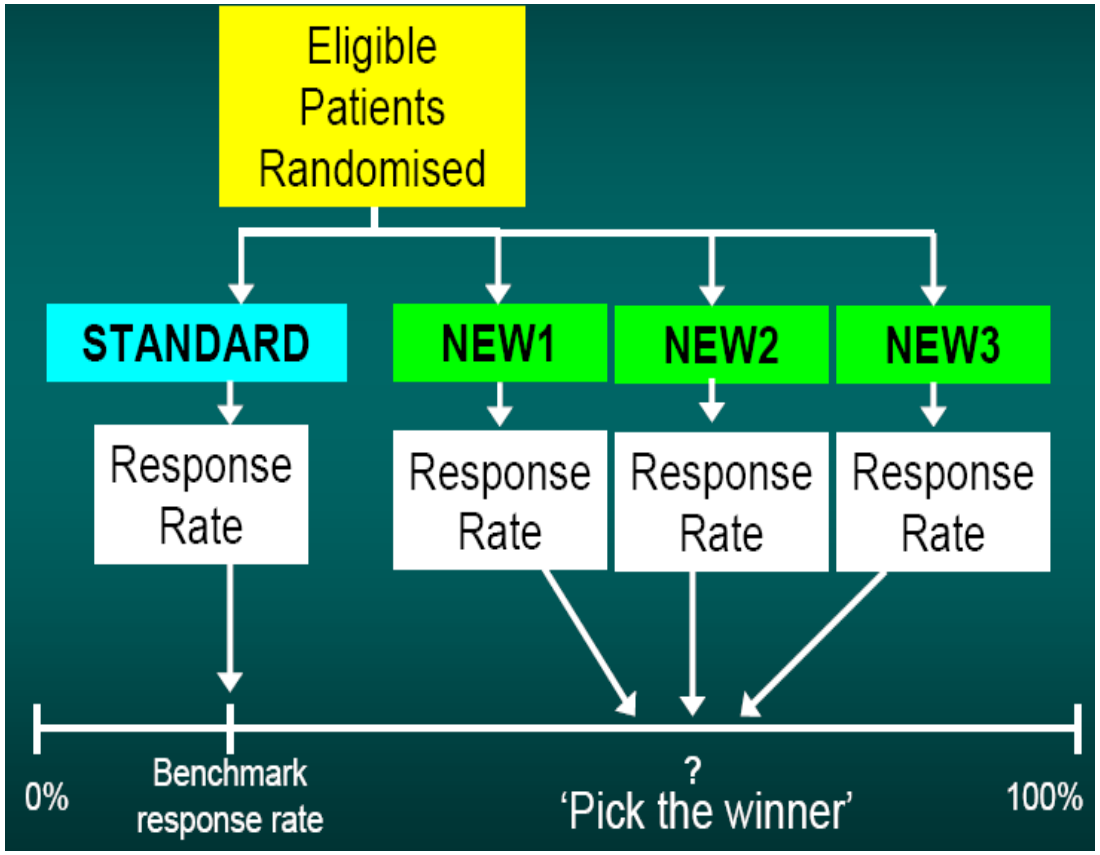
Citation of historical data	Number	Conclusions		Results	
		Unclear	Clear	Reject alternative (agent not worthy of further study)	Reject null (agent worthy of further study)
No historical data cited	32	3	29	6 (21%)	23 (79%)
Historical data cited					
Did not meet criteria	29	2	27	4 (15%)	23 (85%)
Met criteria	9	0	9	6 (67%)	3 (33%)

- **Studies that met the criteria for appropriate citation of prior data were significantly less likely to reject the null (33%) than those cited that did not meet the criteria (85%) *p=0.006***

Courtesy of Billingham C, 2008

Vickers AJ et al, Clin Cancer Res 2007

Pros & Cons (Random. Phase II): Overinterpretation



***Comparative means*:**
The winner enters the Phase III fashion

Pros	Cons
• Control of selection bias • Simultaneous testing of several new treatments, combinations, doses, etc. • Generally preferable: some degree of control is better than none!	• NOT a statistical comparison between randomized groups • The randomized Phase II trials DO NOT replace phase III trials • Over-interpretation of comparative results from small sized randomized Phase II trials

Courtesy of Billingham C, 2008

Dafni U, WCLC 2019

Clinical Trials Methodology: Topics of Phases

Phase	Objectives and Aims	Subjects	Focused on
I	Pharmacokinetics Safety	Healty Subjects Patients	DRUG
II	Activity Safety	Patient	DISEASE
III	**Efficacy**	**Patients**	**INDIVIDUAL PATIENT**

What do we assess in clinical trials?

- **Activity:**
  - ability of the treatment to induce modifications of the disease thanks to which it is assumed that the patient may have a benefit **[Phase II]**
- **Efficacy:**
  - ability of the treatment to induce a clinical benefit in patients who are administered ***in an experimental context*** **[Phase III]**
- **Effectiveness:**
  - ability of a treatment to be effective in a ***non-experimental, concrete and coincident with the clinical practice*** **[Phase IV?]**

- **Phase III clinical trials are the gold standard for evaluating therapeutic interventions.**
- **Randomization:**

  Provides a treatment assignment that is independent of outcome and patient/disease features, thus balancing treatment groups on known and unknown factors associated with outcome.

  The intention-to-treat (ITT) analysis approach is the gold standard for all phase III) randomized, controlled clinical trials: analyzes all patients in the treatment groups as randomized without regard to treatment actually received.

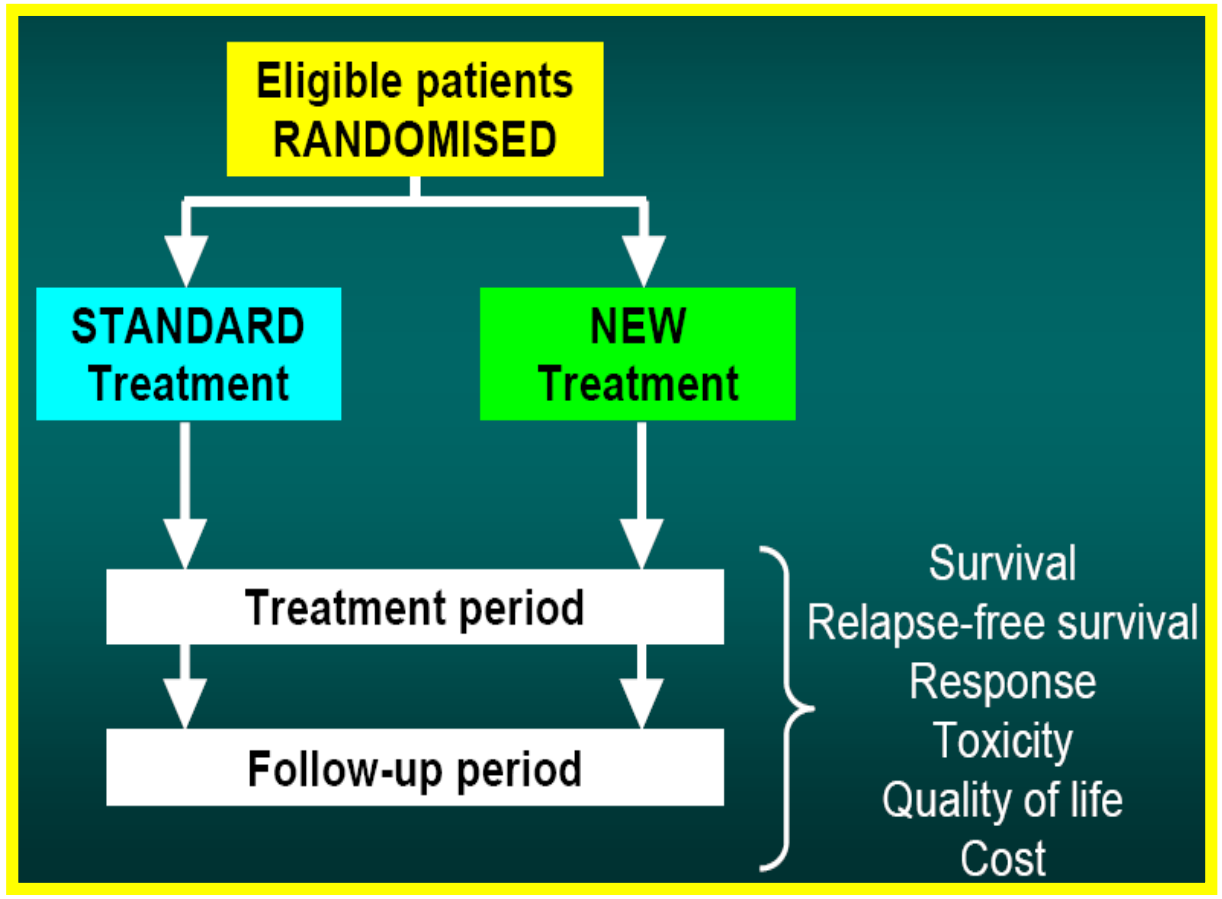
Riwani and Hilsenbeck, J Clin Oncol 2012

Phase III trials: Randomization

- Assignment of subjects to groups compared according to a random sequence
- Why to randomize? To avoid:
  - **Selection bias**
  - **Performance bias**
  - **Detection bias**
  - **Attrition bias**

Jüni P. et al. BMJ 2001

Randomized Phase III Clinical Trials



- **The new drug is more effective than other drugs?**
  - **Superiority trial (is 'statistically' relevant?)**
- **The new drug is as effective as other drugs but with fewer side effects (or minor discomfort or lower costs)?**
  - **Non-inferiority trial**
- **The use of the new drug determines a therapeutic benefit for patients?**
  - **Amount of the benefit (is 'clinically' relevant?)**
- **What categories of patients may derive more benefit from the new drug?**
  - **Subgroups analysis**

Superiority Trials: Principles and ASCO/ESMO Advices

- **Aim:** To demonstrate the superiority of a new therapy compared to an established therapy or placebo
- **Sample Size**: To estimate the sample size one needs to consider
  - the ***Clinical significance (Δ)***: By how much should the new therapy be better than the established?
  - the ***Power (1-β)***: the probability of correctly showing a benefit (usually ≥ 80%)
  - the ***Significance level (α)***: the probability of wrongly concluding that a benefit exists (usually ≤5%)

- **ASCO: Recommended Targets for Meaningful Trial Goals**
- **ESMO: Development of the Magnitude of Clinical Evaluation Scale**

- ***Ellis et al. 2013. JCO. American Society of Clinical Oncology Perspective: Raising the Bar for Clinical Trials by Defining Clinically Meaningful Outcomes***
- ***ASCO 2014, 2015, ESMO 2014, ECCO 2015; Annals of Oncology 2015***

Modified by Dafni U, WCLC 2019

Recommended Targets for Meaningful Trial Goals

Table 1. Summary of Recommended Targets for Meaningful Clinical Trial Goals

			Primary End Point		Secondary End Point	
Cancer Type	Patient Population	Current Baseline Median OS (months)	Improvement Over Current OS That Would Be Clinically Meaningful (months)	Target HRs	Improvement in 1-Year Survival Rate (%)*	Improvement in PFS (months)
Pancreatic cancer	FOLFIRINOX-eligible patients	10 to 11[19]	4 to 5	0.67 to 0.69	48 → 63	4 to 5
Pancreatic cancer	Gemcitabine or gemcitabine/nab-paclitaxel–eligible patients	8 to 9[20,21]	3 to 4	0.6 to 0.75	35 → 50	3 to 4
Lung cancer	Nonsquamous cell carcinoma	13[22]	3.25 to 4	0.76 to 0.8	53 → 61	4
Lung cancer	Squamous cell carcinoma	10[23]	2.5 to 3	0.77 to 0.8	44 → 53	3
Breast cancer	Metastatic triple negative, previously untreated for metastatic disease	18[24,25]	4.5 to 6	0.75 to 0.8	63 → 71	4
Colon cancer	Disease progression with all prior therapies (or not a candidate for standard second- or third-line options)	4 to 6[26]	3 to 5	0.67 to 0.67	25 → 35	3 to 5

Abbreviations: FOLFIRINOX, leucovorin, fluorouracil, irinotecan, and oxaliplatin; HR, hazard ratio; OS, overall survival; PFS, progression-free survival.
*Current → target.

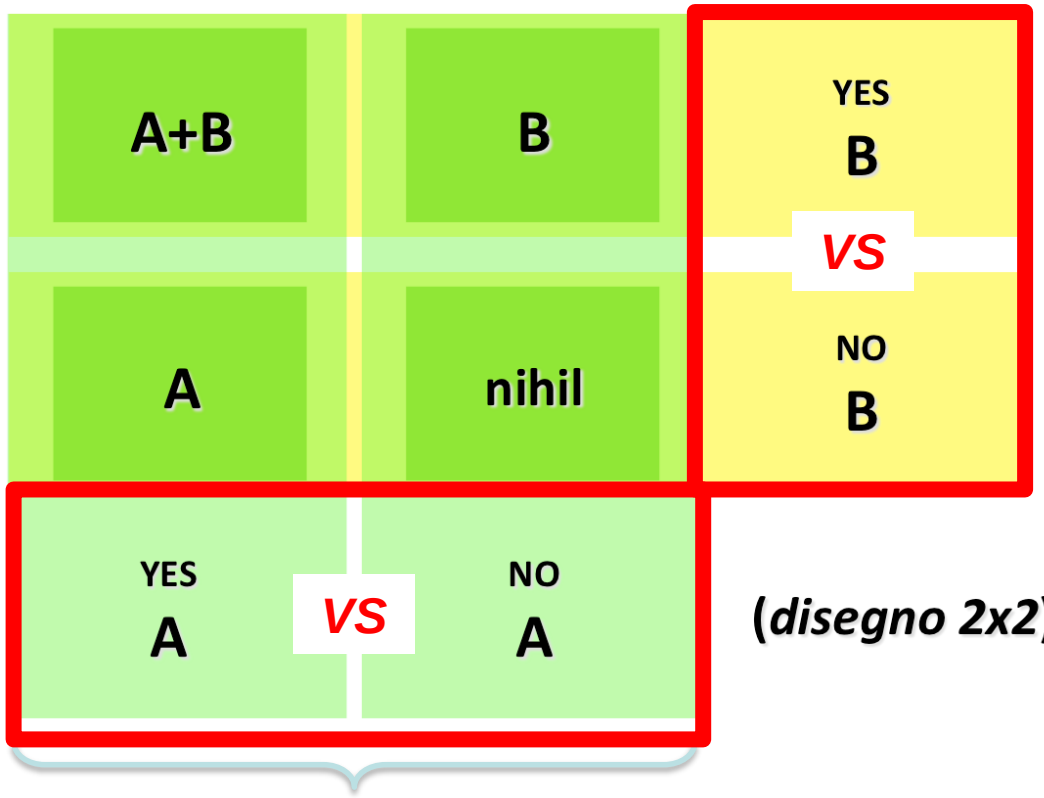
- HR ≤ 0.8, with an improvement in median OS 2.5 to 6 months:
- Minimum incremental improvement over standard therapy that would define a clinically meaningful outcome.

Ellis et al. 2013. JCO. American Society of Clinical Oncology Perspective:

Raising the Bar for Clinical Trials by Defining Clinically Meaningful Outcomes

Dafni U, WCLC 2019

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")

Factorial Design for RCTs

- The sample size of the factorial design is based on a non-interaction assumption between the two aimed questions.
- The application of factorial design allows two independent questions to be answered using the same patients:
  - In other words, it is a simple way to conduct two trials in one.
- If a significant interaction is expected, or it emerges during the trial (ex. other evidence), it should be kept in mind that its occurrence could make the primary result of the trial practically useless, and the unique useful information could come from secondary, less powered analyses.

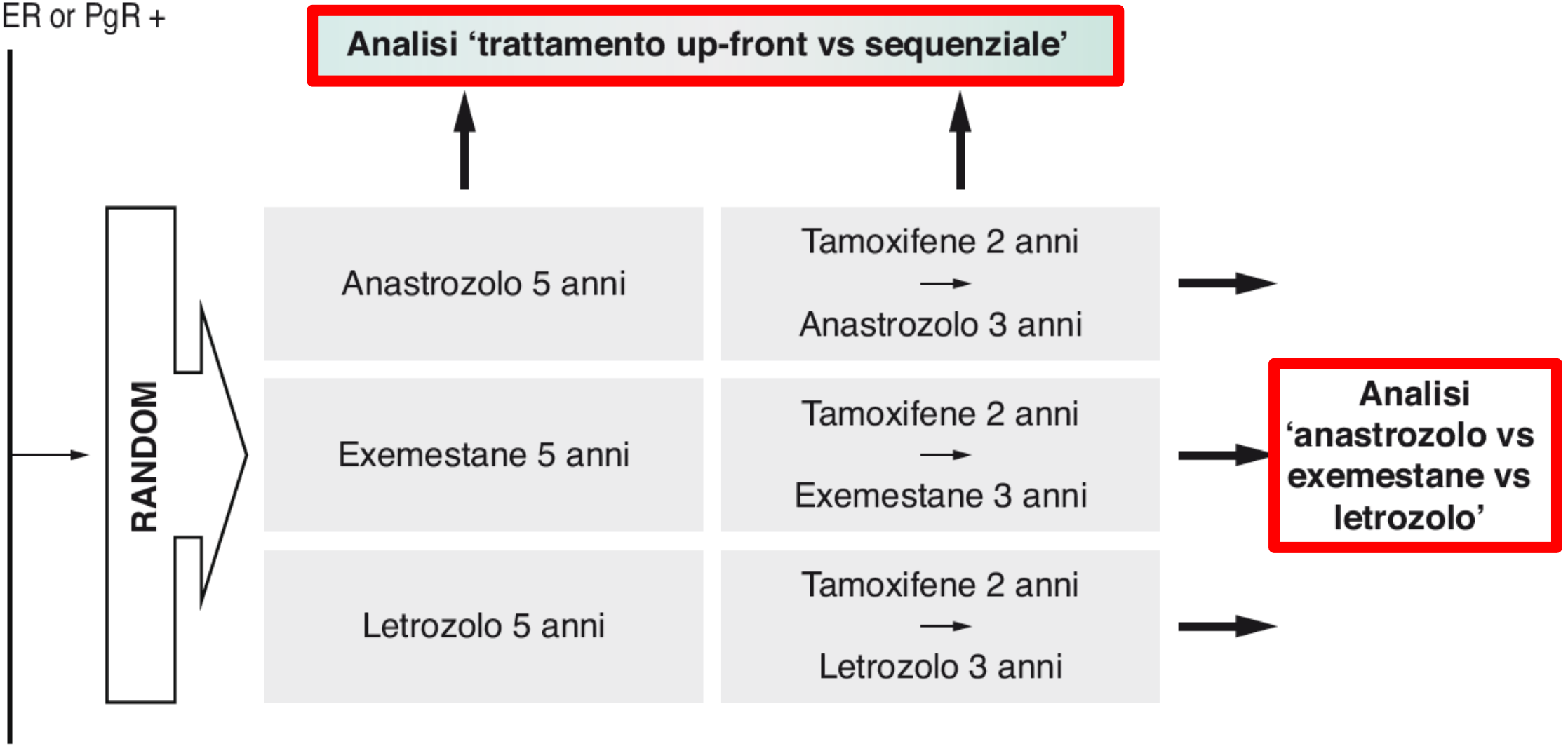
Bria E et al, Ann Oncol 2016

Gruppo Italiano Mammella (GIM) Studies

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

Pazienti

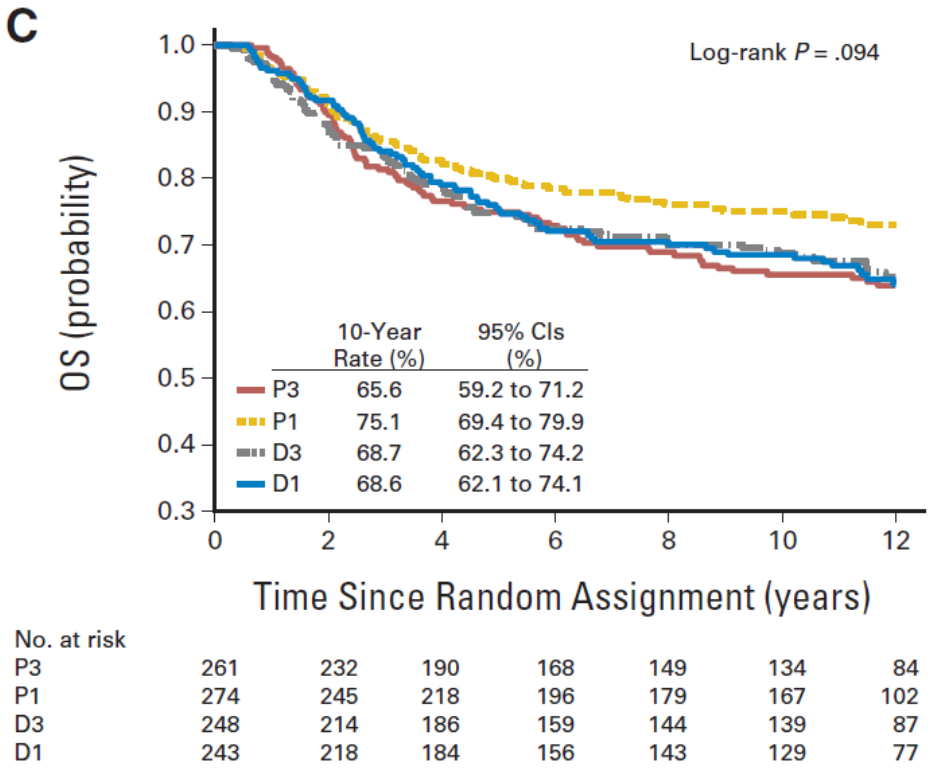
- Carcinoma mammario operato
- Postmenopausa
- ER or PgR +



Paclitaxel in TNBC

Weekly Paclitaxel	Weekly Docetaxel	WEEKLY
		vs
3-Weekly Paclitaxel	3-Weekly Docetaxel	3-WEEKLY
PACLITAXEL *vs*	DOCETAXEL	

E1199 Trial, TNBC subgroup (!!!!)



Sparano JA et al, J Clin Oncol 2015

Factorial Design for RCTs: SPIRIT Guidelines

- **Standard Protocol Items Recommendations for Interventional Trials (SPIRIT) Statement**:
  - Guidance on reporting of trial protocols, for factorial trials
  - Delphy survey with 104 panelists from 14 countries
  - Hybrid consensus with 15 panelists finalizing the statment
  - 33 Items (9 modified this year)
  - Developed for increasing the trials'utility and transparency

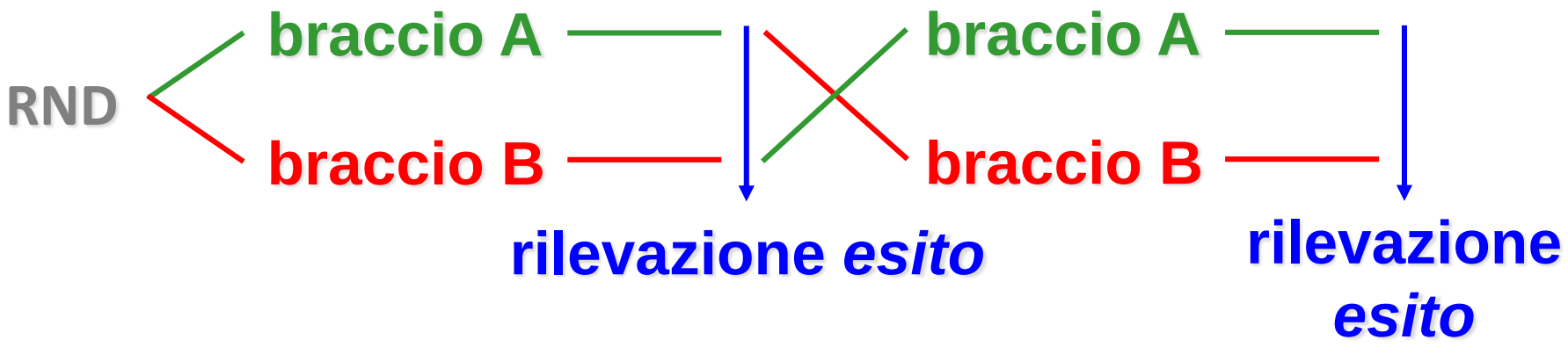
- **Factorial trials offer a powerful means of efficiently evaluating multiple treatments within a single study.**
- **Their methodological intricacies demand transparent reporting of rationale, assumptions, and analytical approaches.**
- **The extension of SPIRIT guidance represents a step toward enhancing the quality of evidence produced.**
- ***Transparency in trial design, robust assessment of interactions, careful sample size determination, and comprehensive reporting practices* are vital to harnessing the full potential of factorial trials and advancing evidence-based health care and policy decisions.**

Chan AW et al, Ann Intern Med 2013

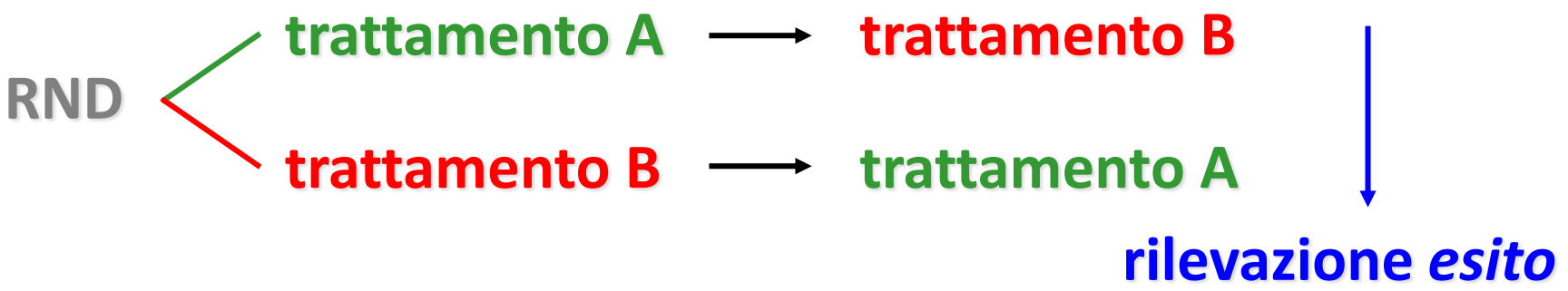
Kahan BC et al, JAMA Network 2023

Mastrantoni L et al, JAMA Network 2023

Attenzione A Non Confondere Un Disegno Crossover…



…Con Un Disegno A Bracci Paralleli Di Tipo Sequenziale

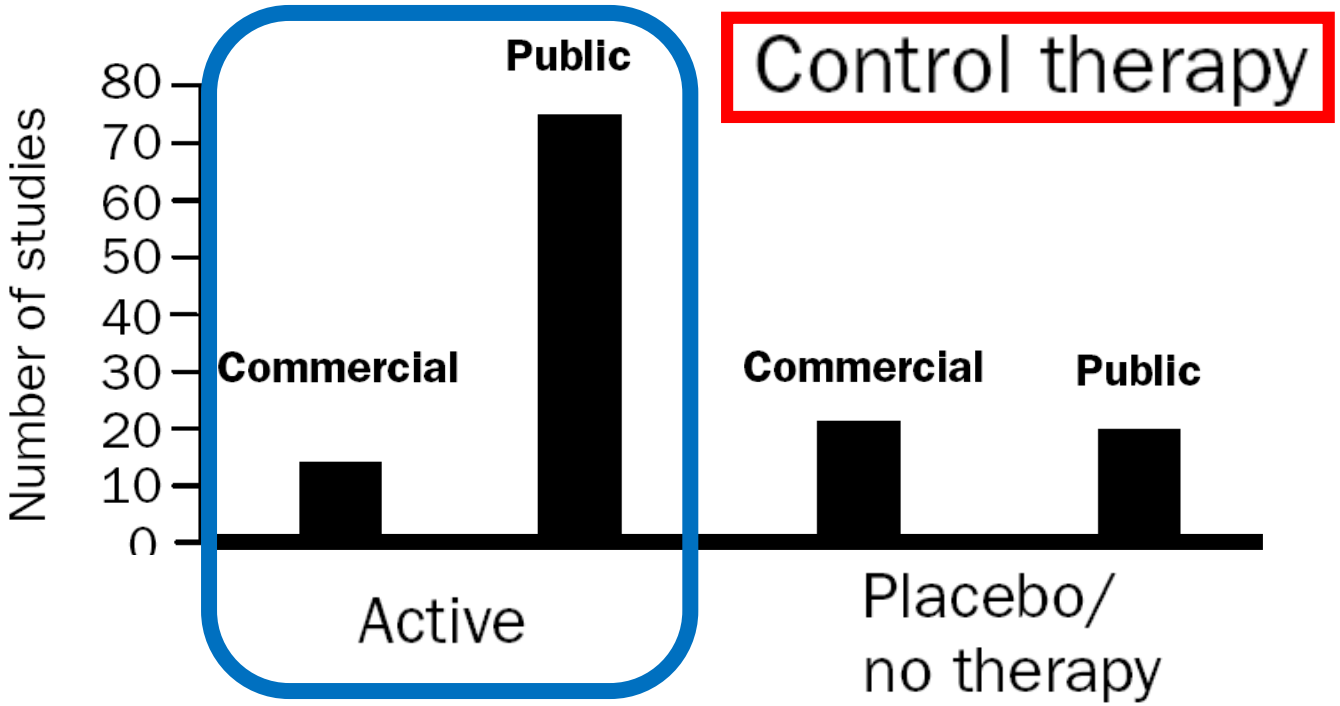


Do not forget the Choice/Performance of the Control Arm!

The uncertainty principle and industry-sponsored research

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.

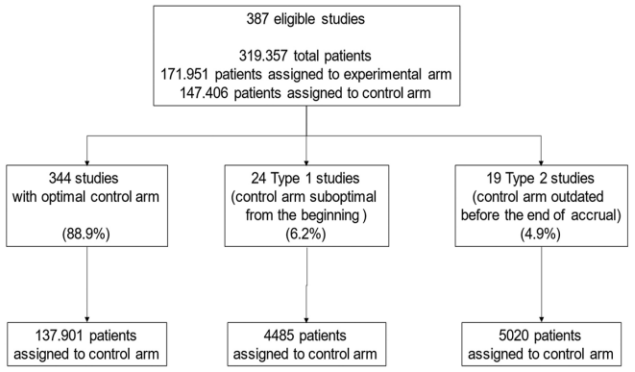


- **Active Control adopted in <20% of COMMERCIAL trials vs >80% of PUBLIC**

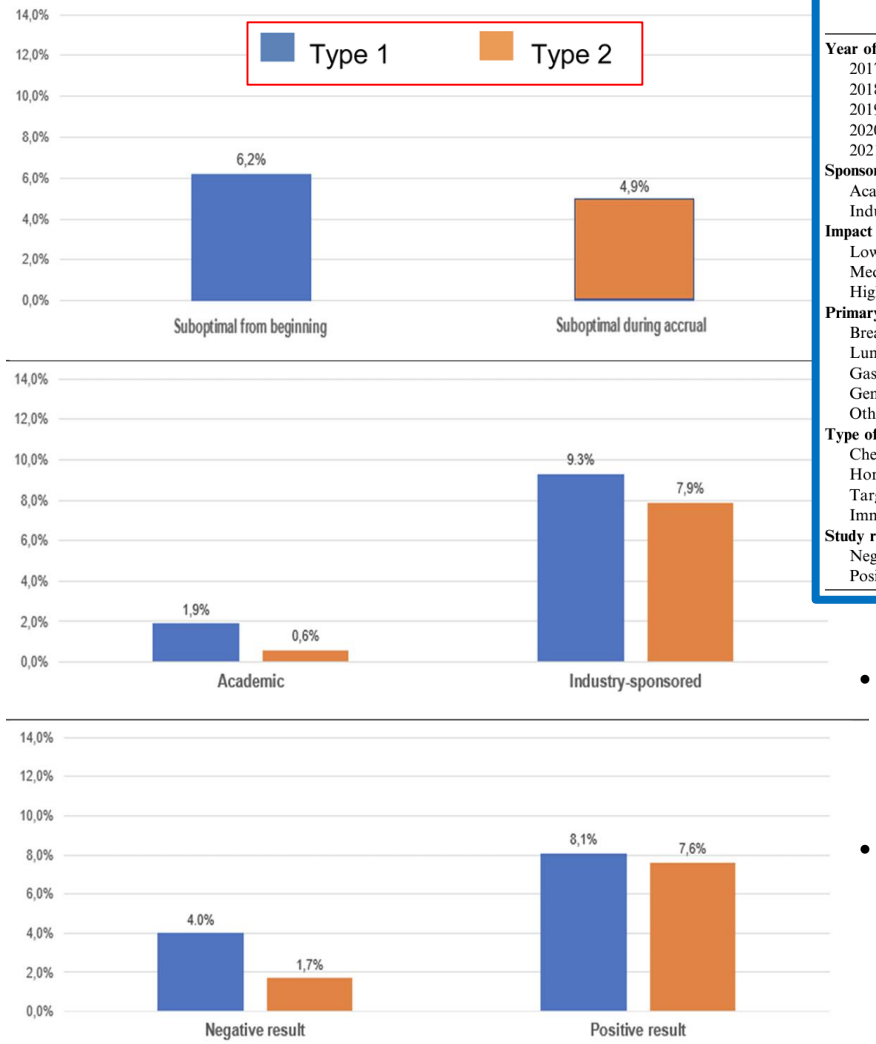
Djulbegovic B et al, Lancet 2000

Do not forget the Choice/Performance of the Control Arm!

Variable	Trials included in the analysis n = 387 (%)
Year of publication:	
2017	100 (25.8%)
2018	77 (19.8%)
2019	67 (17.3%)
2020	63 (16.2%)
2021	80 (20.6%)
Sponsorship:	
Academic	160 (41.3%)
Industry-sponsored	227 (58.6%)
Impact factor:	
Low (< 15)	61 (15.7%)
Medium (15–30)	53 (13.6%)
High (> 30)	273 (70.5%)
Primary site:	
Breast	90 (23.2%)
Lung	67 (17.3%)
Gastrointestinal	87 (22.4%)
Genitourinary	52 (13.4%)
Others	91 (23.5%)
Type of treatment:	
Chemotherapy	108 (27.9%)
Hormonal therapy	27 (7.0%)
Target therapy	166 (42.9%)
Immunotherapy	86 (22.2%)
Study result:	
Negative	177 (45.7%)
Positive	210 (54.3%)



- **Type 1: Studies with Control Arm SUBOPTIMAL FROM THE BEGINNING**
- **Type 2: Studies with Control Arm OUTDATED BEFORE THE END OF THE ACCRUAL**



Variable	Trials with optimal control arm n = 344 (88.9%)	Type 1 studies (control arm suboptimal from the beginning) n = 24 (6.2%)	Type 2 studies (control arm outdated before the end of accrual) n = 19 (4.9%)	P value
Year of publication:				Type 1 p = 0.74
2017	91 (91.0%)	6 (6.0%)	3 (3.0%)	Type 2 p = 0.23
2018	69 (89.6%)	4 (5.2%)	4 (5.2%)	
2019	59 (88.1%)	5 (7.5%)	3 (4.5%)	
2020	57 (90.5%)	3 (4.8%)	3 (4.8%)	
2021	68 (85.0%)	6 (7.5%)	6 (7.5%)	
Sponsorship:				**Type 1 p = 0.003**
Academic	156 (97.5%)	3 (1.9%)	1 (0.6%)	**Type 2 p = 0.001**
Industry-sponsored	188 (82.8%)	21 (9.3%)	18 (7.9%)	
Impact factor:				Type 1 p = 0.10
Low (< 15)	59 (96.7%)	1 (1.6%)	1 (1.6%)	Type 2 p = 0.08
Medium (15–30)	49 (92.5%)	3 (5.7%)	1 (1.9%)	
High (> 30)	236 (86.4%)	20 (7.3%)	17 (6.2%)	
Primary site:				Type 1 p = 0.77
Breast	82 (91.1%)	5 (5.6%)	3 (3.3%)	Type 2 p = 0.09
Lung	56 (83.6%)	5 (7.5%)	6 (9.0%)	
Gastrointestinal	78 (89.7%)	5 (5.7%)	4 (4.6%)	
Genitourinary	42 (80.8%)	5 (9.6%)	5 (9.6%)	
Others	86 (94.5%)	4 (4.4%)	1 (1.1%)	
Type of treatment:				Type 1 p = 0.06
Chemotherapy	104 (96.3%)	3 (2.8%)	1 (0.9%)	Type 2 p = 0.06
Hormonal therapy	26 (96.3%)	0	1 (3.7%)	
Target therapy	141 (84.9%)	16 (9.6%)	9 (5.4%)	
Immunotherapy	73 (84.9%)	5 (5.8%)	8 (9.3%)	
Study result:				Type 1 p = 0.09
Negative	167 (94.4%)	7 (4.0%)	3 (1.7%)	**Type 2 p = 0.007**
Positive	177 (84.3%)	17 (8.1%)	16 (7.6%)	

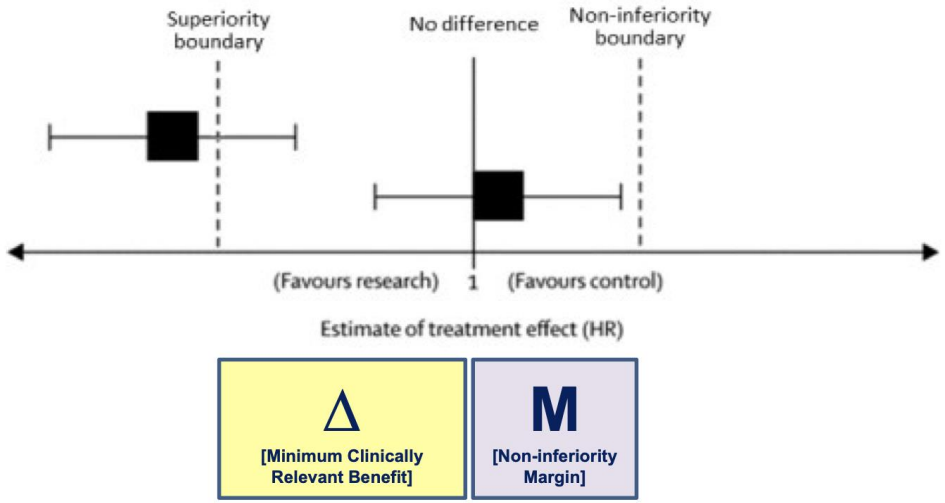
- **Many trials have SUBOPTIMAL CONTROL ARMS, even in journals with high IF, leading to suboptimal treatment of control patients and biased evaluation of trial results.**
- **In some cases, our daily clinical practice is influenced by RCTs with SUBOPTIMAL CONTROL ARMS, with rates that are not negligible.**

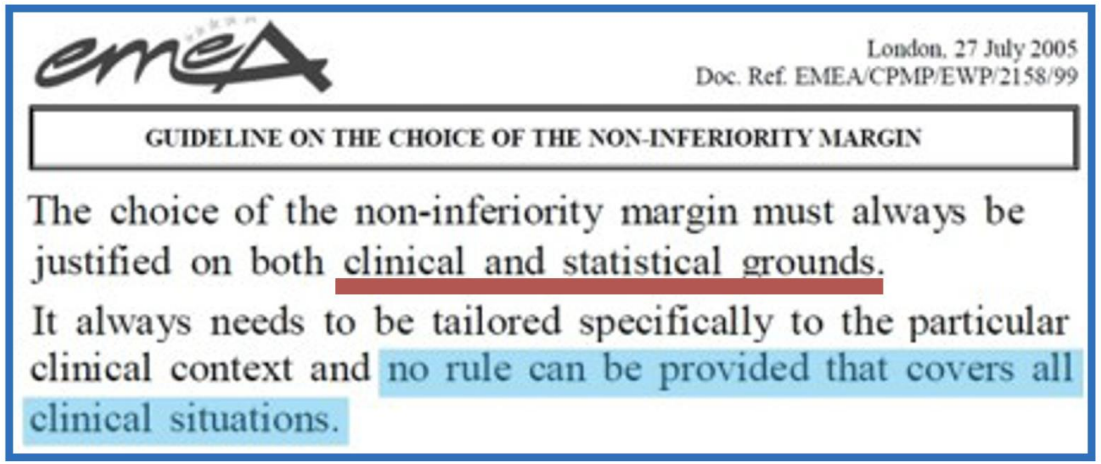
Rossi A et al, EJC 2023

Sizing a trial (A [Exp.] *vs.* B [Control]) For…

- **Superiority**:
  - **A** is better than **B** in terms of the primary outcome, and ***this expected difference must be clinically relevant [Minimum Clinically Relevant Benefit: ∆]***
- **Non-inferiority**:
  - **A** is ***inferior*** to **B** in terms of the primary outcome, but ***not lower than a clinically pre-specified 'margin' [M] which is considered clinically relevant***.
    - **Ex. You can accept a small degree of lower benefit, if patients experience better tollerability**

Superiority & Non-inferiority





emea

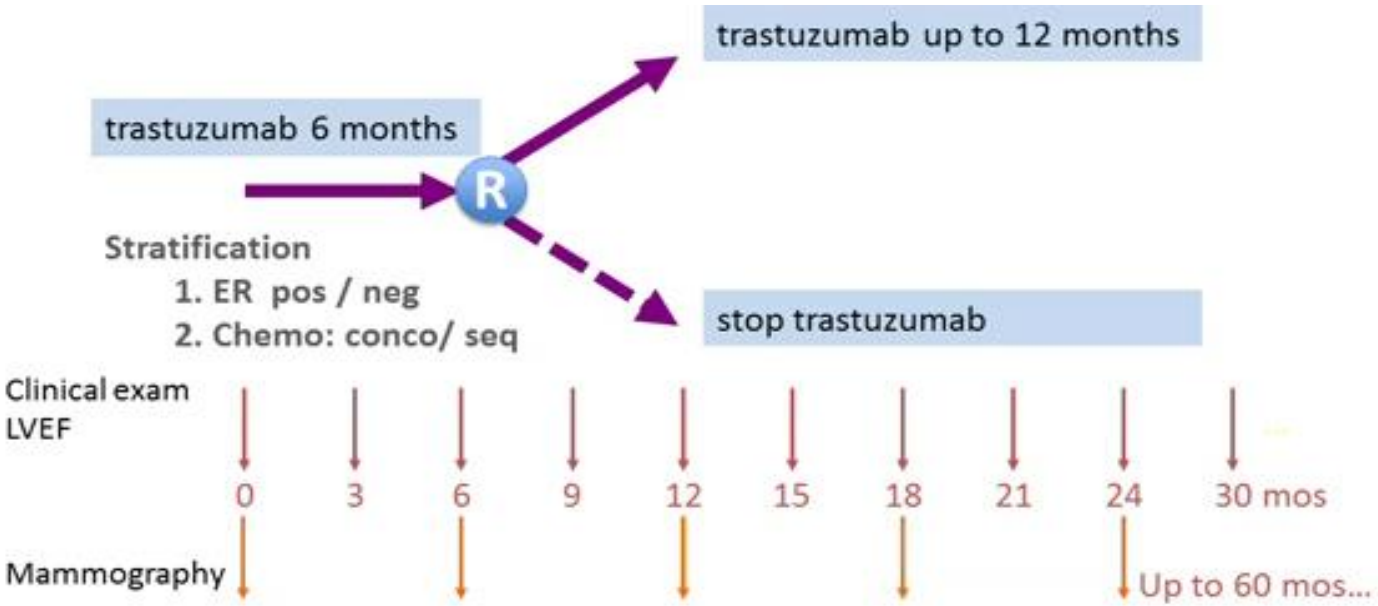
London, 27 July 2005
Doc. Ref. EMEA/CPMP/EWP/2158/99

GUIDELINE ON THE CHOICE OF THE NON-INFERIORITY MARGIN

The choice of the non-inferiority margin must always be justified on both clinical and statistical grounds.

It always needs to be tailored specifically to the particular clinical context and no rule can be provided that covers all clinical situations.

Non-Inferiority in Breast Cancer: Trastuzumab De-Escalation

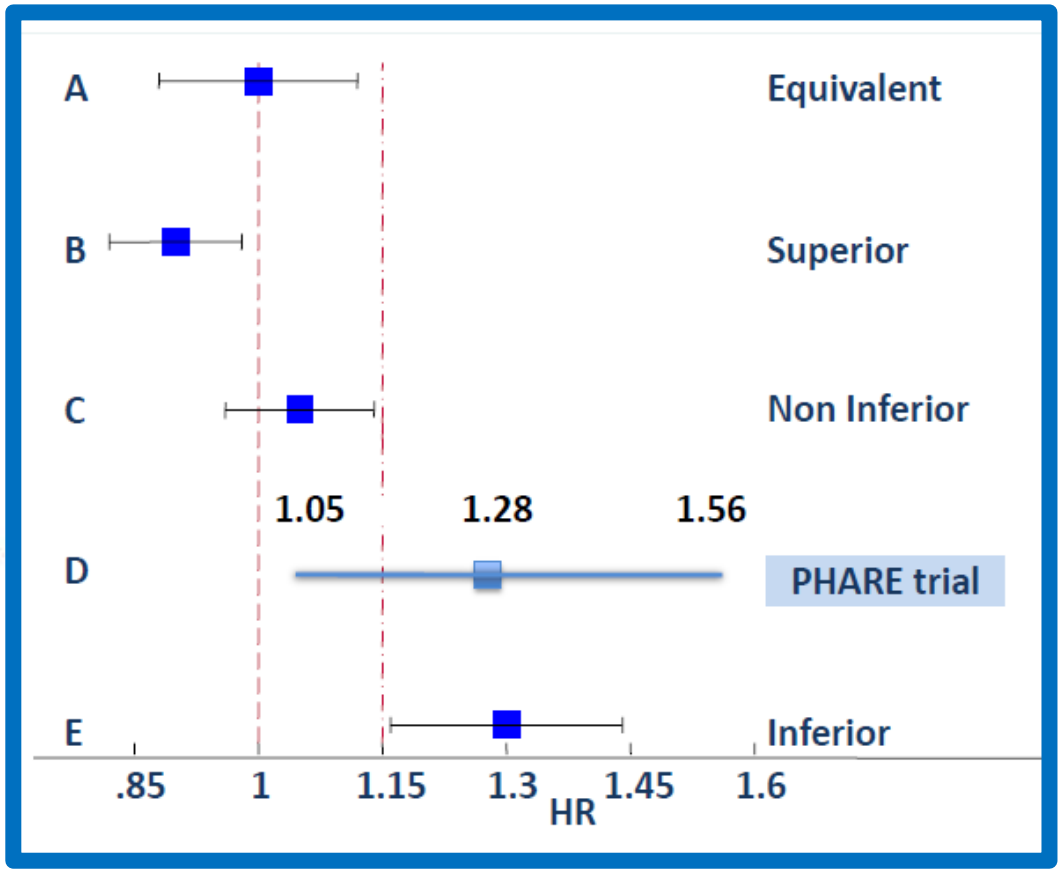


- **Non inferiority randomized trial**
  - 2% variation in terms of absolute difference of recurrence
  - The 95% CI HR margins should not cross the 1.15 boundary
  - 1040 DFS events required for 80% power at 5% level

  or

  4 years of accrual and at least 2 years of follow-up
  - HR were estimated from the stratified Cox model

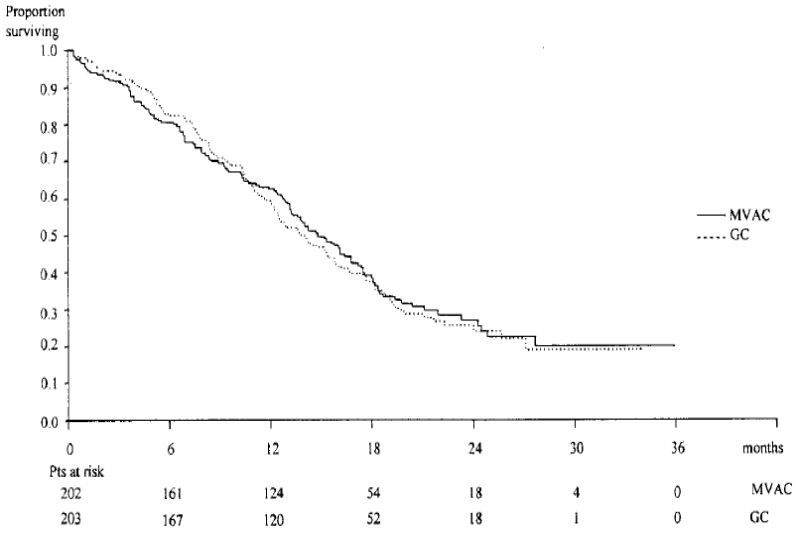
- **Accrual target: 3400 patients**



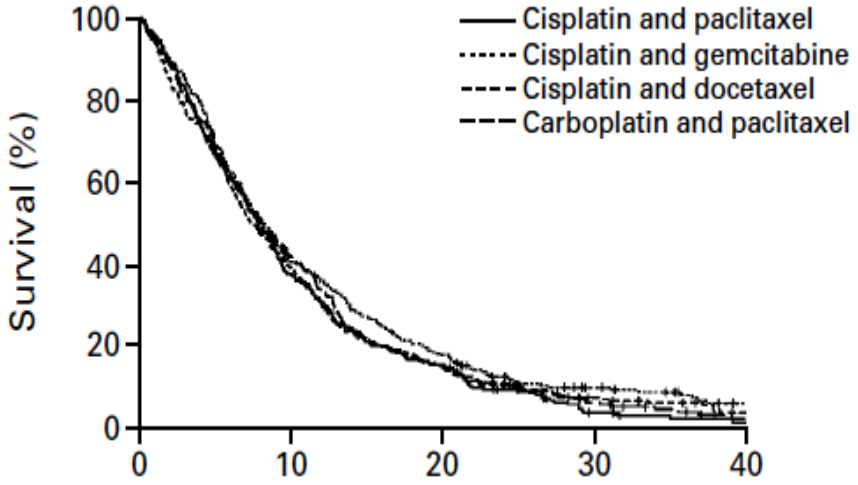
Pivot X et al, ESMO 2012

Key Examples of Mis-intepretation of SUPERIORITY RCTs with Relevant Implication for Clinical Practice

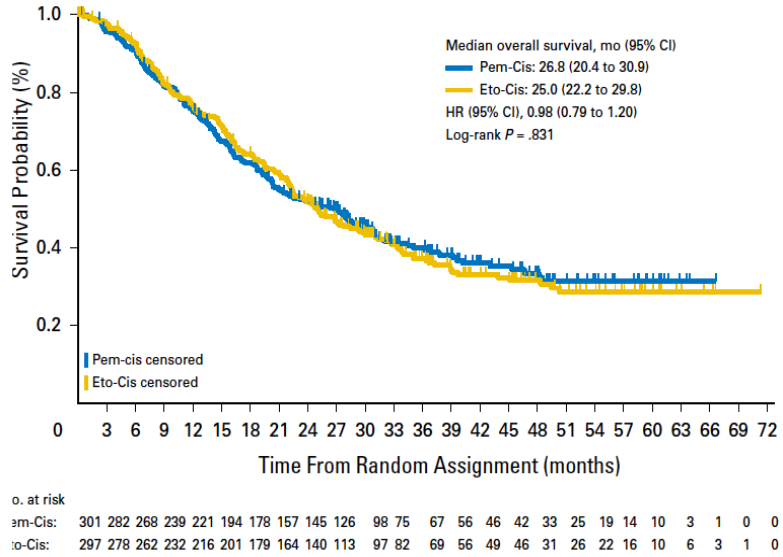
Bladder [Stage IV-Advanced]: Testing SUPERIORITY of CIS-GEM *vs.* MVAC



NSCLC [Stage IIIB-IV - ECOG 1594]: Testing SUPERIORITY of One *vs.* Other(s)



NSCLC [Stage III – PROCLAIM] Testing SUPERIORITY of CIS-PEM *vs.* CIS-VP-16



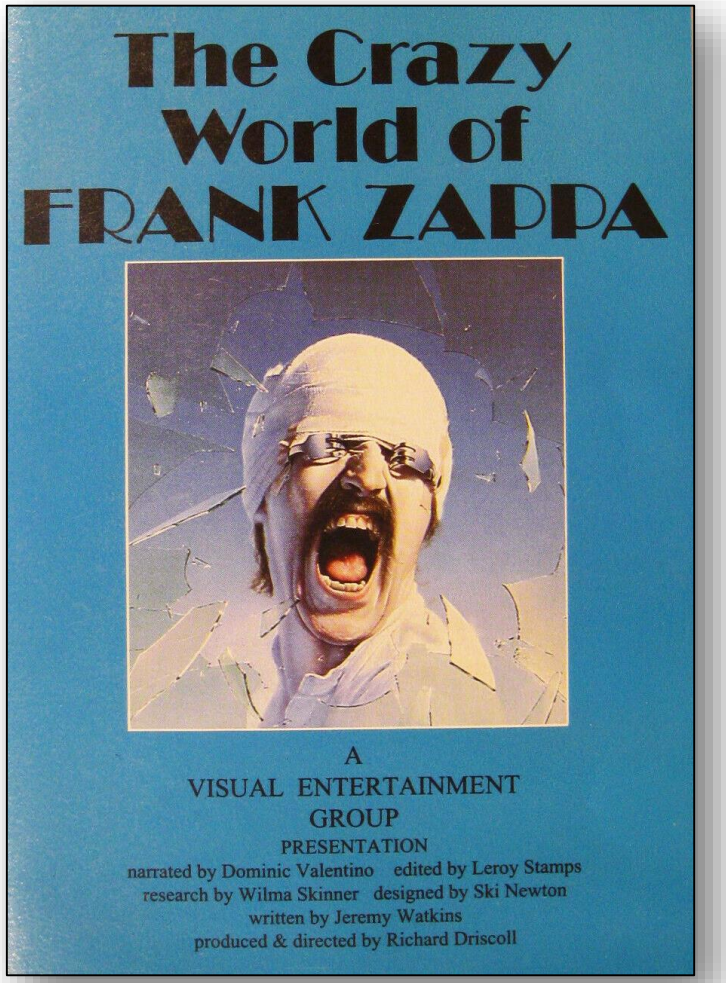
Non-Inferiority & Equivalence MUST be Pre-Specified and Powered! Beware of Wrong Conclusions for Clinical Practice!

Van De Maase H et al, JCO 2000

Schiller J et al, NEJM 2002

Senan S et al, J Clin Oncol 2015

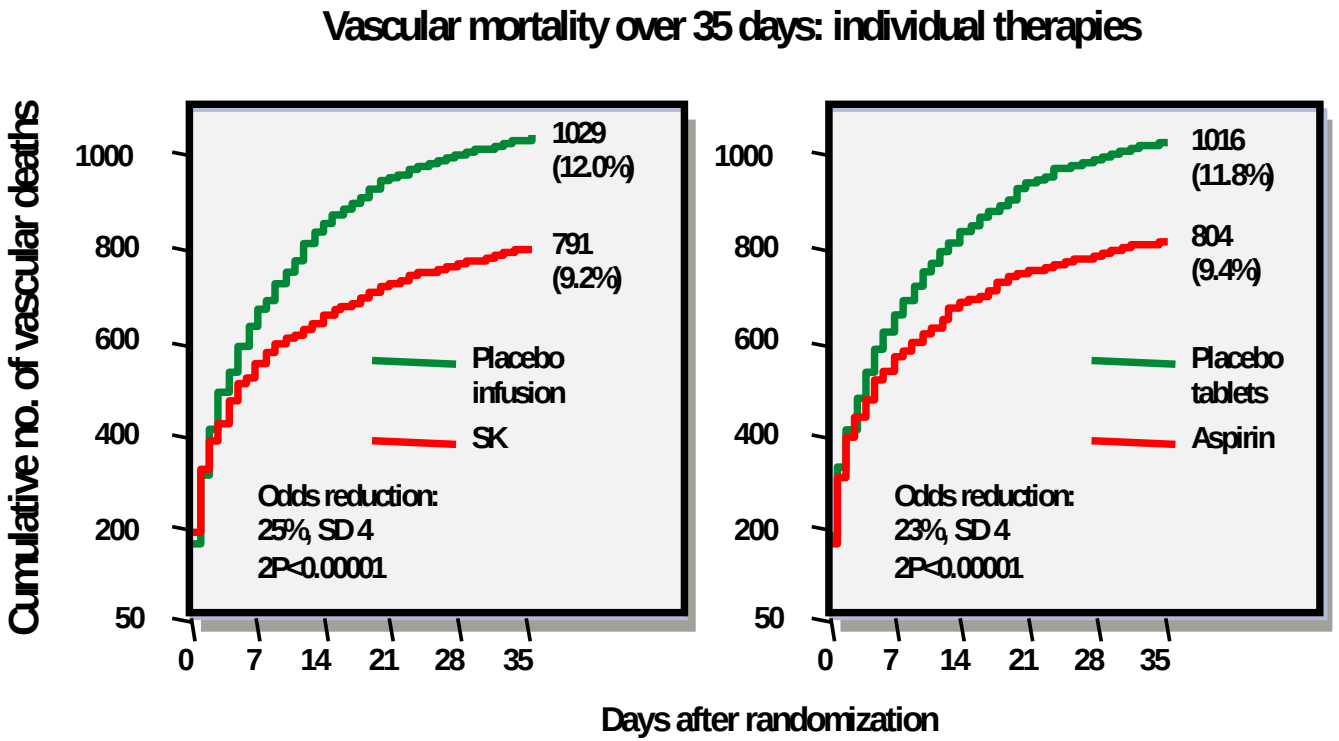
Big Issues (in a given trial)

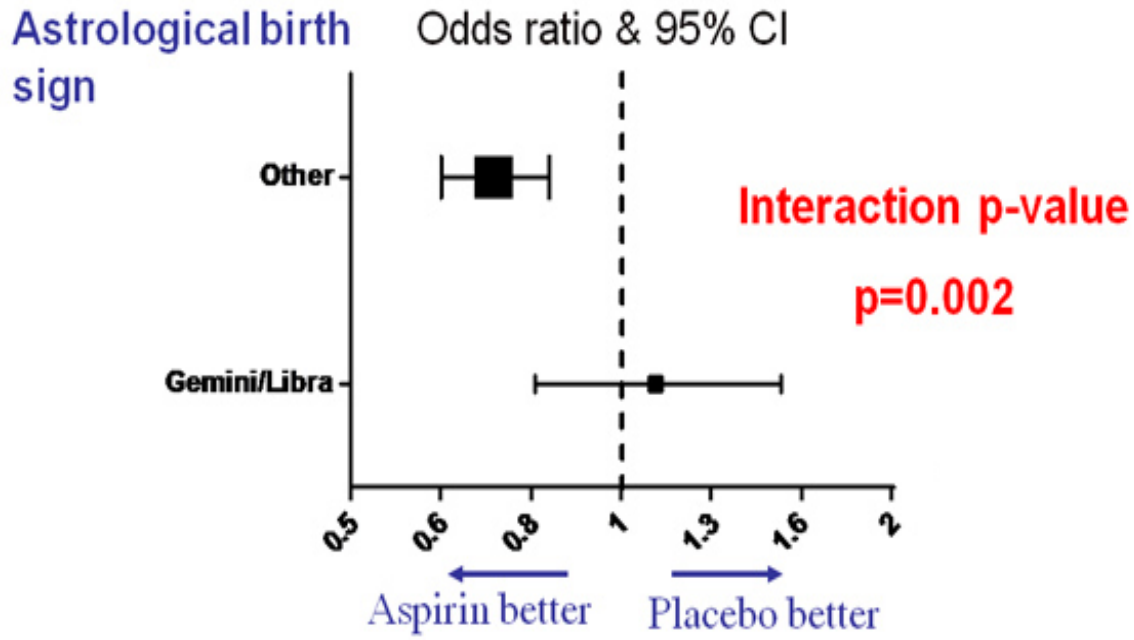


- **Can we *reliably trust* in a significant difference between arms in the (*pre-* or *post-* planned) *analysis of a specific patients' subgroup?***
  - **Risks of Subgroup Analysis**

Subgroup Analyses:
The Case of '*Gemini & Libra*' in ISIS2 Trial

ISIS-2: aspirin vs control - effects on vascular death in 17,187 patients with acute myocardial infarction (Peto et al, Lancet 1988)





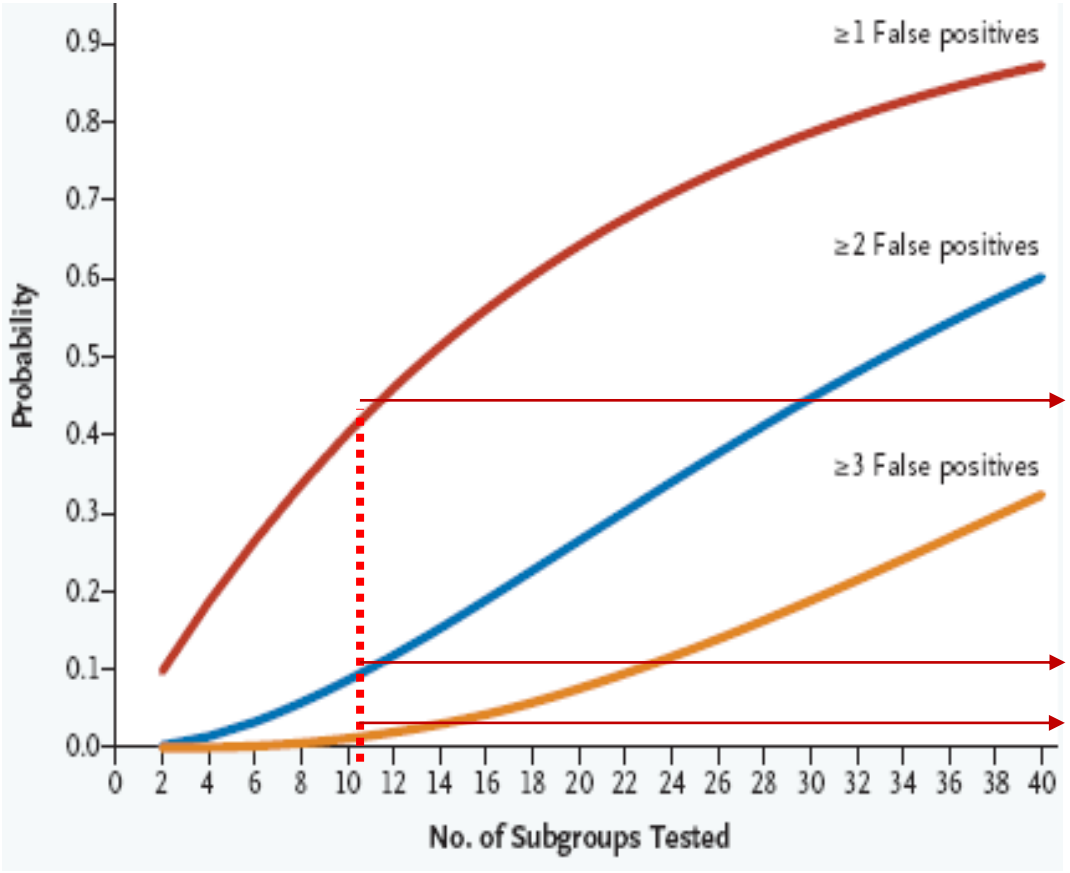
The ISIS-2 Collaborative Group. Lancet 1988

Courtesy of Michiels S, ESMO 2008

False Positive (F.P.) Risk when Testing Subgroups in RCTs

Ex.: if you test 10 subgroups, your F.P. chance is:

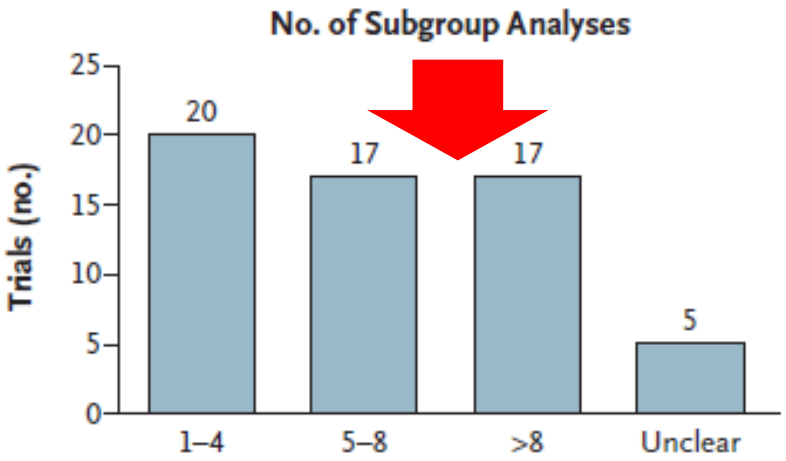
The Challenge of Subgroup Analyses — Reporting without Distorting

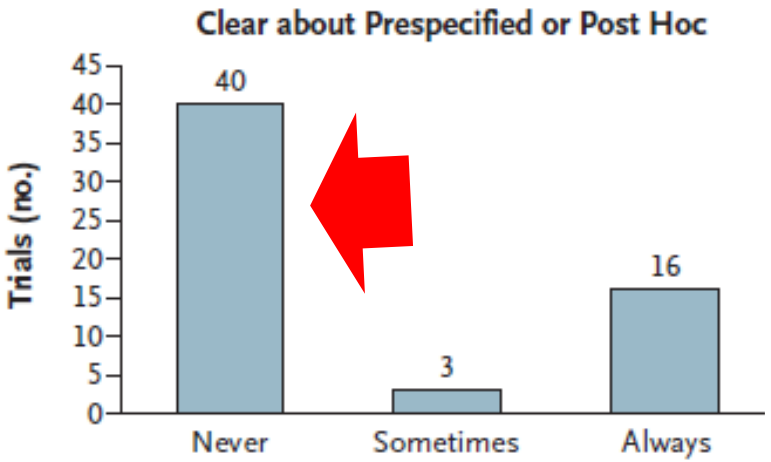


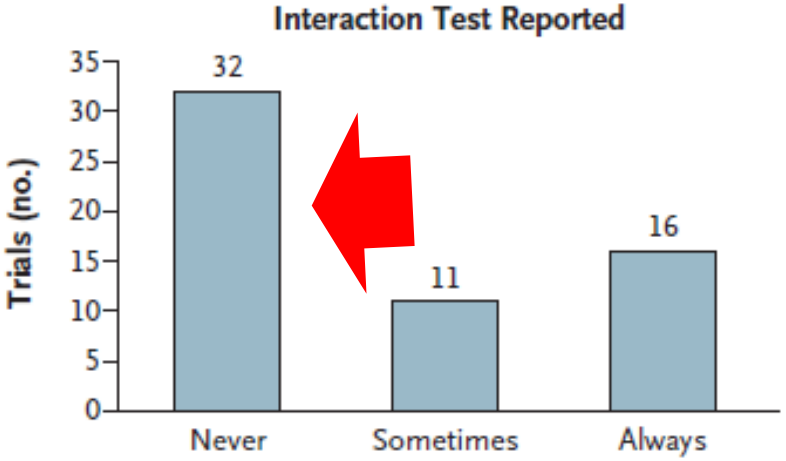
Probability That Multiple Subgroup Analyses Will Yield at Least One (Red), Two (Blue), or Three (Yellow) False Positive Results.

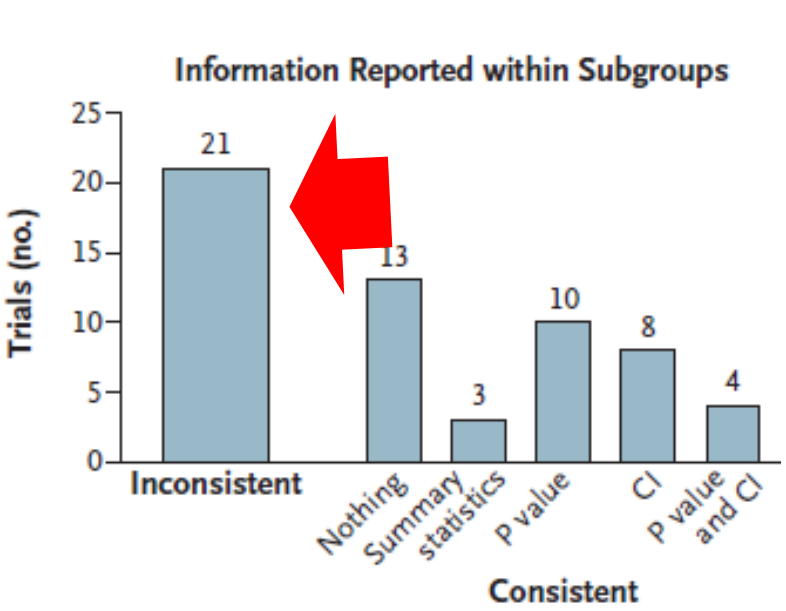
Lagakos et al, NEJM 2006

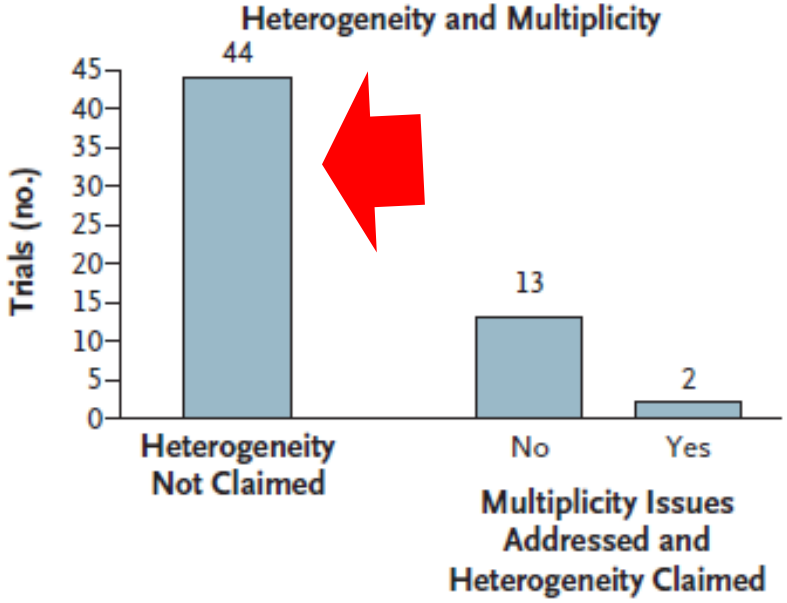
Reporting of Subgroup Analyses in Clinical Trials

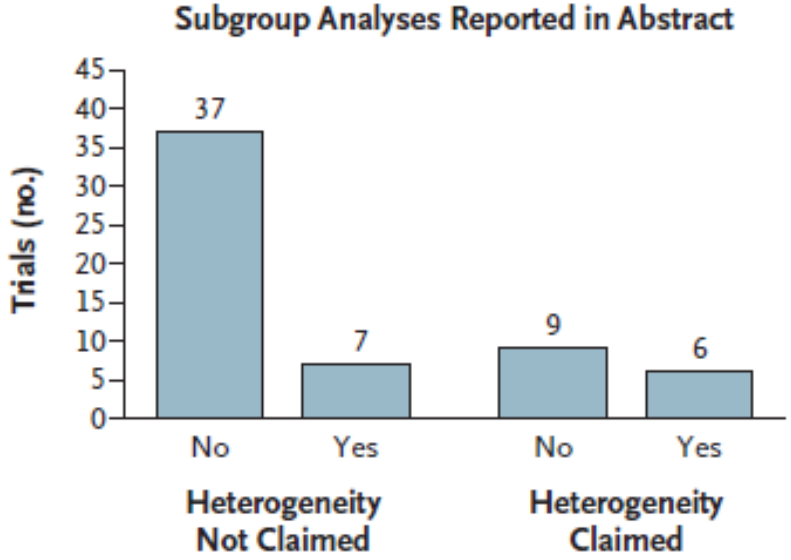












Wang R et al, NEJM 2007

Subgroup Analyses and Interactions

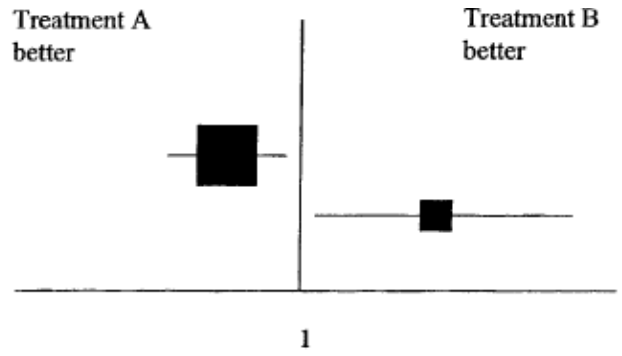
Subgroup analyses in RCTs: risk of subgroup-specific analyses

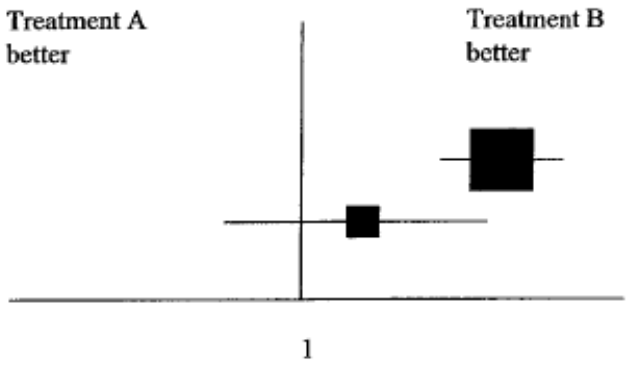
- A test for interaction between treatment and subgroup is the appropriate way to examine whether treatment effects differ between subgroups.
- This approach tests and estimates the difference between treatment effects across subgroups directly.
- It involves 1 statistical test irrespective of the number of subgroups, whereas subgroup specific analyses involve 2 or more.

Brookes ST el, J Clin Epid 2004

Types of Interaction

- **Qualitative Interaction: the direction of true treatment differences varies among subsets of patients**
  - **also called *crossover interaction***
- **Quantitative Interaction: variation in the magnitude but NOT direction of treatment effects among patient subgroups**
  - **also called a *non-crossover interaction***





Modified - Amy Wagaman, 2008

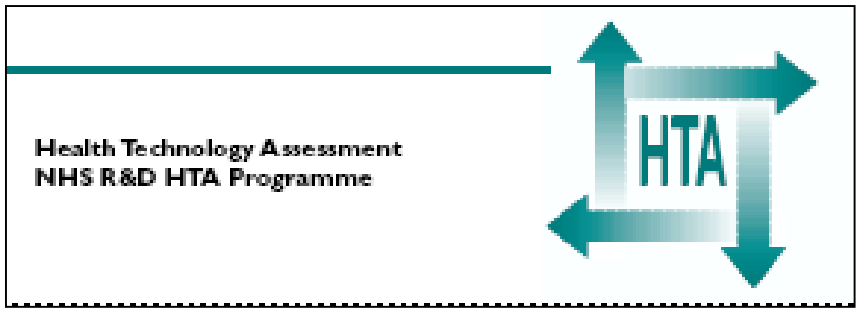
Recommendations for Subgroup Analyses

Interpretation

- Subgroup analyses **should**, as far as possible, **be restricted to those proposed before data collection**…….
- Trials should ideally be **powered** with subgroup analyses in mind……
- Subgroup-specific analyses are unreliable and affected by many factors.
- These analyses should always be **based on formal tests of interaction** although even these should be interpreted with caution.
- Unless there is strong supporting evidence, they are best viewed as a **hypothesis-generation exercise**.

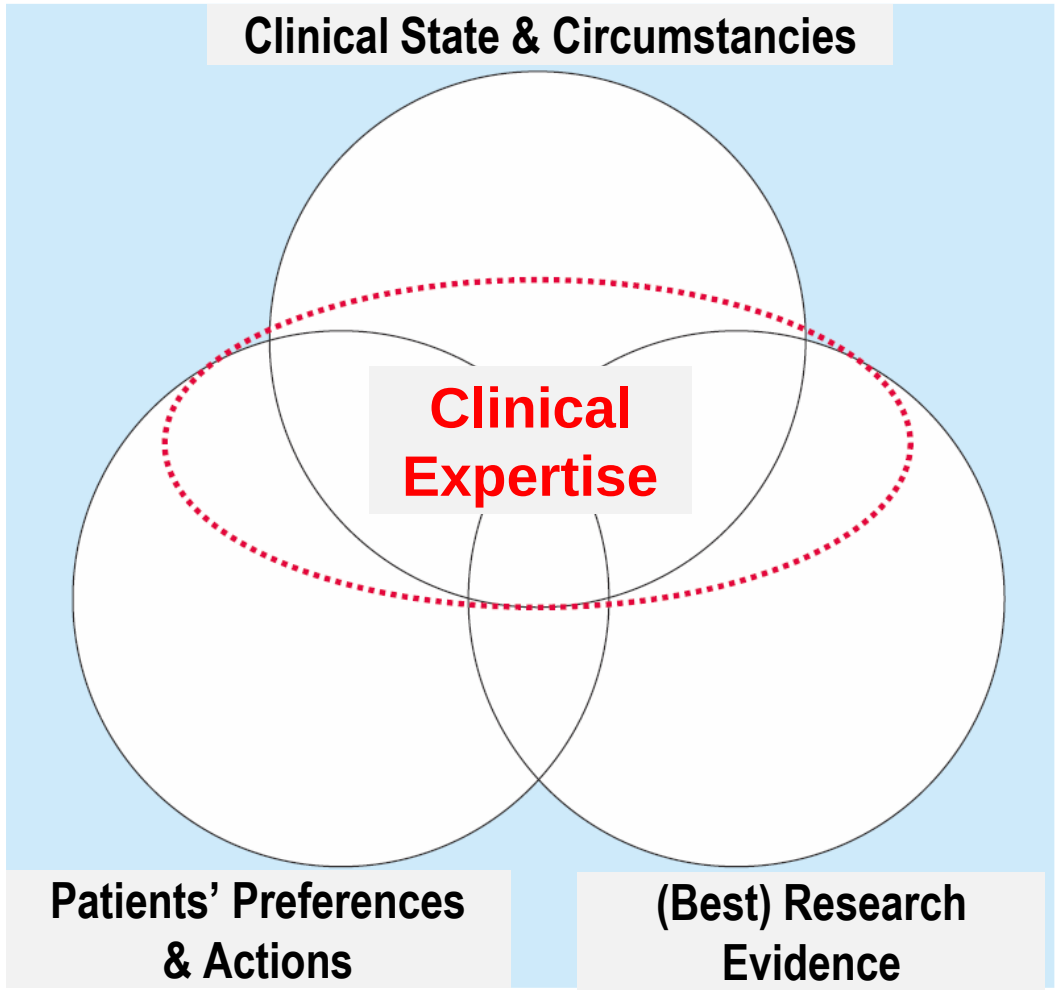
Research

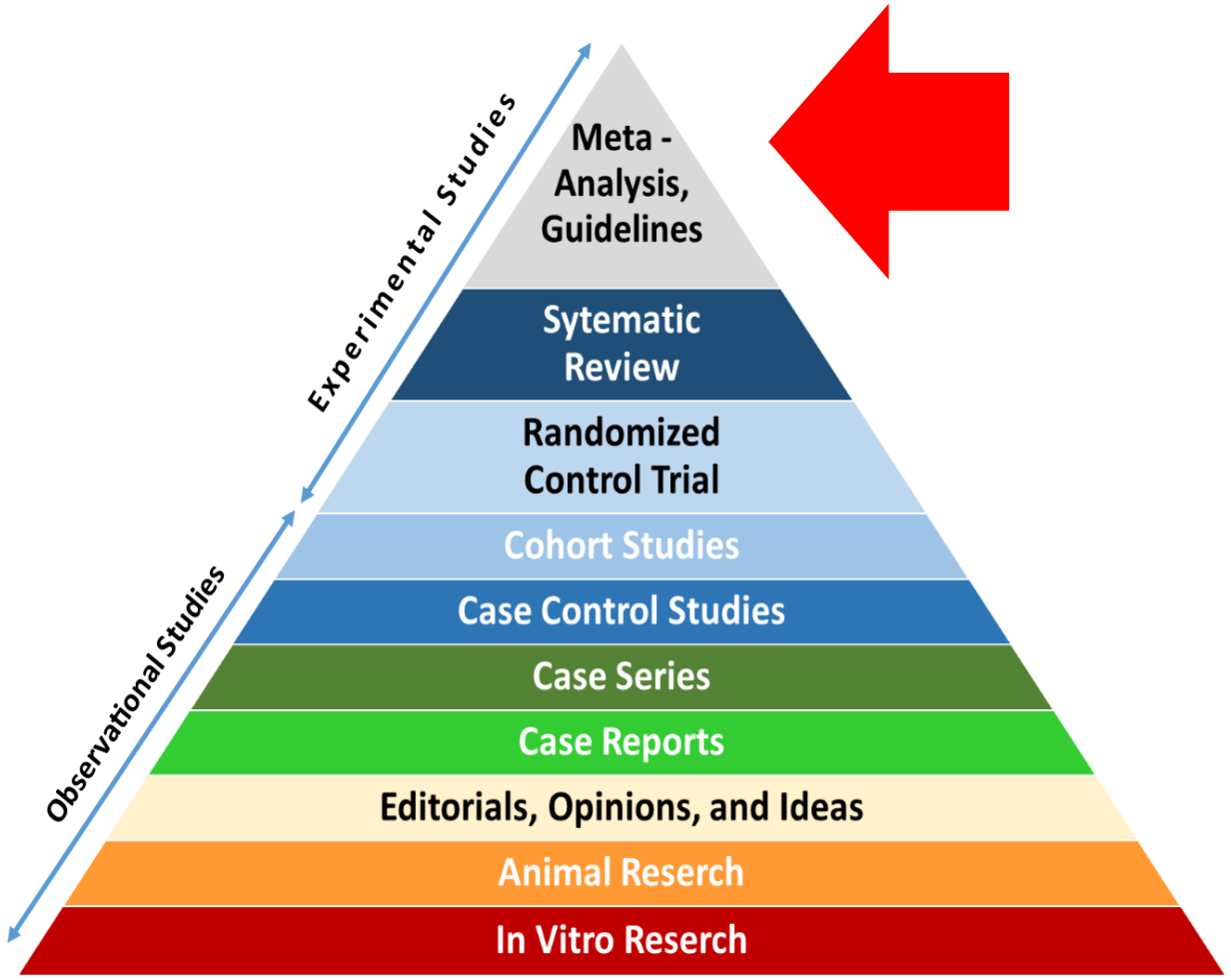
- The implications of **considering confidence intervals** rather than p-values could be considered.
- The same approach as in this study could be applied to contexts other than RCTs, such as observational studies and meta-analyses.



Brookes ST et al, Health Technology Assessment 2001

Evidence Based Medicine

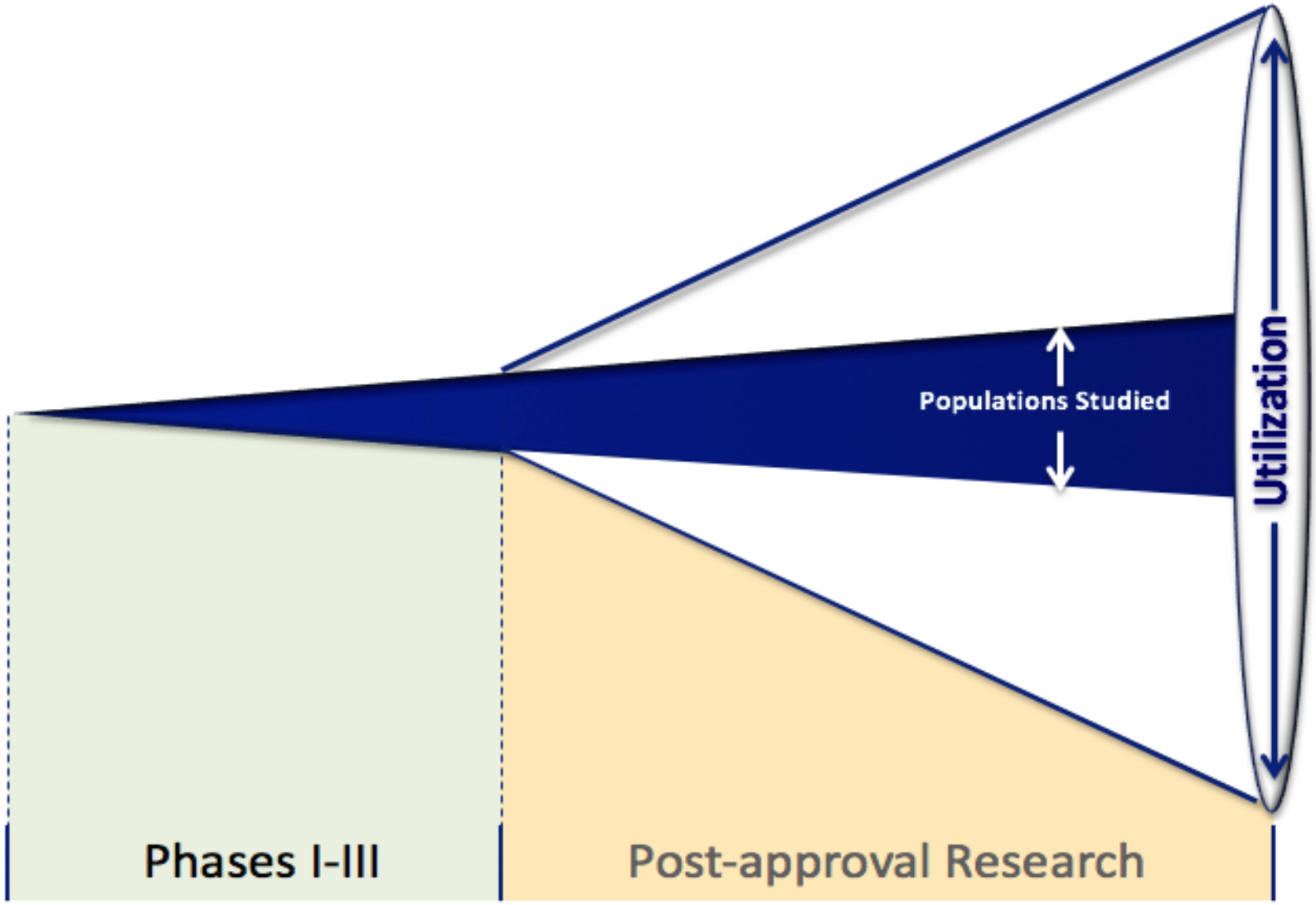




Modified from Haynes RB et al, BMJ 2002; 324: 1350

NCI Web Site, 2004

The ‘Pyramid’ of Evidence is Evolving

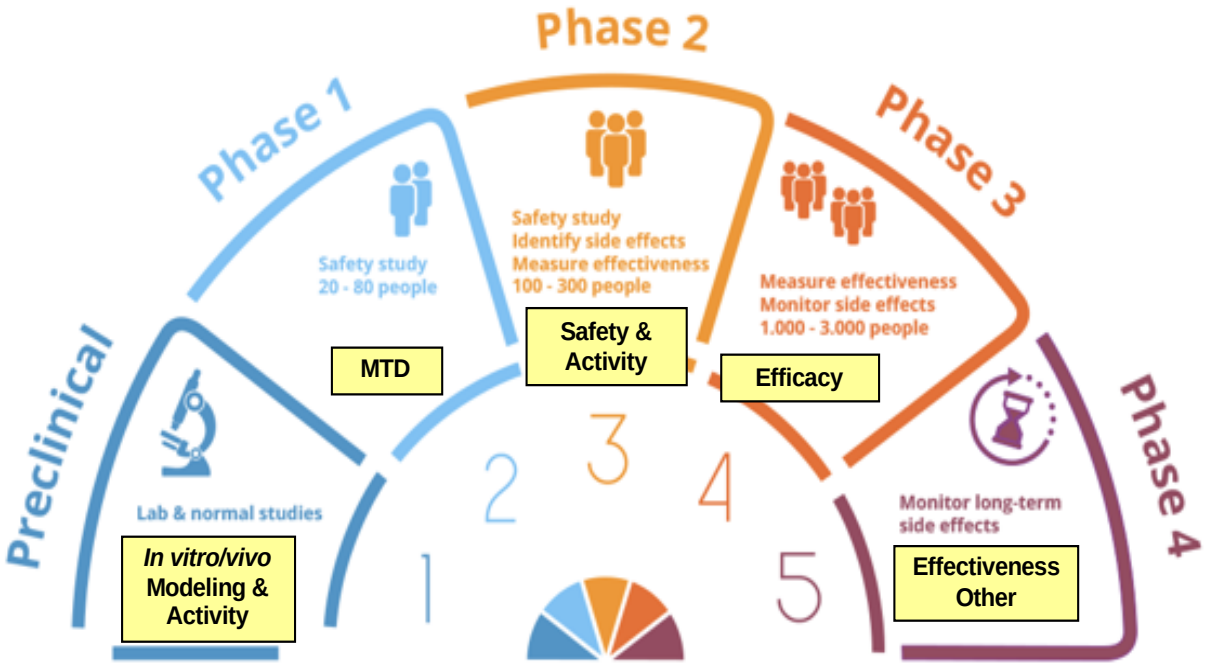


Variances in populations utilizing technology versus the populations studied

- Differing age groups (elderly, pediatrics)
- Race, ethnicity & gender variances
- Unstudied co-morbid conditions
- Differing concomitant drugs (including OTC)
- Lifestyle variances including smoking, dietary habits
- Differences in disease severity
- Varying levels of compliance

Modified from Di Maio M

Clinical Studies: Phases and Objectives

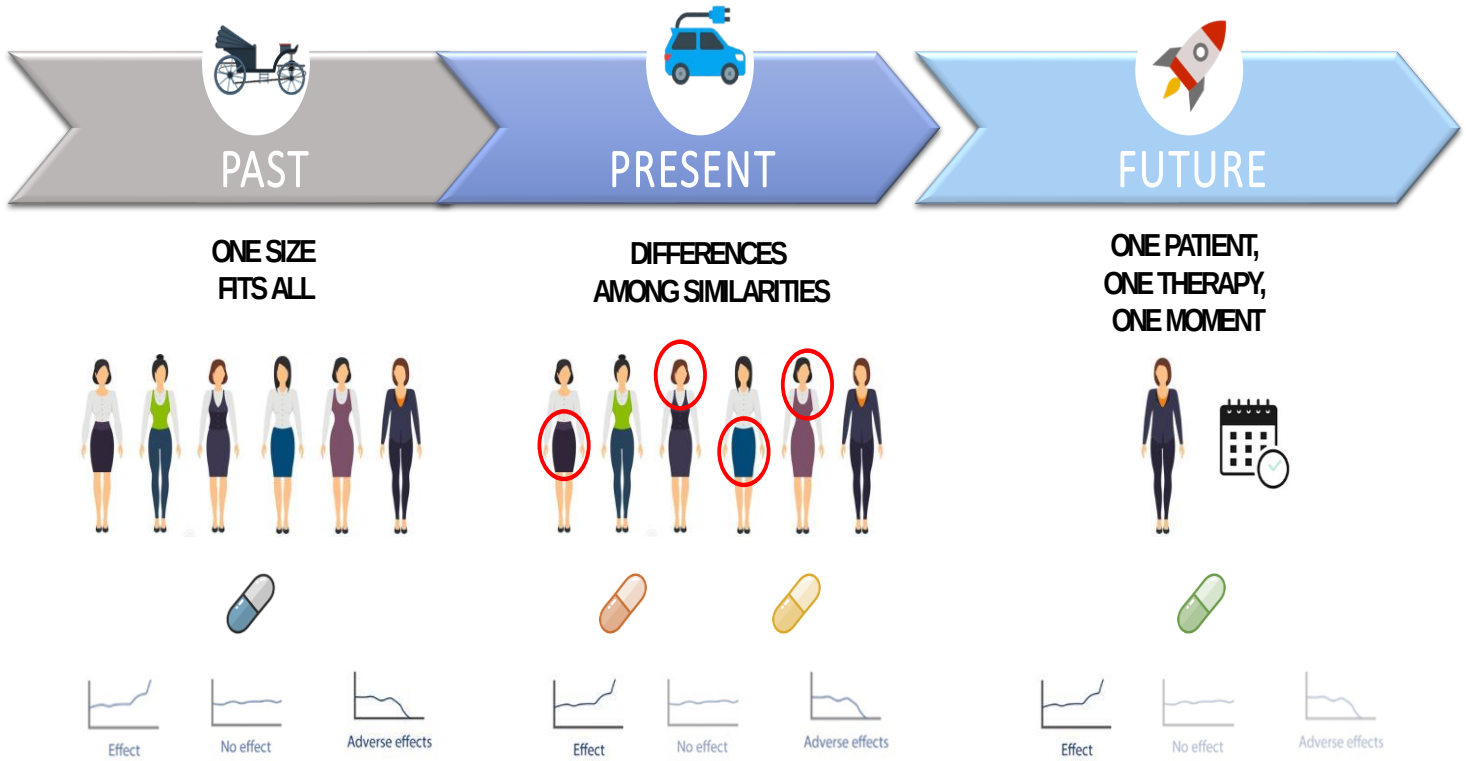


- **Activity:**
  - ability of the treatment to induce modifications of the disease thanks to which it is assumed that the patient may have a benefit **[Phase II]**
- **Efficacy:**
  - ability of the treatment to induce a clinical benefit in patients who are administered ***in an experimental context*** **[Phase III]**
- **Effectiveness:**
  - ability of a treatment to be effective in a ***non-experimental, concrete and coincident with the clinical practice [Phase IV?]***

The Humanity Perspective and Vision……

To Cure the Environment…….

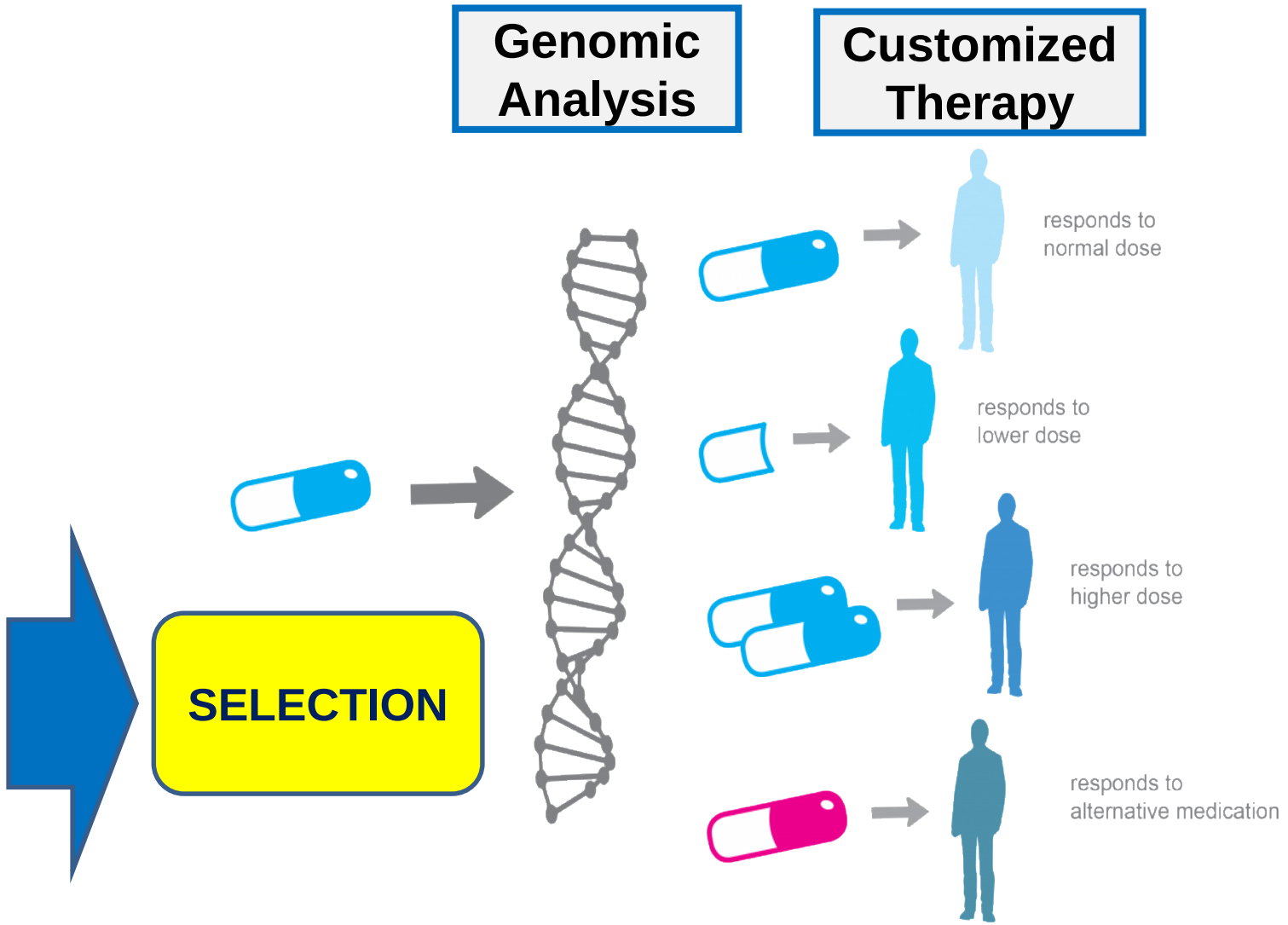
To Cure Cancer…….



Modified from Nero C, 2022

The Evolution of Biomarker-Driven Oncology

Therapy	*Era of application*
Locoregional treatments	**1940 - 50**
Chemotherapeutic- based systemic treatments	**1960 - 80**
Targeted treatments	**Latest 6-7 years**
Genomic-based treatments	**Going to start**



Modificato da Tortora GP &Schilsky RL. *Nature Rev. Drug Discovery* 9, 363-366, 2010

How MUST we proceed from now on? NCI Statment

Cancer-
To-Do List

- Performing Genomic Analyses

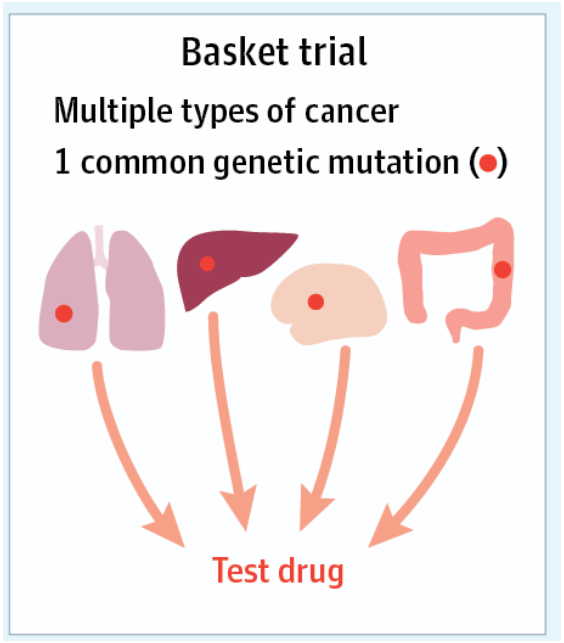
- Developing Unconventional Trials
- Analysis of Exceptional Responders

Varmus H, Science 2013

Molecular Profiling dealing with '*Unconventional*' Trials

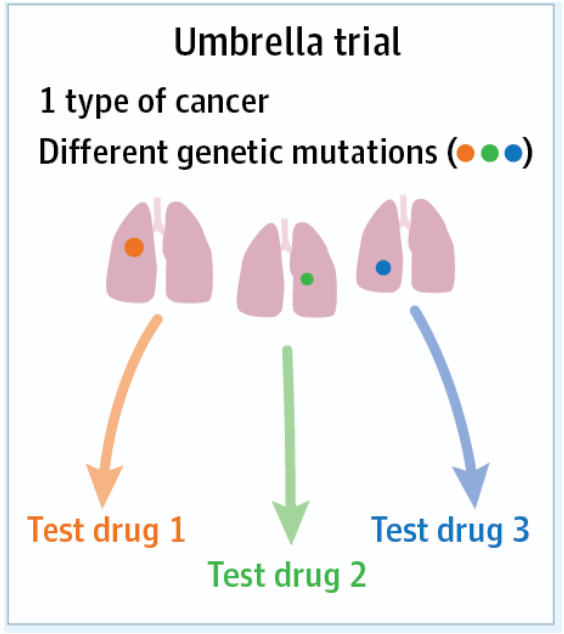
3 Types of Trials/Studies:

Basket (or Bucket)



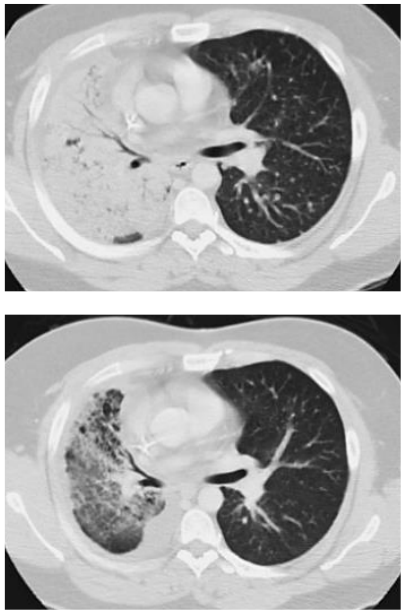
Ex. Herbst R et al, CCR 2015

Umbrella



Ex. Kummar et al, JNCI 2015

'Exceptional' Responders



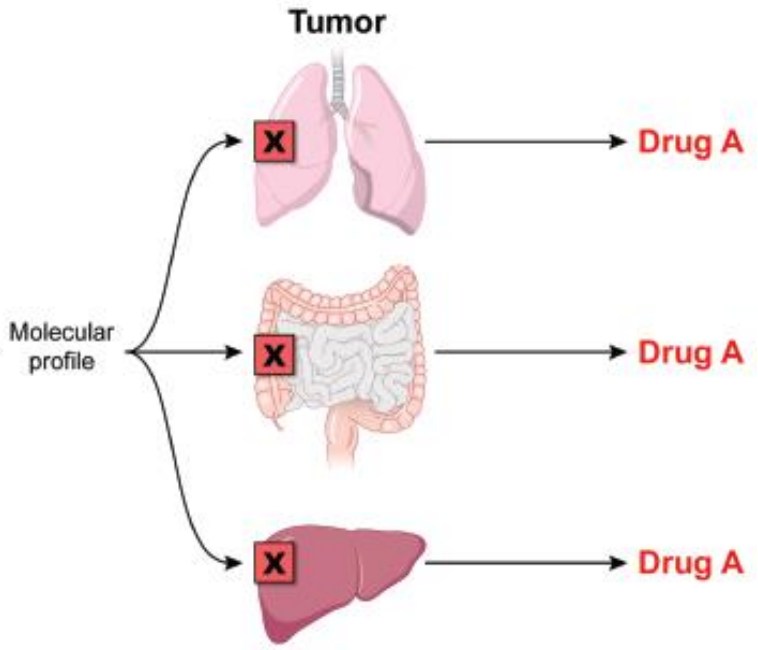
Ex. Lynch T et al, NEJM 2004

'*Unconventional*' Trials: Basket (or Bucket) Trials

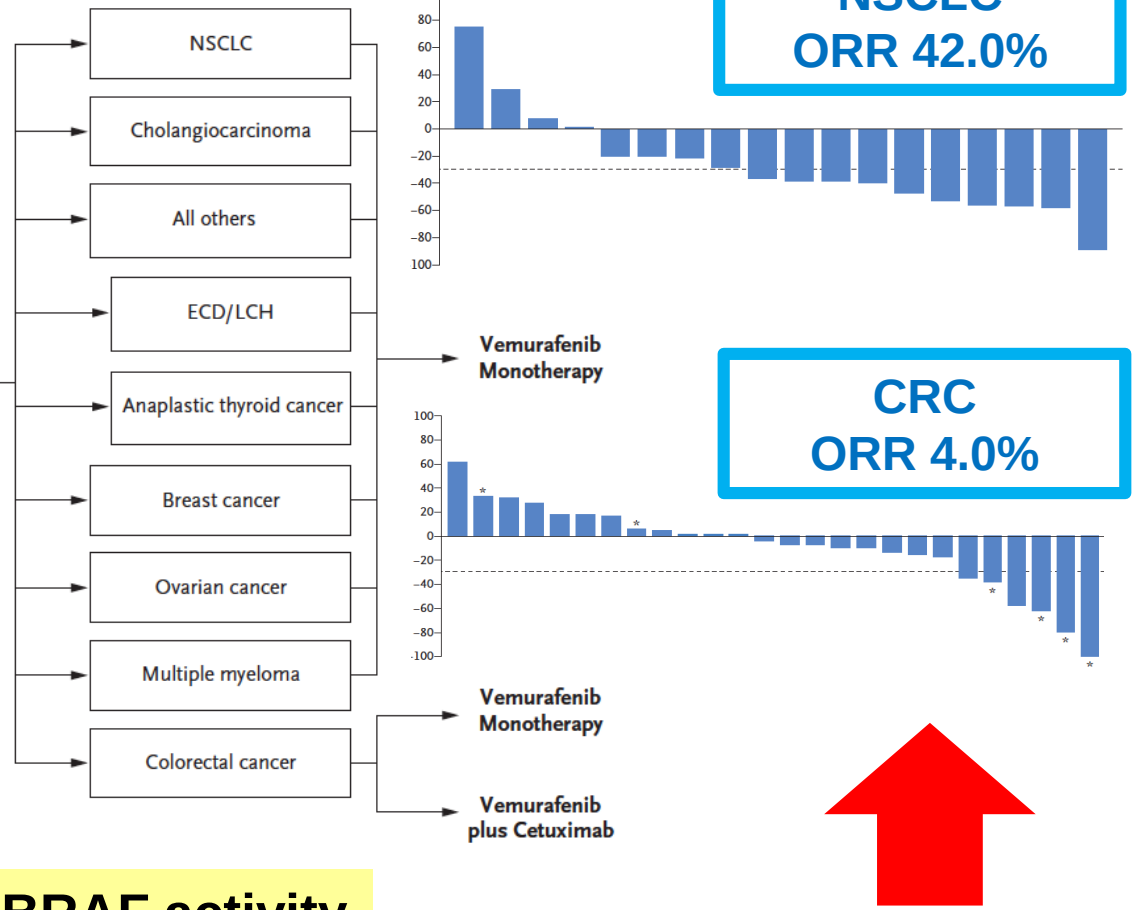
Basket or bucket trials

Single drug targeting a single mutation

Variety of tumors carrying genetic aberration X



BRAF V600–positive (testing per local methods)
Vemurafenib, 960 mg twice daily orally
Primary end point
Response rate at wk 8
Secondary end points
Progression-free survival
Time to progression
Best overall response
Time to response
Duration of response
Clinical benefit rate
Overall survival
Safety

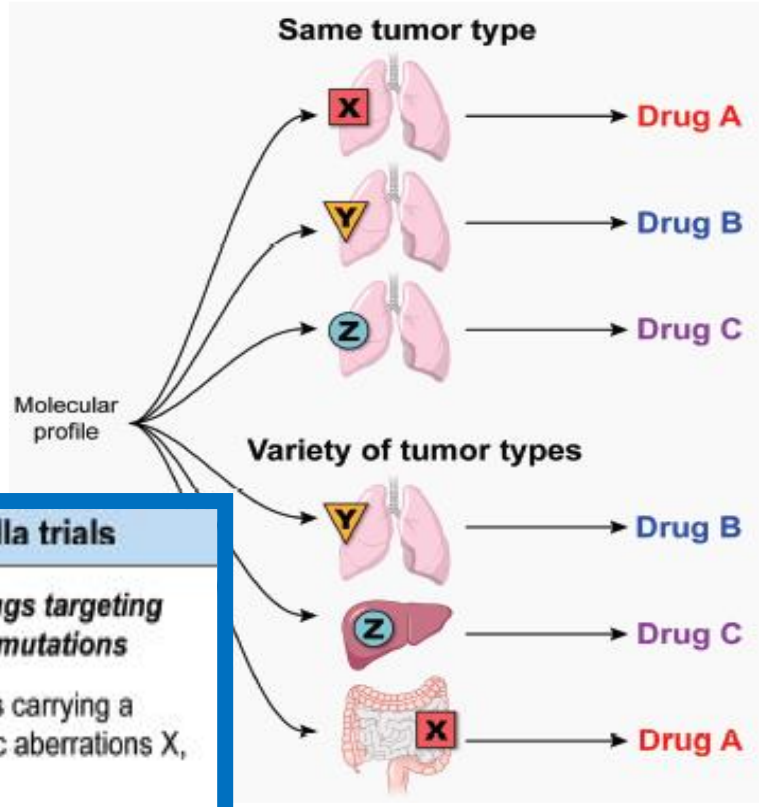


Differences for anti-BRAF activity for BRAF V600E mutations between CRC and NSCLC

Kummar et al, JNCI 2015

Hyman DM et al, NEJM 2015

'*Unconventional*' Trials: 'Umbrella' Trials

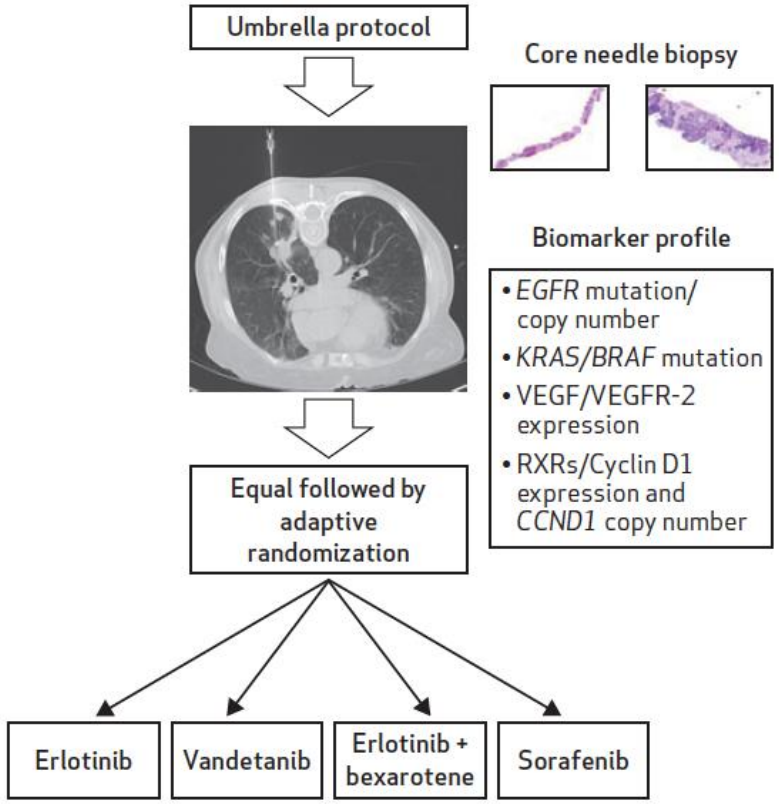


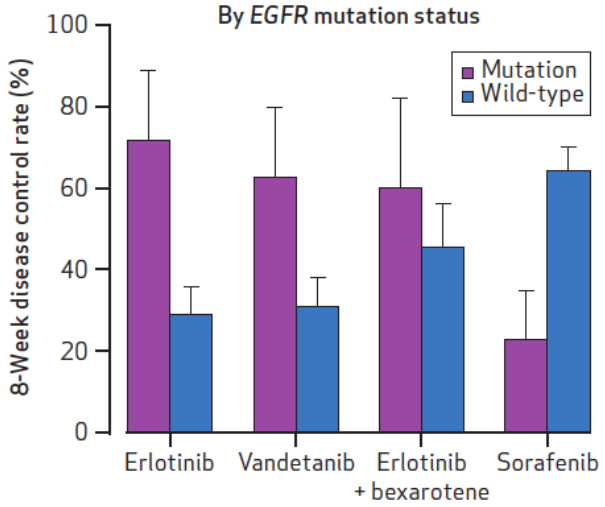
Umbrella trials

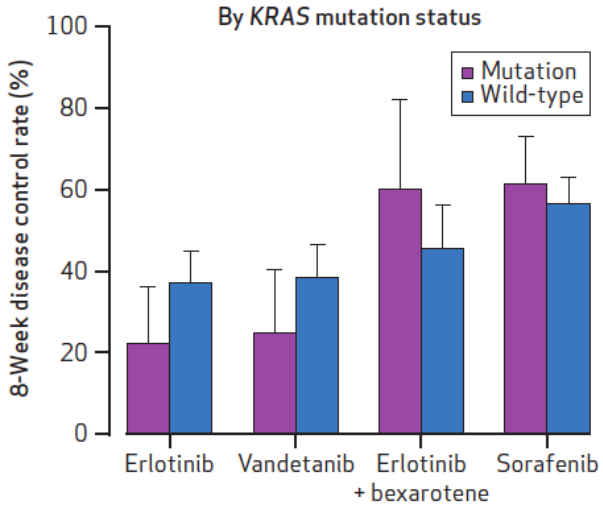
Multiple drugs targeting multiple mutations

Variety of tumors carrying a variety of genetic aberrations X, Y, & Z

- Randomized or nonrandomized
- Rules-based treatment assignment or per patient based on review of individual profile data







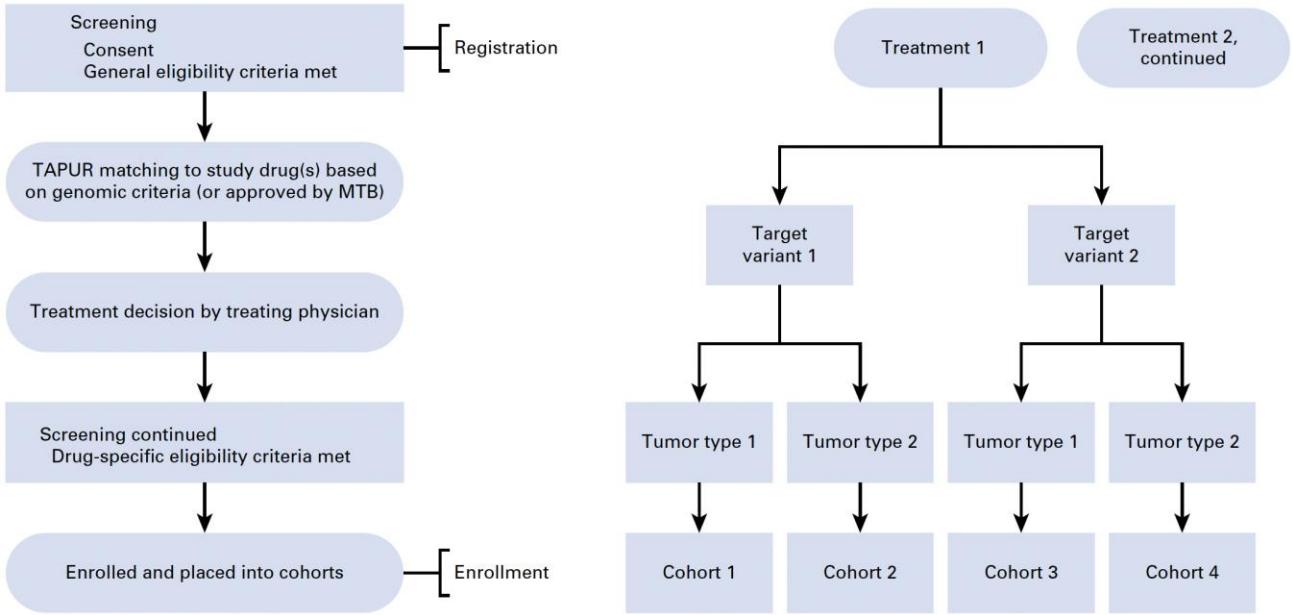
- **What about driving hypotheses and sample size?**
- **What about prognostic effect?**

Kummar et al, JNCI 2015

Kim WK et al, Cancer Disc 2011

Targeted Agent and Profiling Utilization Registry Study [TAPUR]

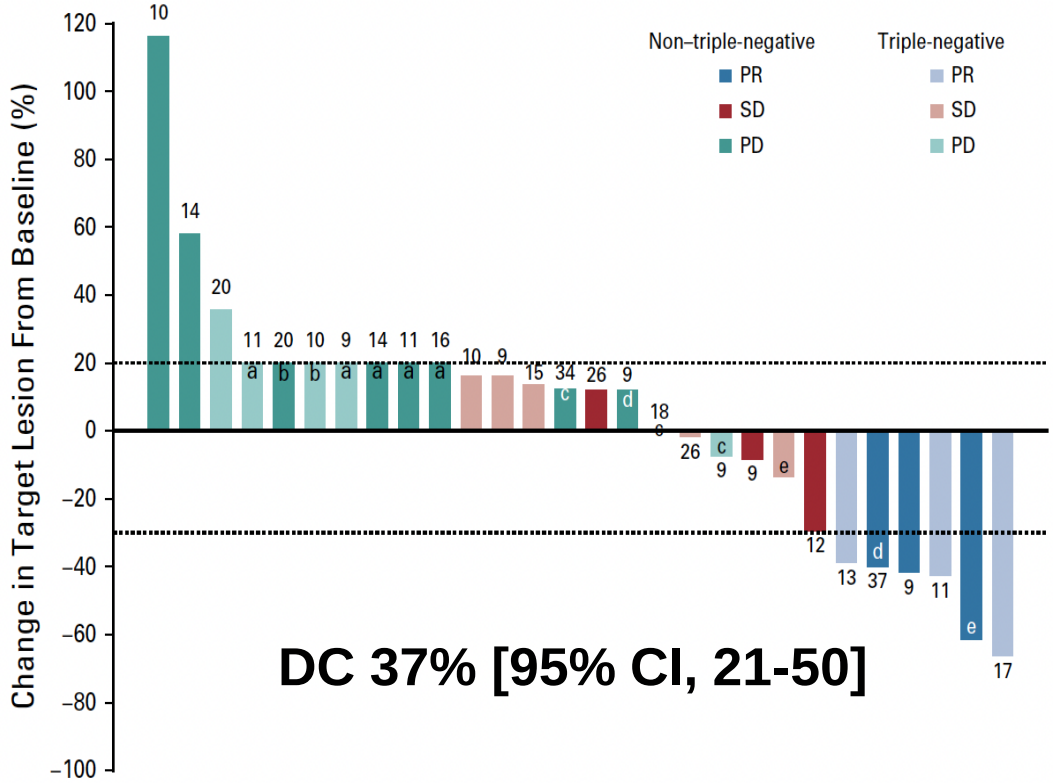
- **Phase II prospective, nonrandomized, multibasket pragmatic clinical trial**
- **AIM: to identify signals of activity for FDA–approved drugs matched to prespecified targets**



- **Primary End-Point: Disease Control (DC), ORR at ≥ 8 weeks or documented SD at ≥ 16 weeks - H0 (ORR/SD)=15%, H1=35% (Power 85%, Alpha 0.10, one sided)**

Mangat PK et al, JCO Prec Oncol 2028

Pembrolizumab in Patients With Metastatic Breast Cancer With High Tumor Mutational Burden (H-TMB)



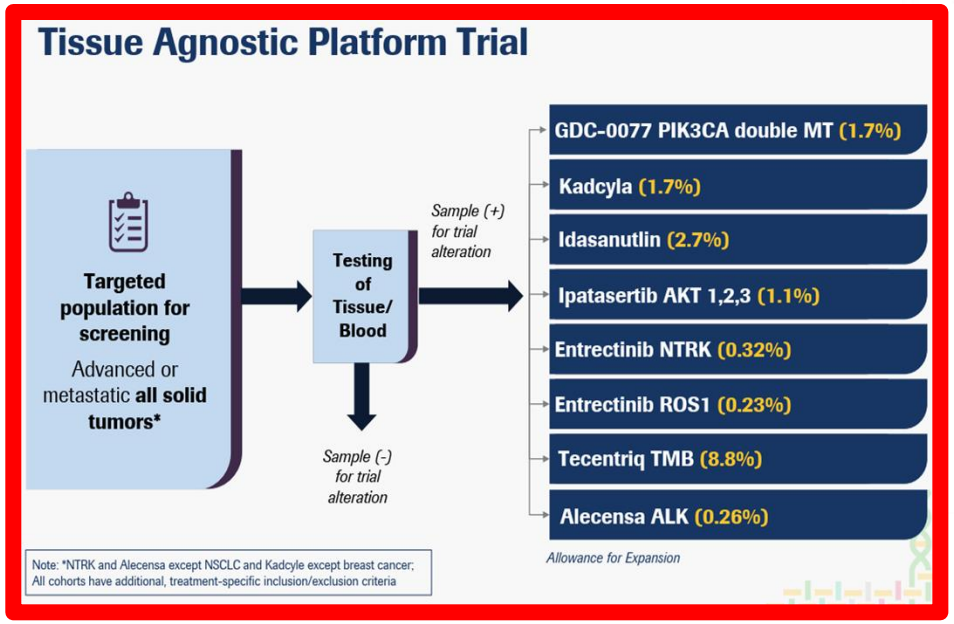
DC 37% [95% CI, 21-50]

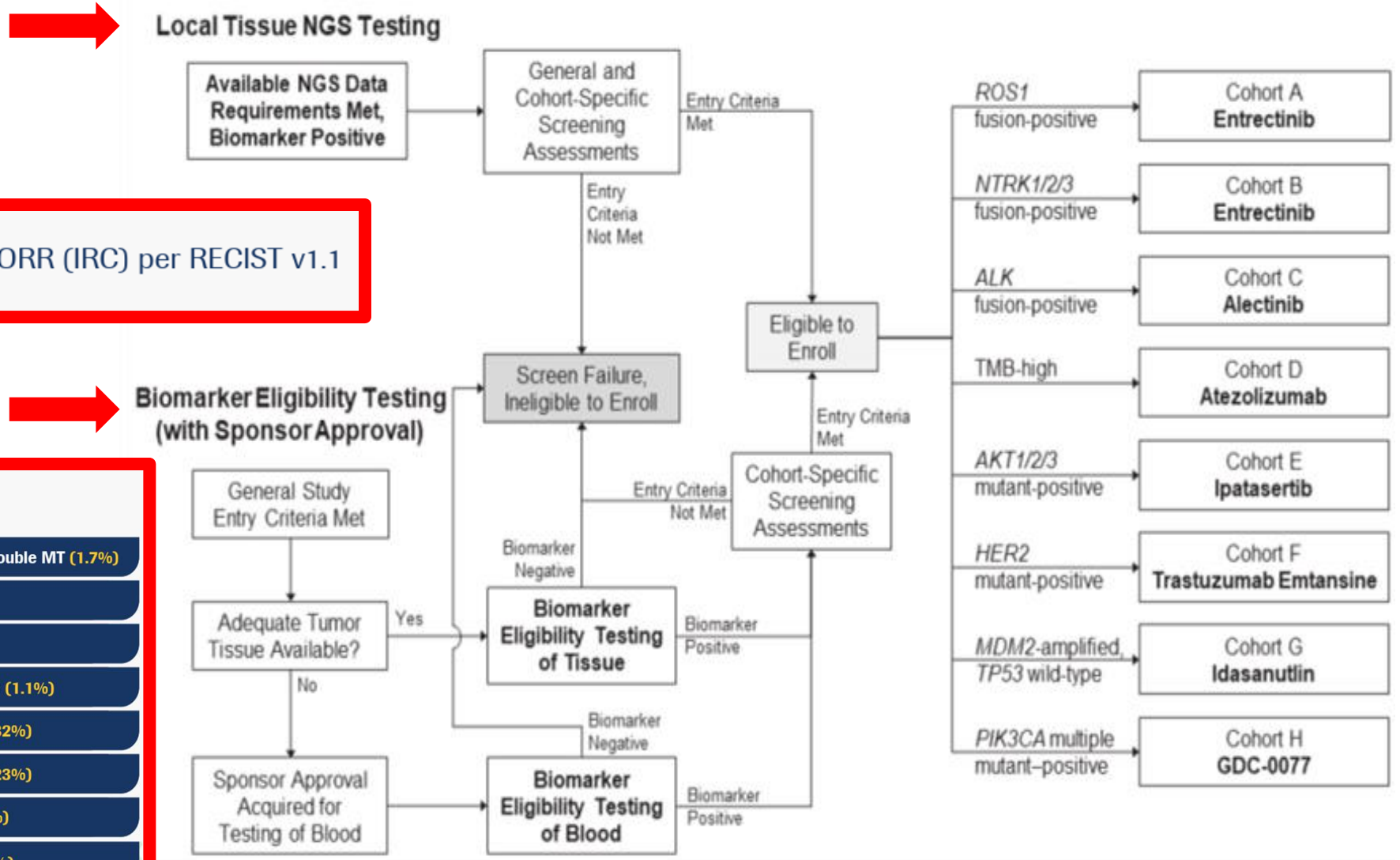
Alva AS et al, JCO 2021

Tumor-Agnostic Precision Immuno-oncology and Somatic Targeting Rational for You [TAPISTRY] - Phase II Platform Trial

Primary Endpoints

- Confirmed ORR (IRC) per RECIST v1.1

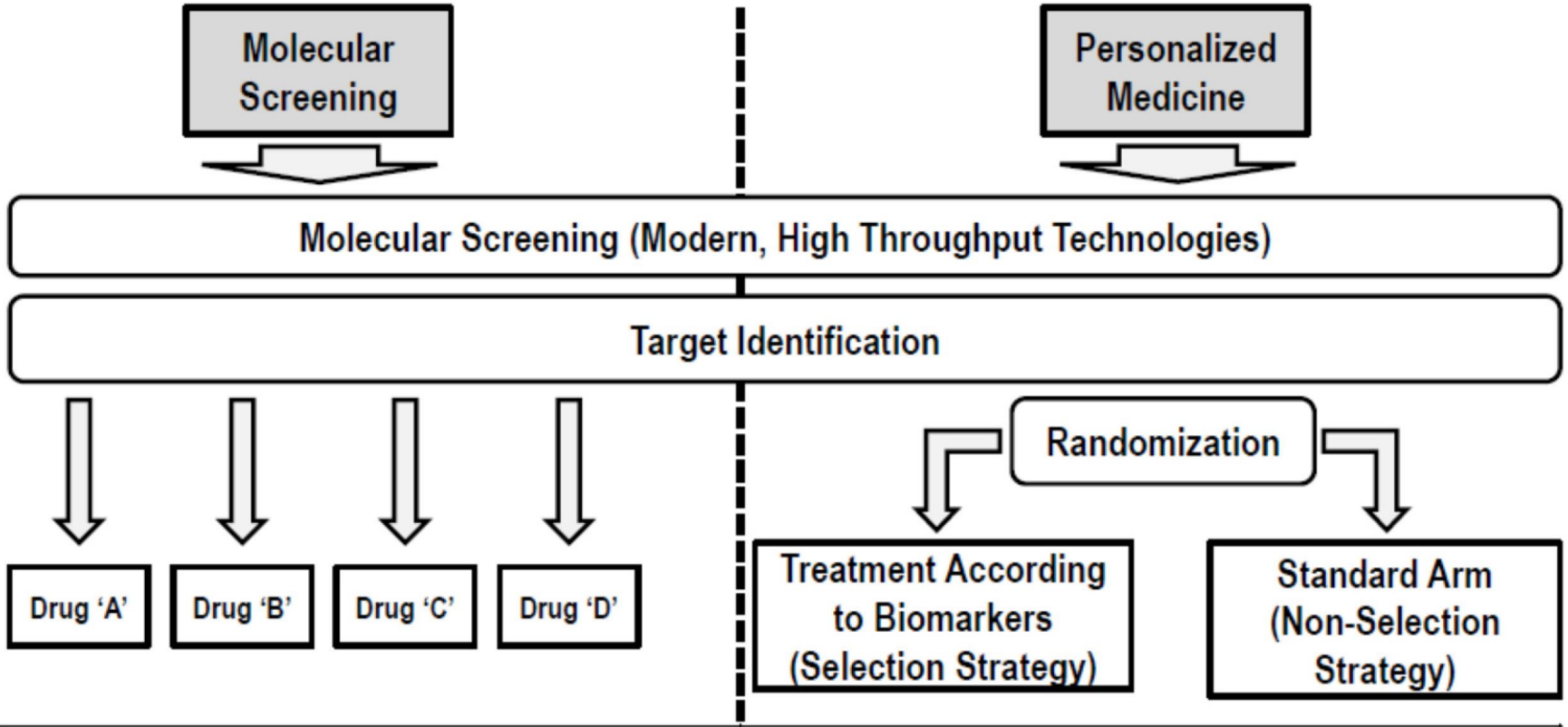




P.I. Daniele G

Unità di fase I – fase1@policlinicogemelli.it

'*Unconventional*' Trials: MS and PM Studies



Goal	Parallel drug development according to biomarkers	Efficacy of molecular screening (selection versus non-selection strategy)
Ideal Genomic Alteration	The best candidates are those alteration occurring in ≥ 10% overall population	The higher the incidence of each alteration, the stronger the impact on efficacy
Hypotheses	One hypothesis for each downstream trial (i.e. activity in enriched parallel subgroups)	Using NGS (or genomics overall) regardless of the target improves outcome

Rossi A et al, J Molecular Pathology 2021

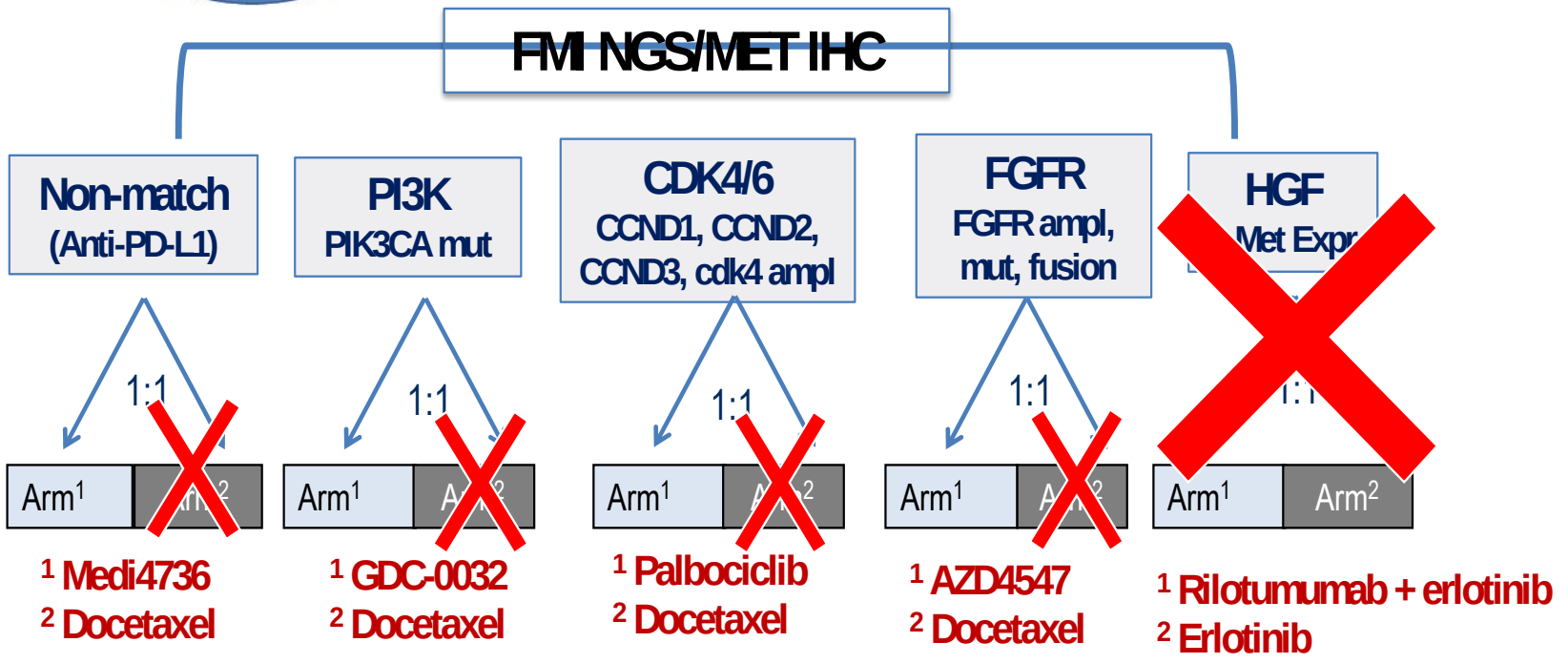
LUNG MAP Molecular Profiling: Master Lung [Squamous] Study

S1400 Master Protocol
Unique Private-Public Partnerships with the NCI

Genetic Alterations Identified

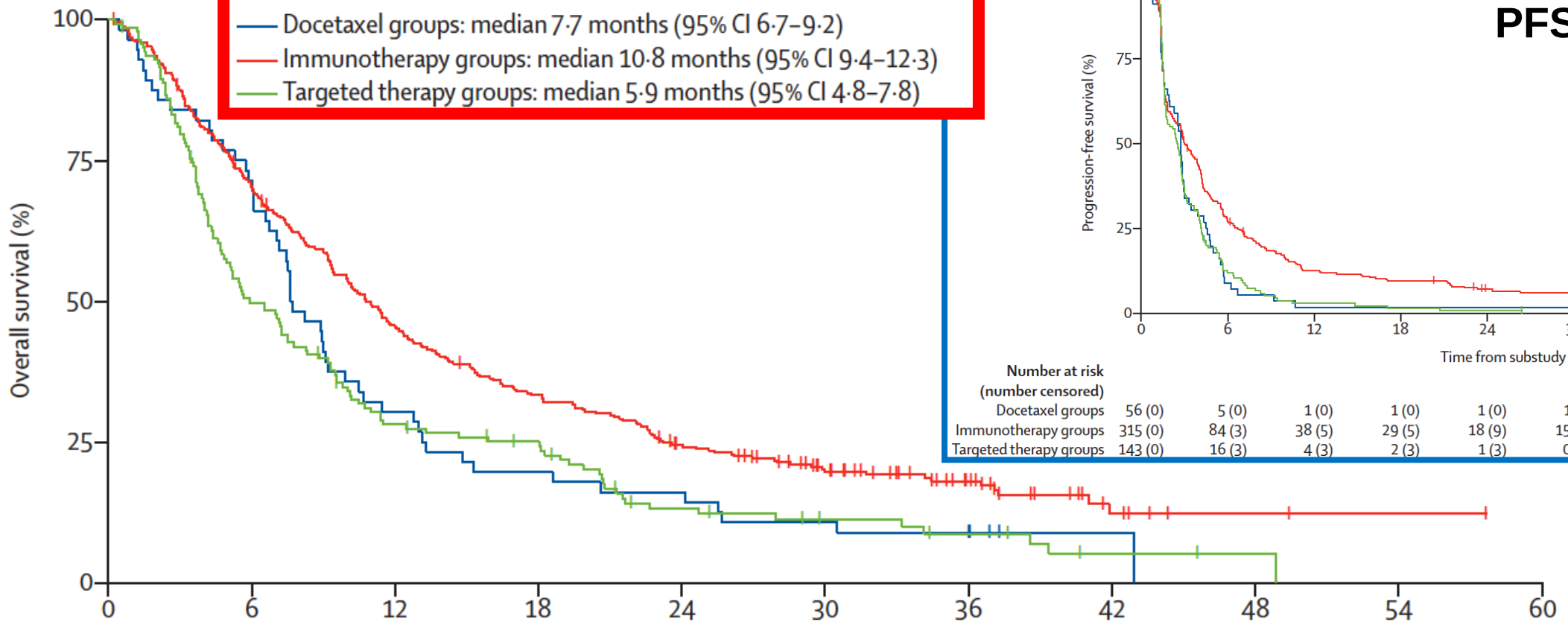
Gene	Event Type	Frequency
FGFR1	Amplification	20-25%
FGFR2	Mutation	5%
PIK3CA	Mutation	9%
PTEN	Mutation/Deletion	18%
CCND1	Amplification	8%
CDKN2A	Deletion/Mutation	45%
CMET	Amplification/Mutation	40%
PDGFRA	Amplification/Mutation	9%
EGFR	Amplification	10%
MCL1	Amplification	10%
BRAF	Mutation	3%
DDR2	Mutation	4%
ERBB2	Amplification	2%

- **Participants:** Entire North American Lung Intergroup (SWOG, Alliance, ECOG-Acrin, NRG, NCI-Canada)
- **Screening:** 500-1,000 patients/year
- **With 4-6 arms open simultaneously, anticipate a "hit rate >50% in matching a patient with a drug/biomarker arm**



Herbst R, WCLC 2017

LUNG MAP Molecular Profiling: Master Lung [Squamous] Study

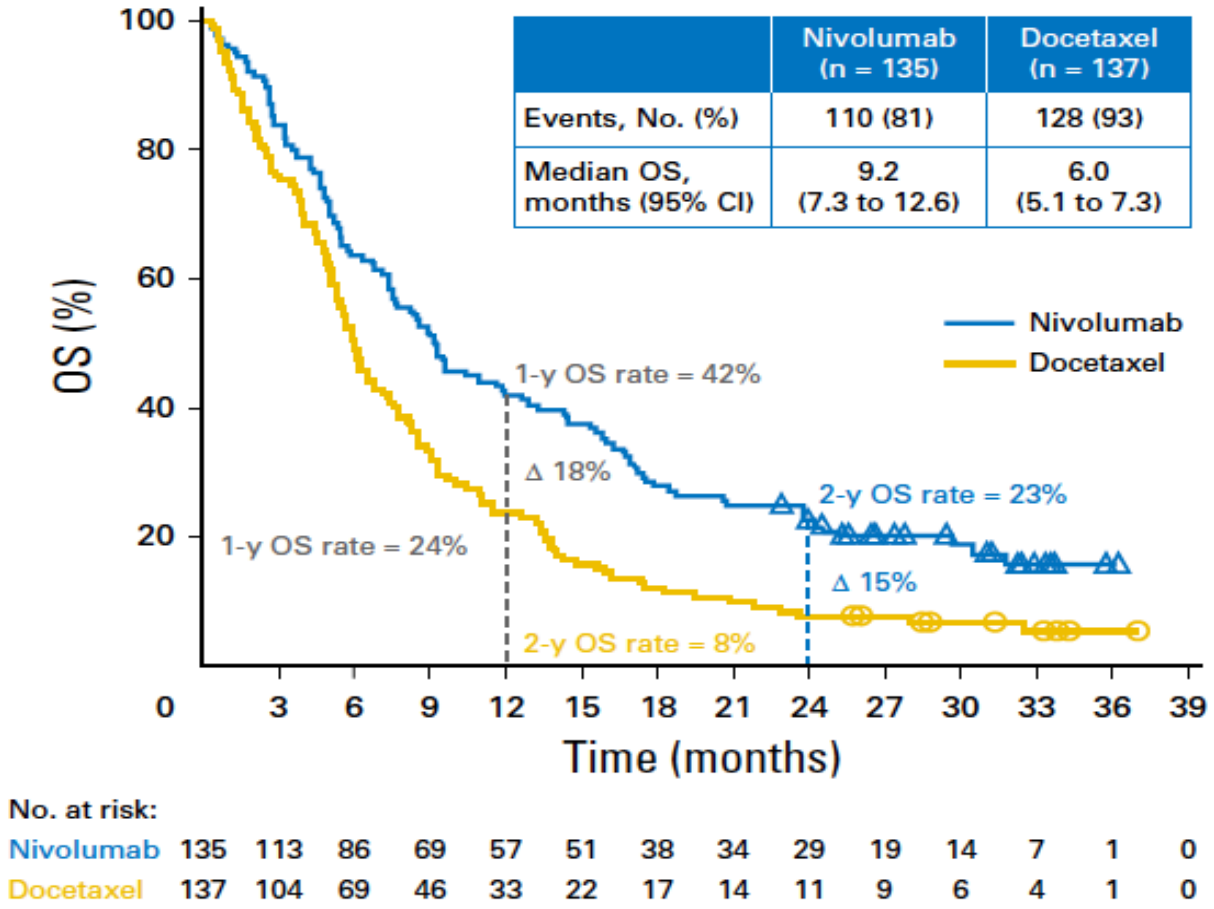


Number at risk (number censored)

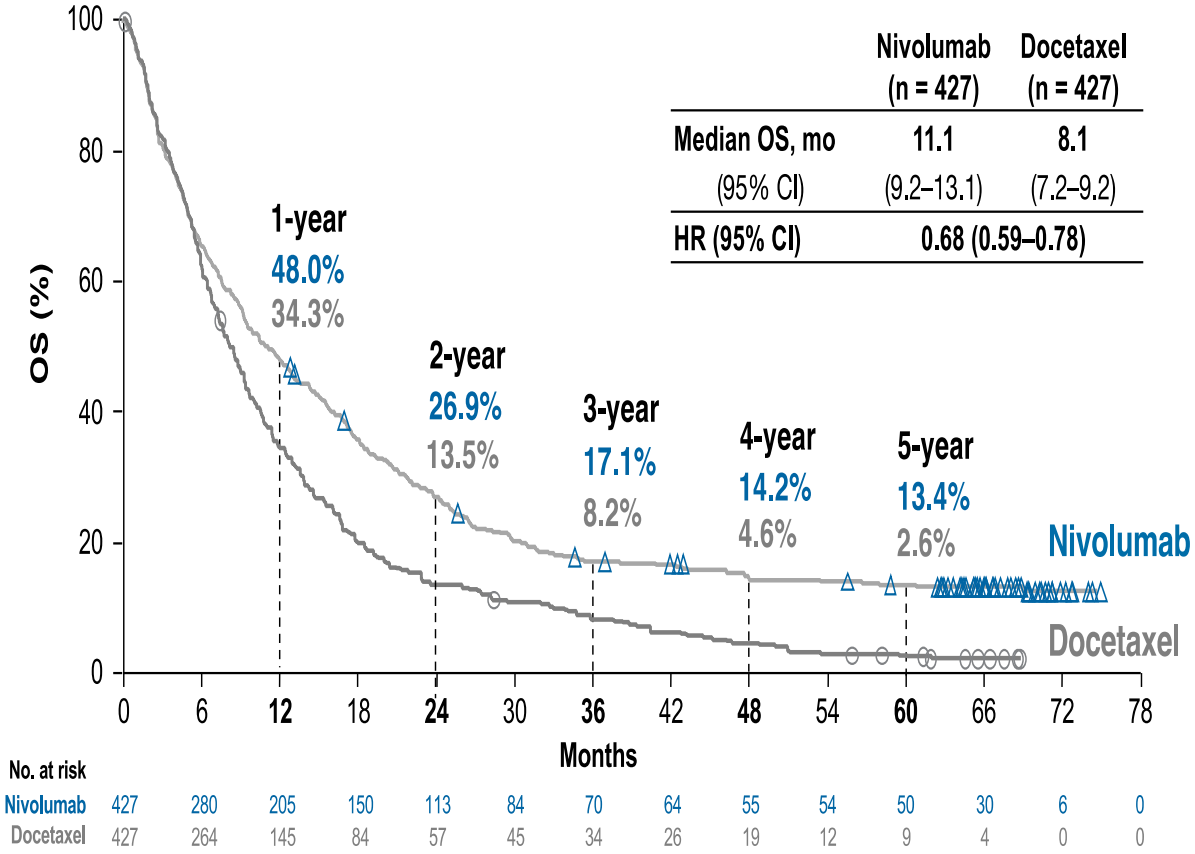
Docetaxel groups	56 (0)	40 (0)	17 (0)	11 (0)	9 (0)	6 (0)	5 (0)	1 (4)	0 (4)	0 (4)	0 (4)
Immunotherapy groups	315 (0)	219 (3)	140 (5)	102 (6)	71 (10)	47 (21)	25 (40)	7 (53)	2 (58)	1 (59)	0 (60)
Targeted therapy groups	143 (0)	70 (2)	38 (4)	31 (7)	14 (10)	9 (13)	6 (14)	2 (16)	1 (17)	0 (17)	0 (17)

Redman RW et al, Lancet Oncol 2020

Survival Update of CM 017 & Cum. 017/057



	Nivolumab (n = 135)	Docetaxel (n = 137)
Events, No. (%)	110 (81)	128 (93)
Median OS, months (95% CI)	9.2 (7.3 to 12.6)	6.0 (5.1 to 7.3)



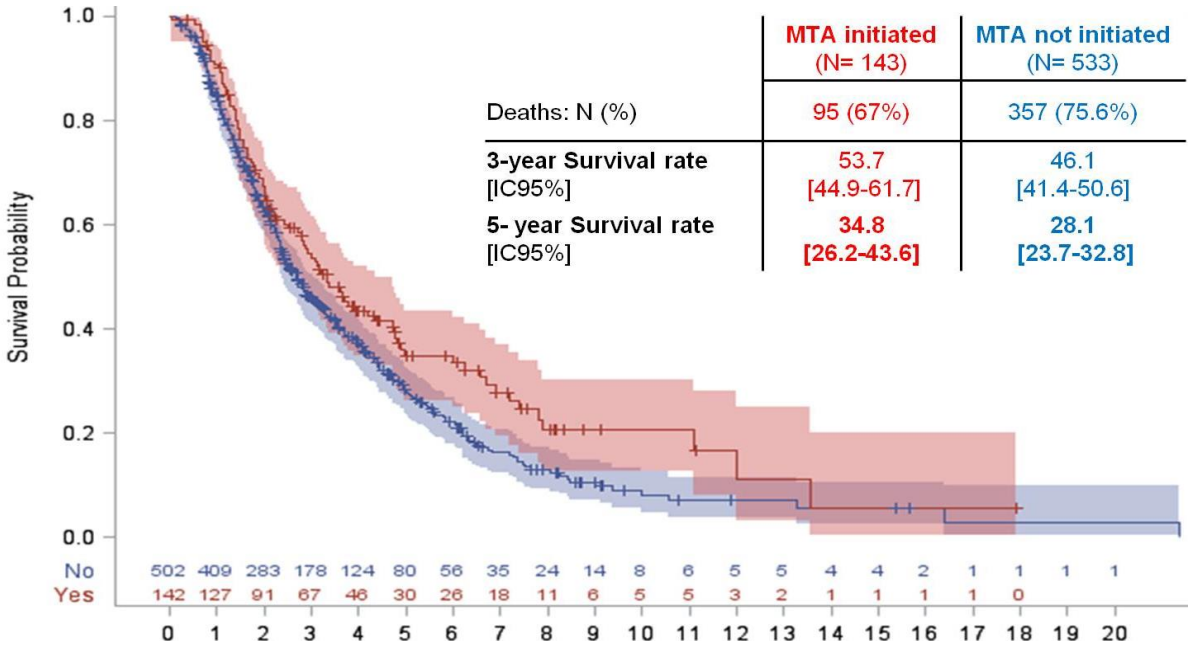
	Nivolumab (n = 427)	Docetaxel (n = 427)
Median OS, mo	11.1	8.1
(95% CI)	(9.2–13.1)	(7.2–9.2)
HR (95% CI)	0.68 (0.59–0.78)	

Horn L et al, J Clin Oncol 2017

Gettinger S et al, WCLC 2019

'*Unconventional*' Trials: MS and PM Studies

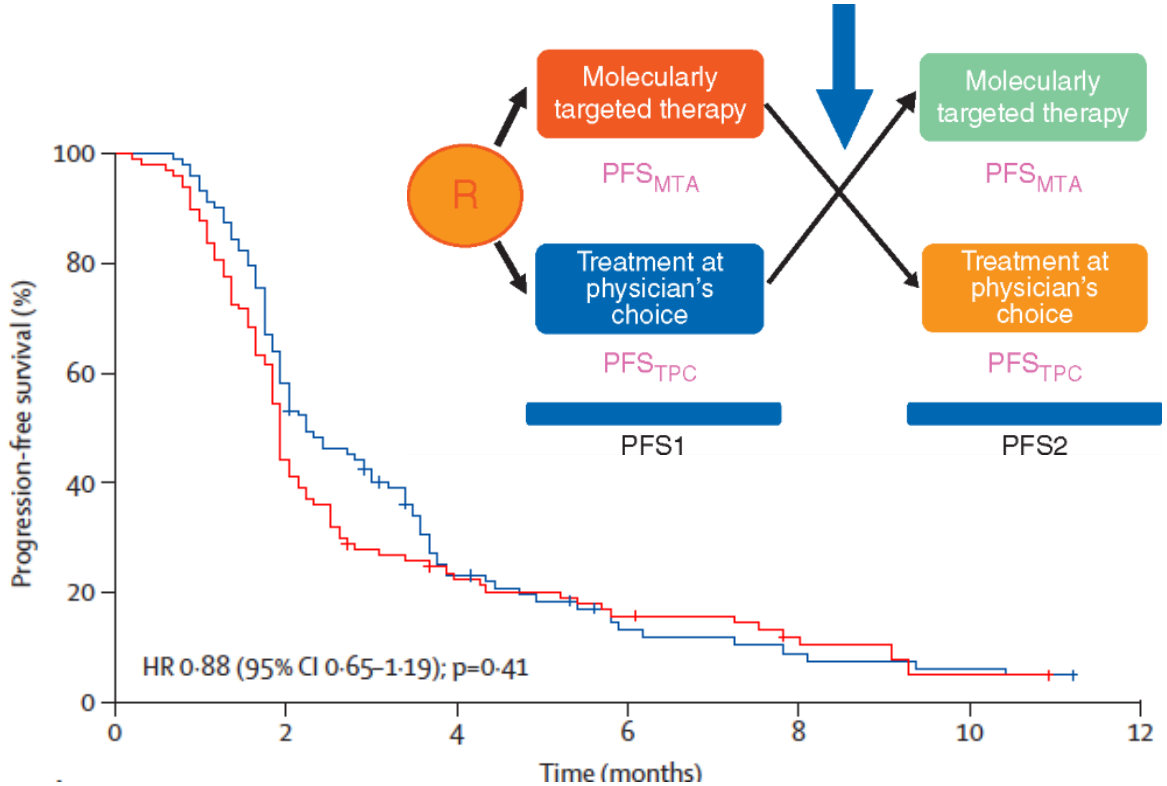
MS: PROFILER Study [Non-Randomized]



	MTA initiated (N= 143)	MTA not initiated (N= 533)
Deaths: N (%)	95 (67%)	357 (75.6%)
3-year Survival rate [IC95%]	53.7 [44.9-61.7]	46.1 [41.4-50.6]
5- year Survival rate [IC95%]	34.8 [26.2-43.6]	28.1 [23.7-32.8]

Le Tourneau E et al, Lancet Oncol 2015

PM: SHIVA Trial [Randomized]

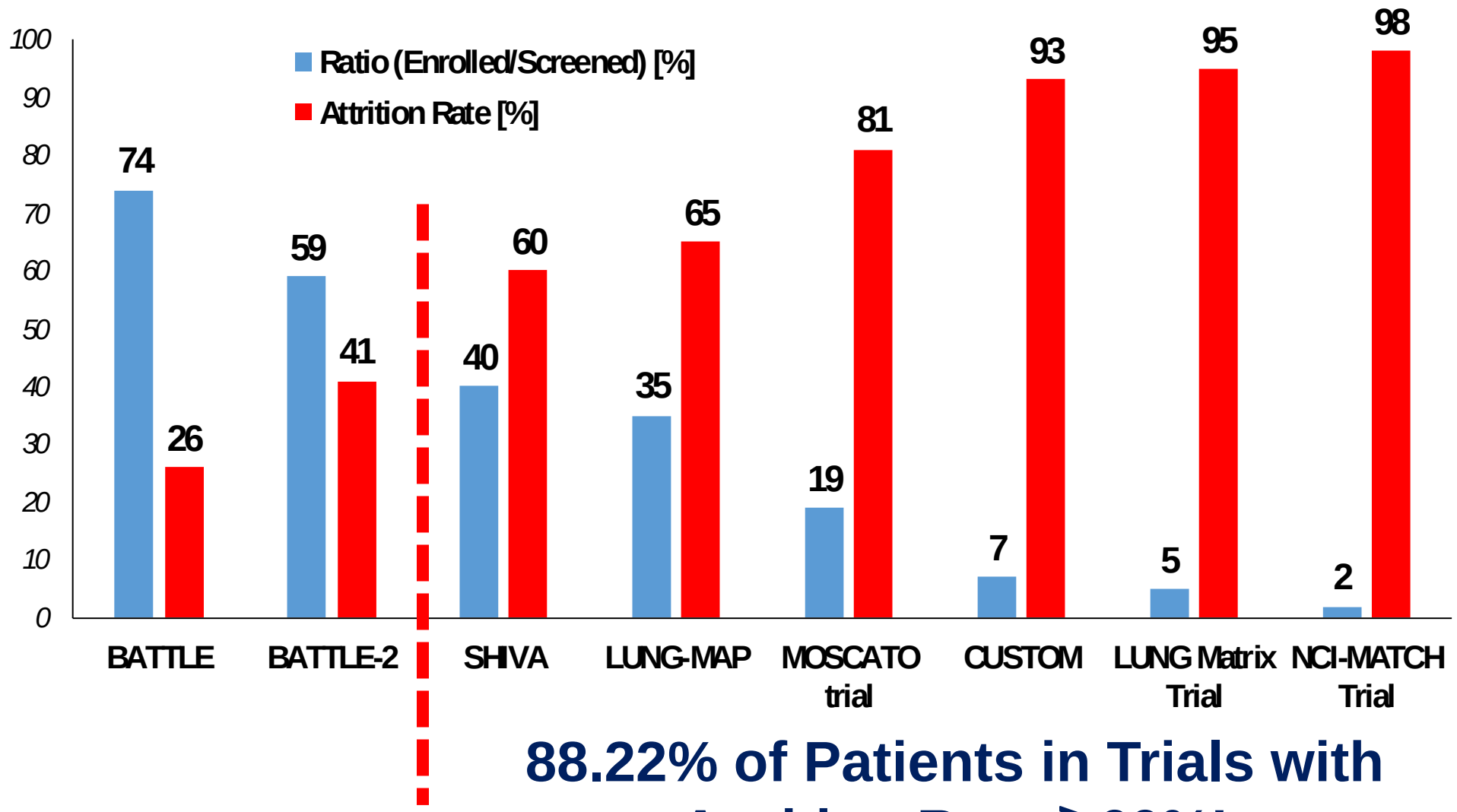


Tredan O et al, ASCO 2017, Ann Oncol 2019

The use of Molecularly Targeted Agents/Therapies [MTA/MTT] outside their indications does not improve outcome compared with standard in heavily pretreated patients with cancer.

Systematic Review of Precision Oncology Studies

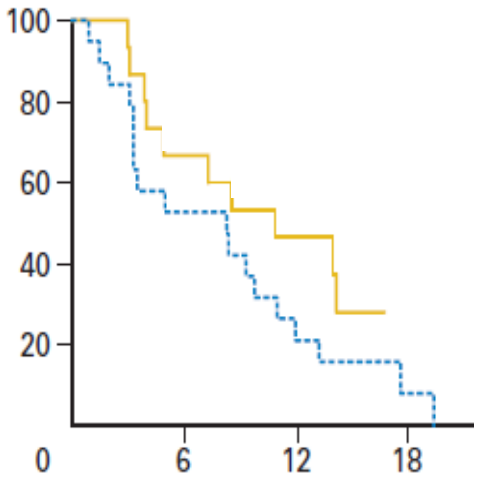
Study	Screened Pts (n.)
LUNG Matrix Trial	5467
LUNG-MAP	1864
MOSCATO trial	1035
NCI-MATCH Trial	795
SHIVA	741
CUSTOM	647
BATTLE	341
BATTLE-2	334



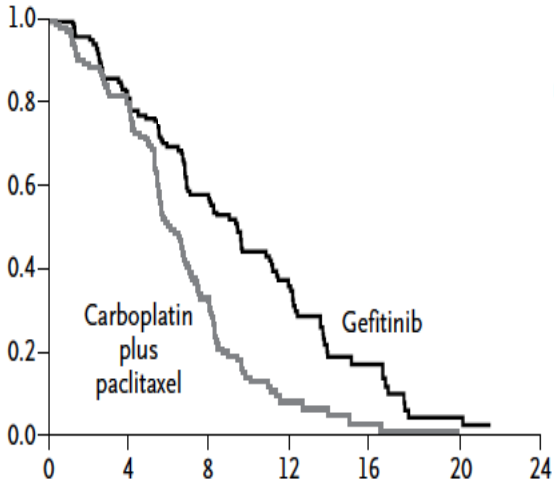
88.22% of Patients in Trials with Attrition Rate ≧60%!

Why Attrition Rate is So Relevant for Survival Analysis of Target Agents in Biomarker-Enriched Population?

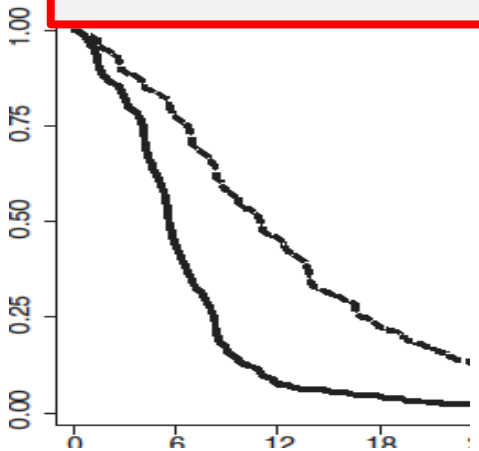
BR.21: ERL *vs.* PLB



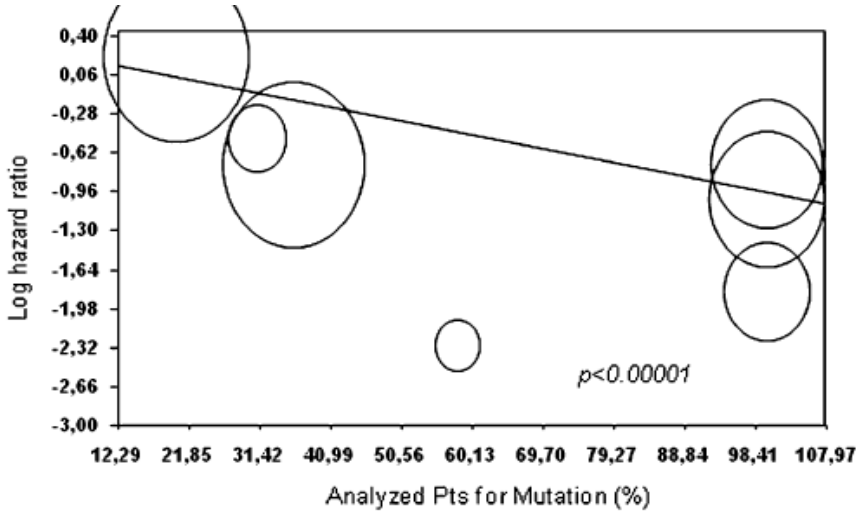
IPASS: GEF *vs.* Chemo



IPD-Meta-Analysis: TKIs *vs.* Chemo



Metaregression Analysis correlating PFS with the rate of patients analyzed for the EGFR mutation (Meta-Analysis of 6227 patients in 11 RCTs; EGFR TKIs *vs.* Chemo)

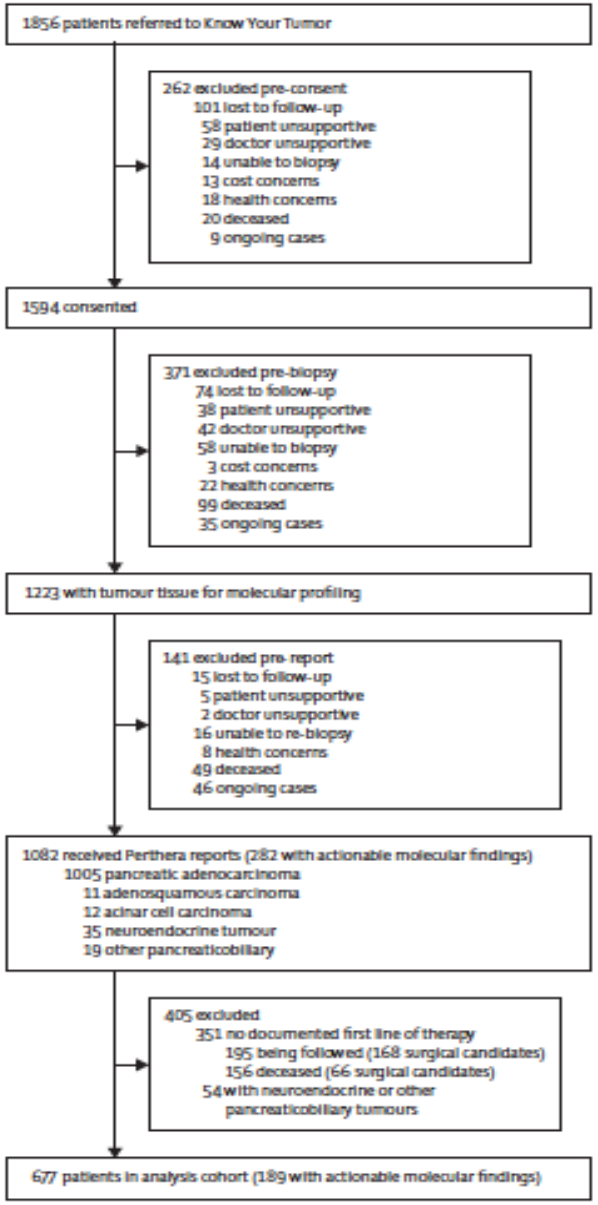


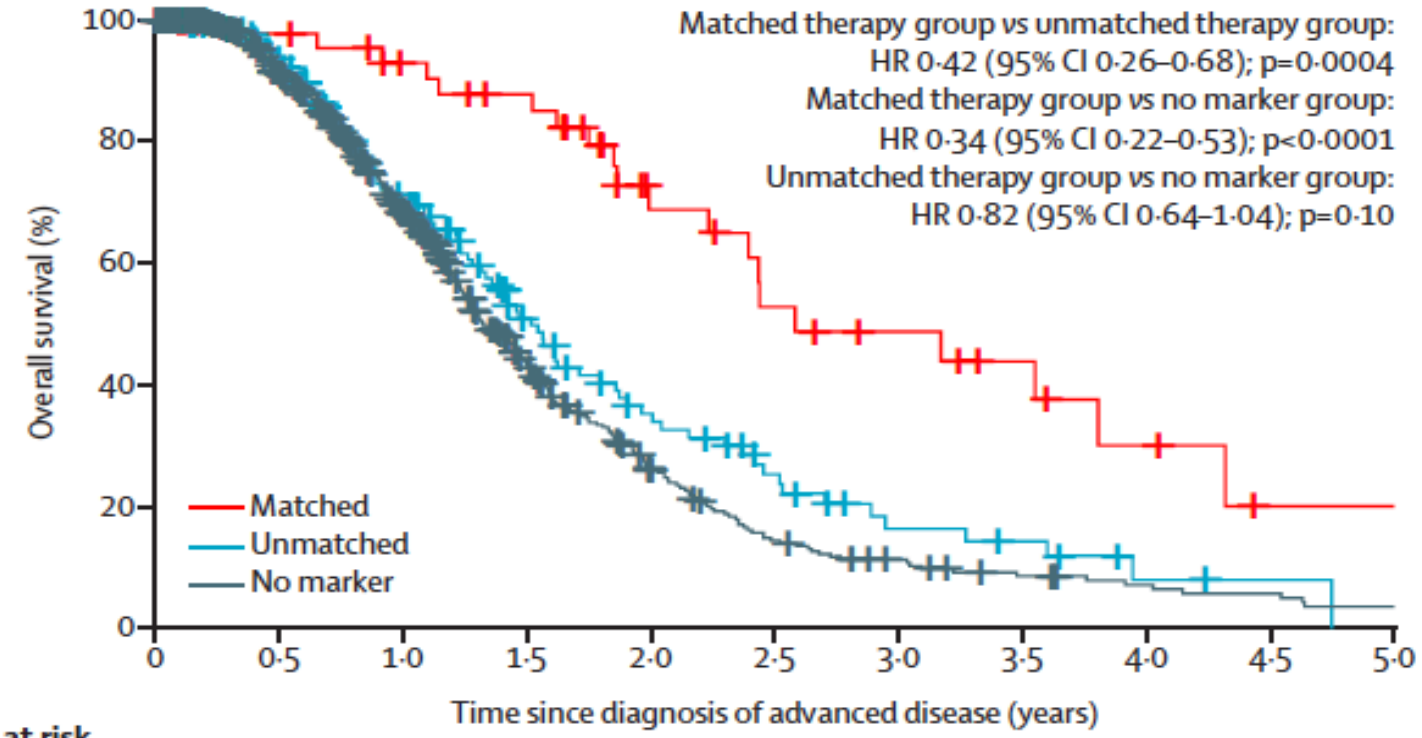
Pts Analyzed	**20%**	**36%**	**100%**
EGFR+	**3%**	**20%**	**100%**
Attrition Rate	**80%**	**67%**	**0%**

The Higher the Number of Patients Analyzed (*i.e. the lower the Attrition*) the lower the HR (i.e. the larger the benefit of TKIs)

Wheatley-Price P et al, JCO 2008 *Mok T et al, NEJM 2009* *Lee CK et al, JNCI 2017* *Bria E et al, Ann Oncol 2011*

Know Your Tumor Registry (NYTR) Trial: Advanced Pancreatic Cancer





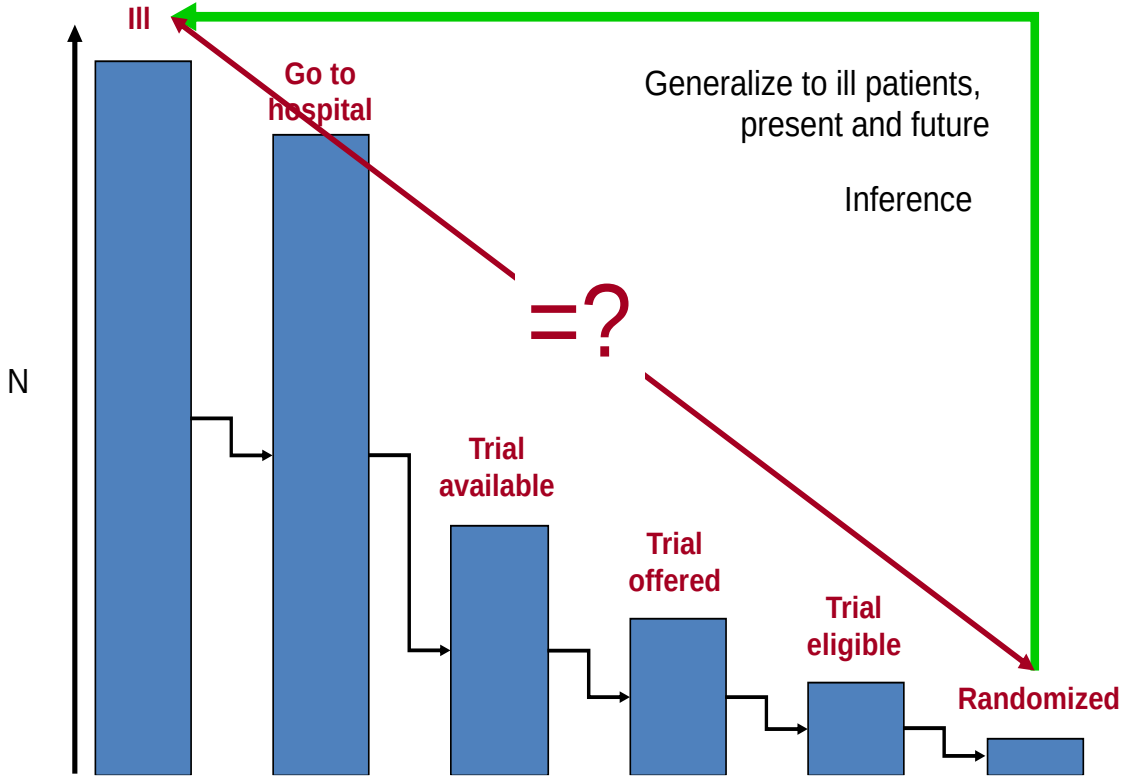
Number at risk (number censored)											
Matched therapy	46 (0)	42 (3)	36 (4)	32 (2)	18 (8)	13 (1)	10 (2)	7 (2)	4 (1)	1 (2)	1 (0)
Unmatched therapy	143 (0)	116 (19)	78 (11)	44 (15)	27 (4)	16 (4)	8 (3)	6 (1)	2 (2)	1 (1)	0 (0)
No marker	488 (0)	384 (66)	241 (55)	124 (39)	63 (15)	33 (4)	22 (4)	14 (3)	10 (2)	8 (0)	5 (0)

Best Scenario, <u>Attrition: 63%</u>

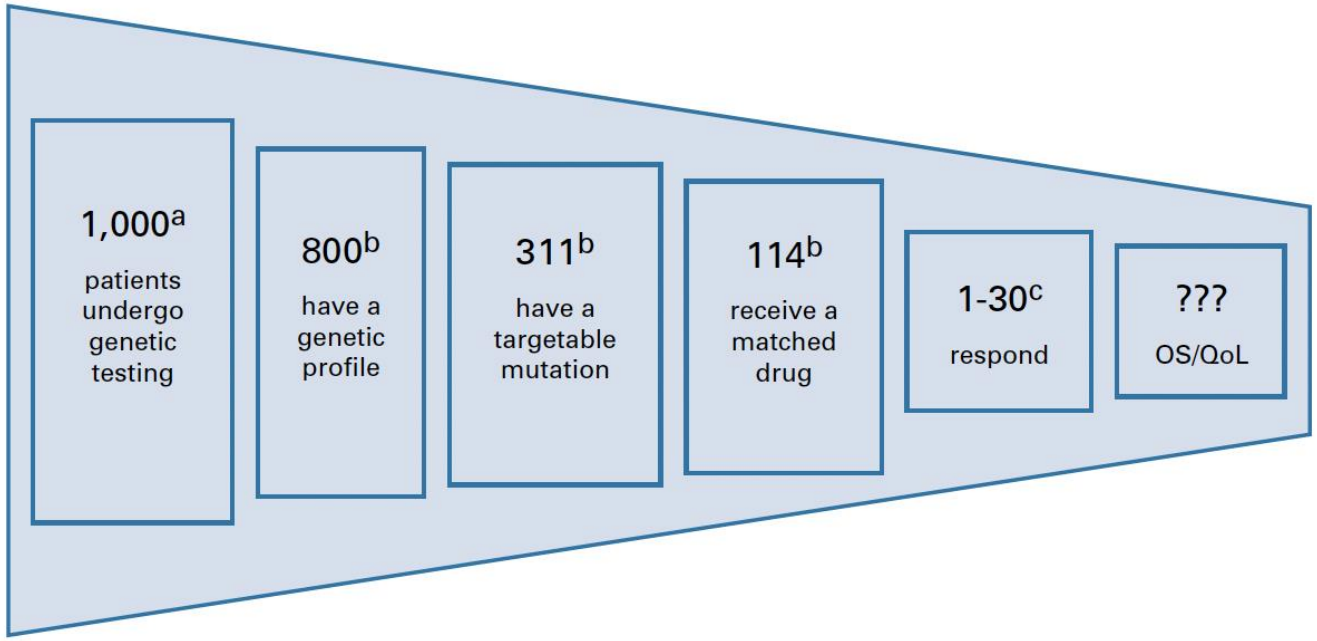
Pishvaian MJ et al, Lancet Oncol 2020

Generalizability of Data from RCTs/Basket & PM Trials

External Validity?



'The shrinking denominator'



Representative results derived from a large series of patients who have undergone genetic testing and where an attempt has been made to treat with a drug targeting a genetic mutation.

Courtesy of Di Maio M & Perrone F

Chan KKW et al, J Clin Oncol 2021

- Etica e integrità della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- **Disegni di studio osservazionali e interventistici**
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

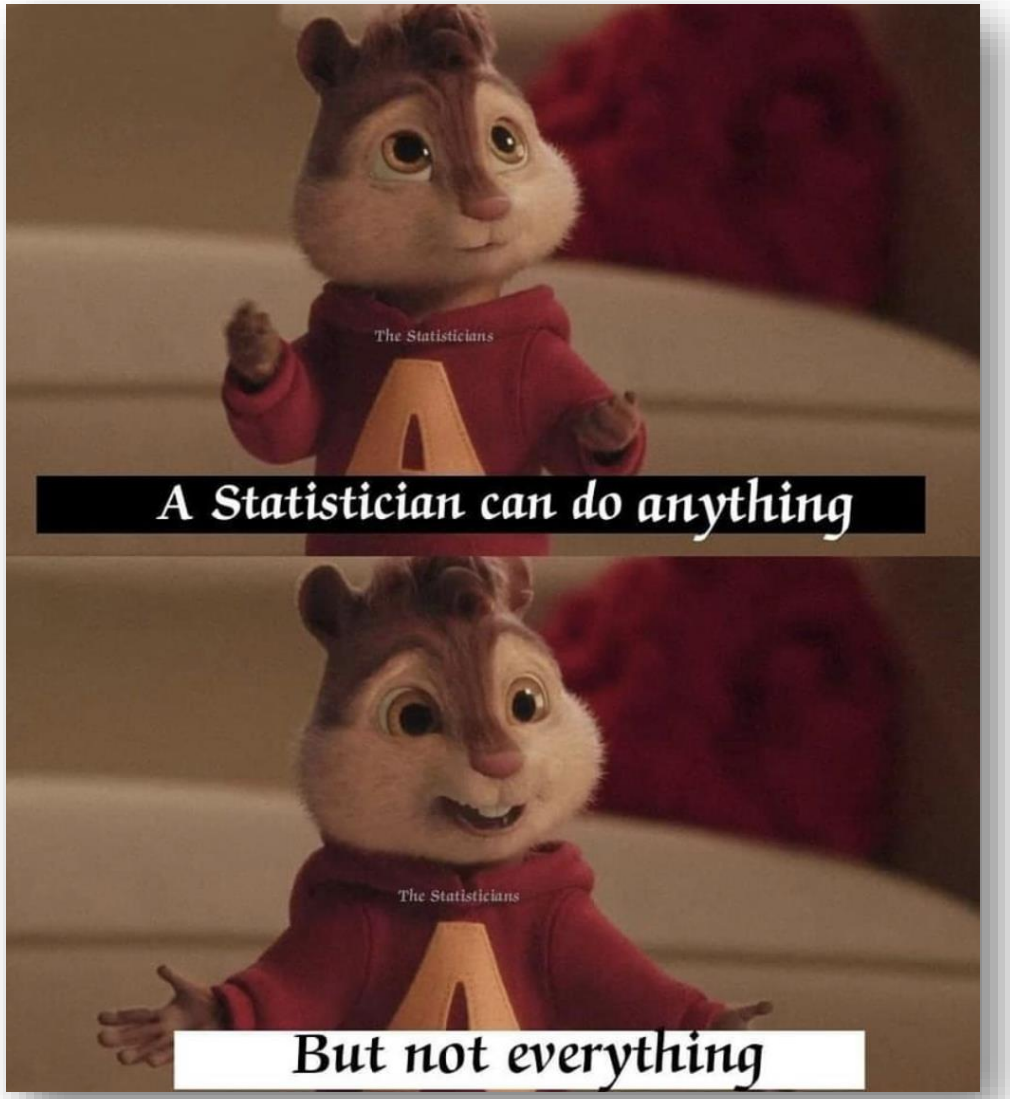
Dall'Ipotesi al Disegno di una Sperimentazione Clinica

- **Premessa: la buona ricerca clinica**
- **L'importanza <u>del disegno di studio</u>**
  - Descrizione dei dati e inferenza statistica
  - L'importanza della numerosità campionaria
  - La rappresentatività del campione
  - La tentazione delle analisi per sottogruppi

E' la struttura operativa dello studio, che permette di giungere a risultati:

1. **Credibili**
2. **Riproducibili**
3. **Generalizzabili**

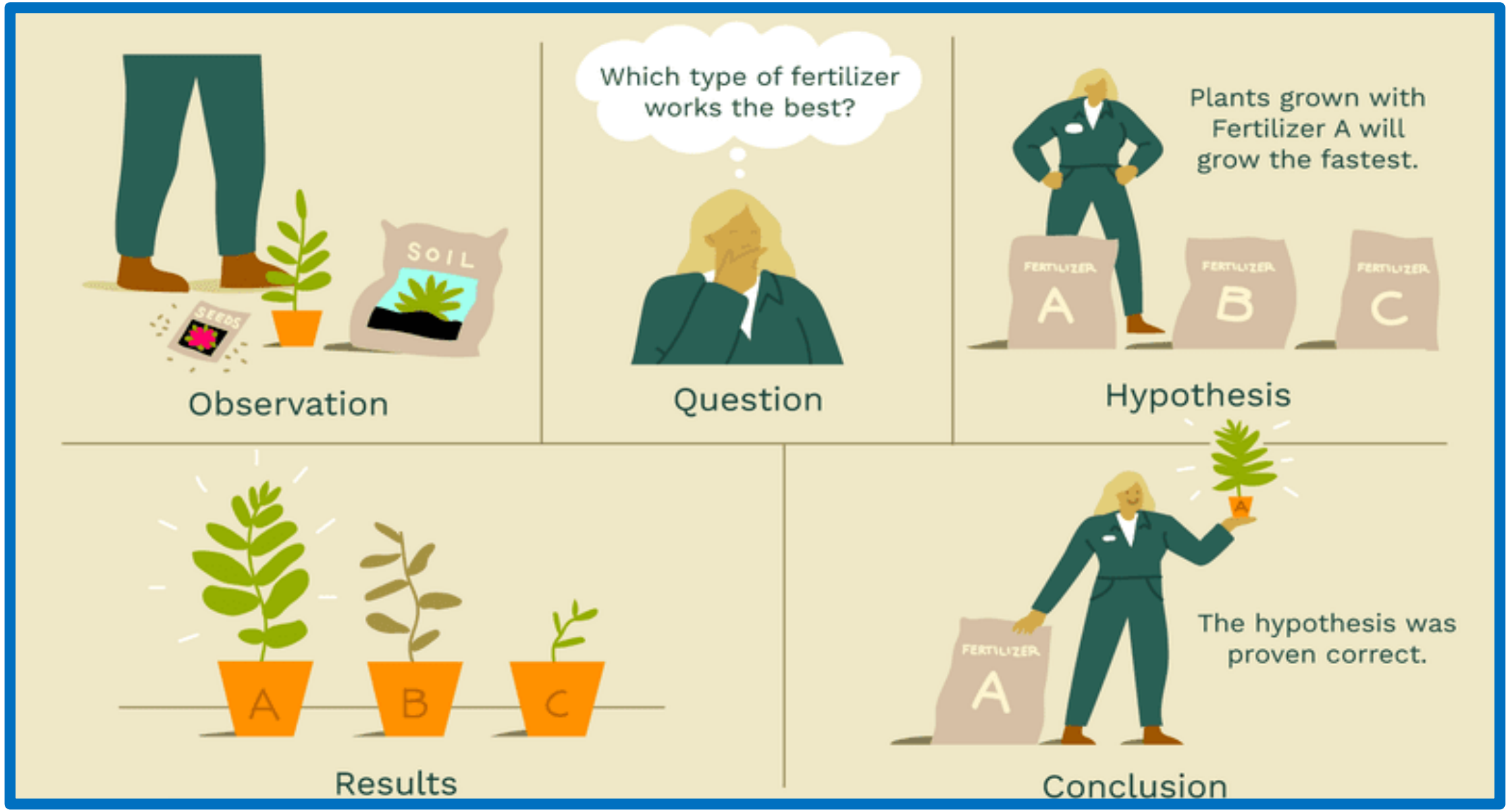
L'analisi Dovrebbe Essere Coerente Con Il Disegno Dello Studio!



'If you torture your data long enough, they will tell you whatever you want to hear'

Mills, N Engl J Med 1993

Steps of the Scientific Method



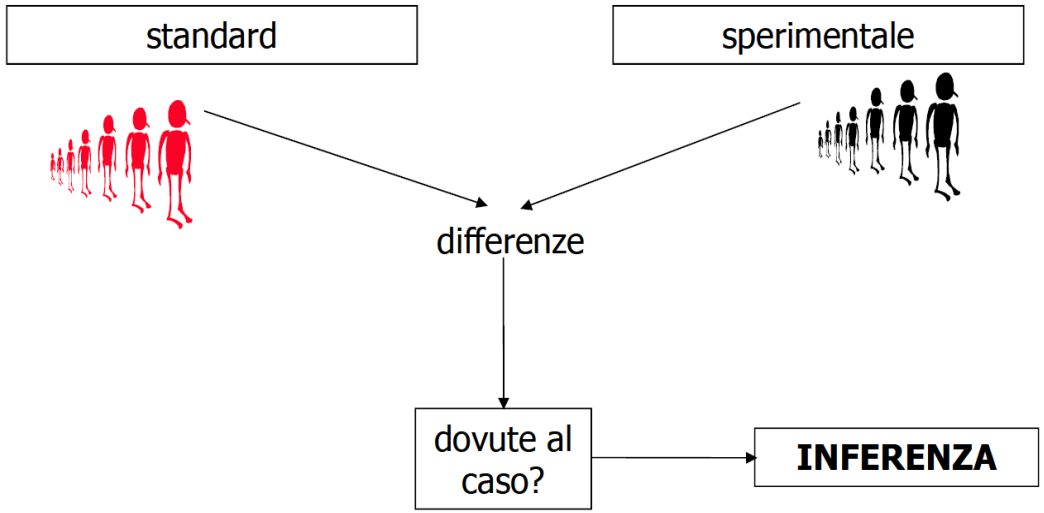
Dall'Ipotesi al Disegno di una Sperimentazione Clinica

- **Premessa: la buona ricerca clinica**
- **L'importanza del disegno di studio**
  - **Descrizione dei dati e inferenza statistica**
  - L'importanza della numerosità campionaria
  - La rappresentatività del campione
  - La tentazione delle analisi per sottogruppi

- **Statistica descrittiva: prende in considerazione gli aspetti di presentazione (es. tabelle e grafici) dei dati**
- **Statistica inferenziale: deriva conclusioni riguardanti le popolazioni a partire dallo studio di un campione**

Courtesy of Valter Torri

Inferenza Statistica



Ipotizziamo che in un campione di pazienti la % risposte del trattamento A sia migliore rispetto a quella di B

Come interpretiamo questa differenza?

3 possibili ipotesi per giustificare la diversità osservata:

1. **Il confronto è viziato perché qualche fattore non considerato è responsabile della differenza osservata**
2. **La differenza è dovuta alla variabilità campionaria**
3. **A è effettivamente più attivo di B**

L'ipotesi 1 può essere esclusa disegnando bene l'esperimento

L'ipotesi 2 può essere esclusa applicando un test di significatività statistica

Soltanto dopo avere escluso le ipotesi 1 e 2 possiamo concludere che A è migliore di B

Studi Osservazionali *vs.* Studi Sperimentali

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

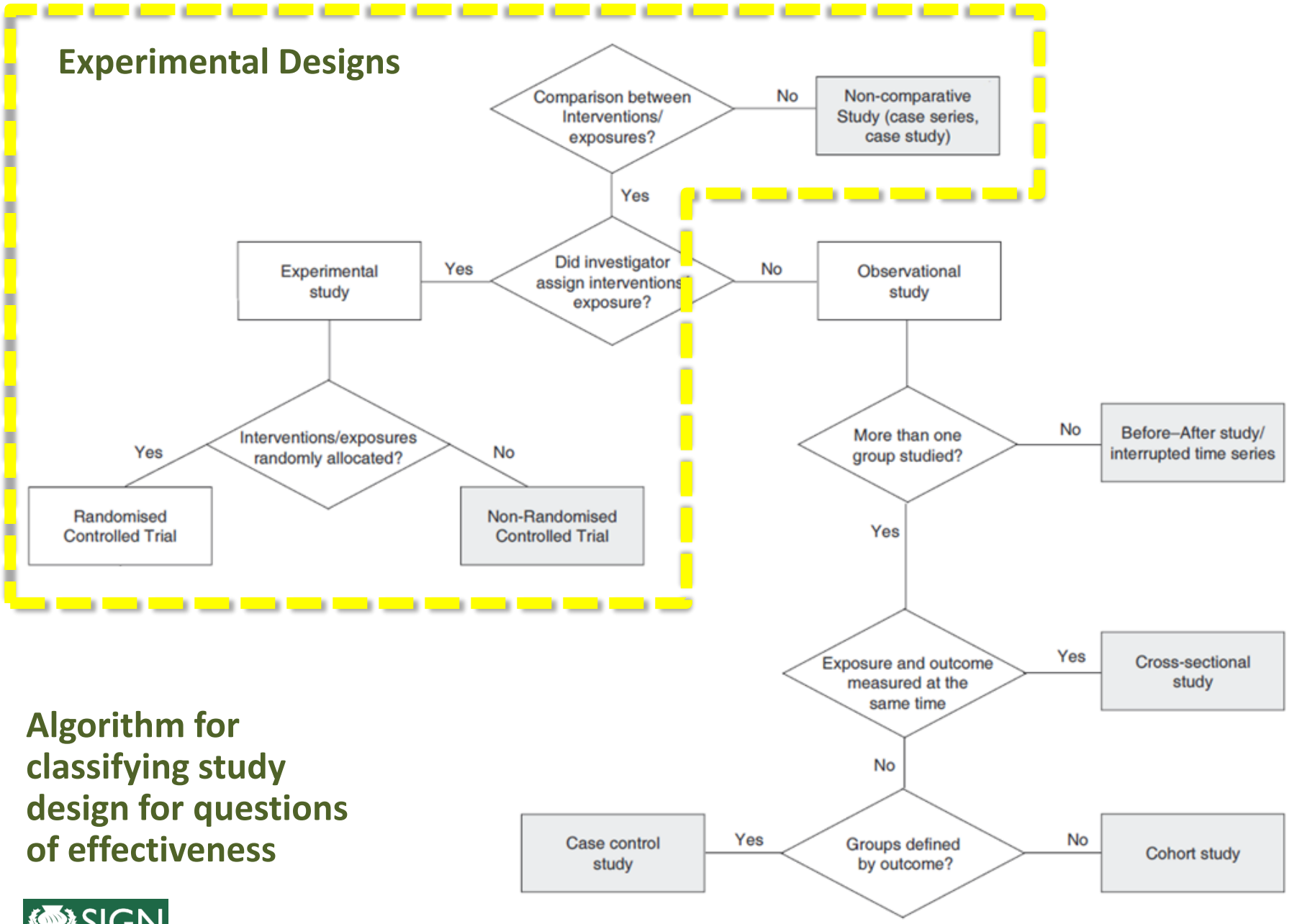
- The term observational study covers a wide range of study designs, a common feature of which is that they are non-interventional, in the sense that the study protocol does not determine the precise features of any therapy given to the participants in the study.

Studi non controllati

Studi controllati non randomizzati

Studi controllati randomizzati

Experimental Designs



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

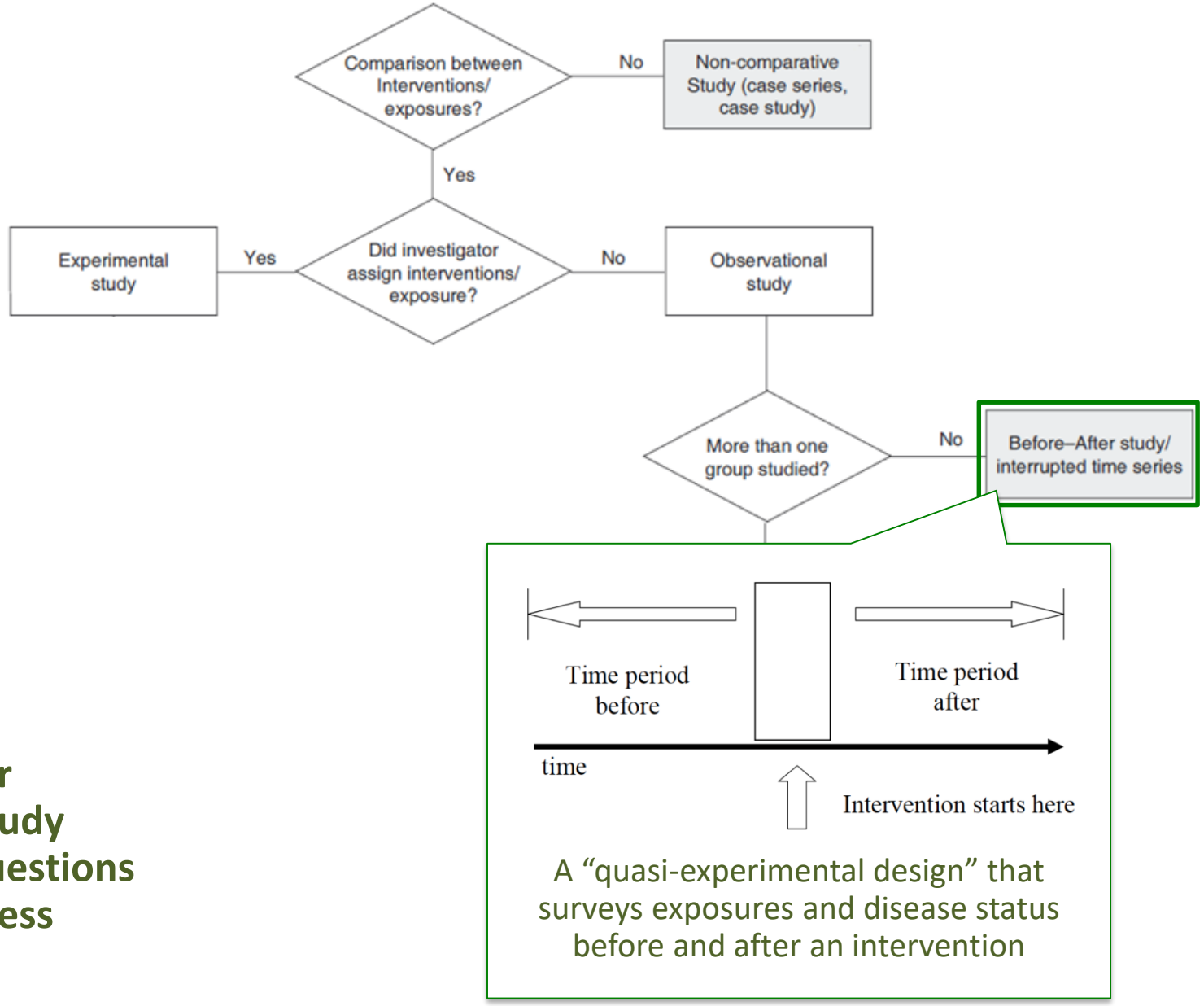
Identify the incidence of a particular outcome over time

Non-comparative case series

Report outcomes of patients who received a specific intervention

Comparative case series

Compare outcomes between patients who received different interventions



A "quasi-experimental design" that surveys exposures and disease status before and after an intervention

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

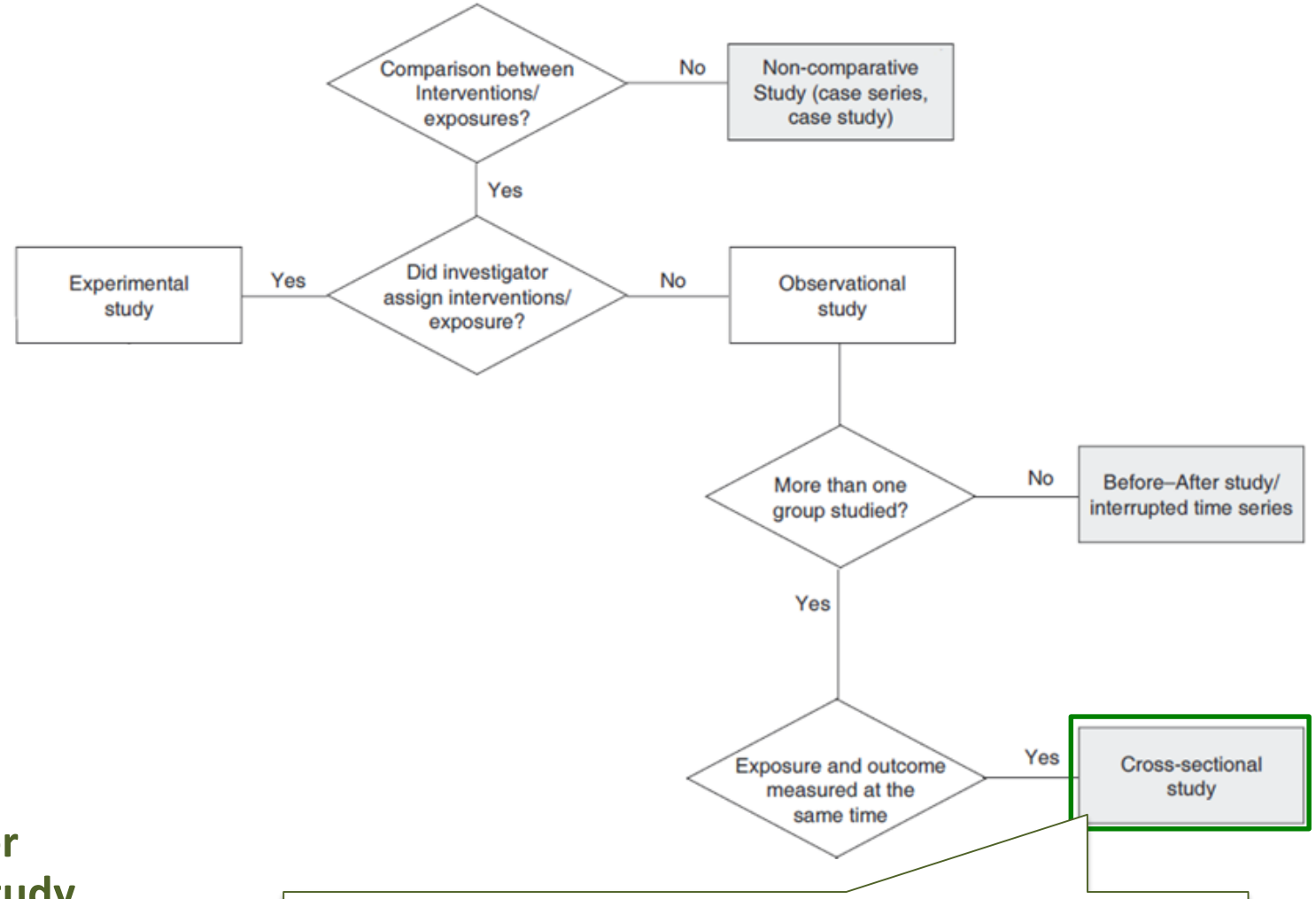
Non-comparative case series

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

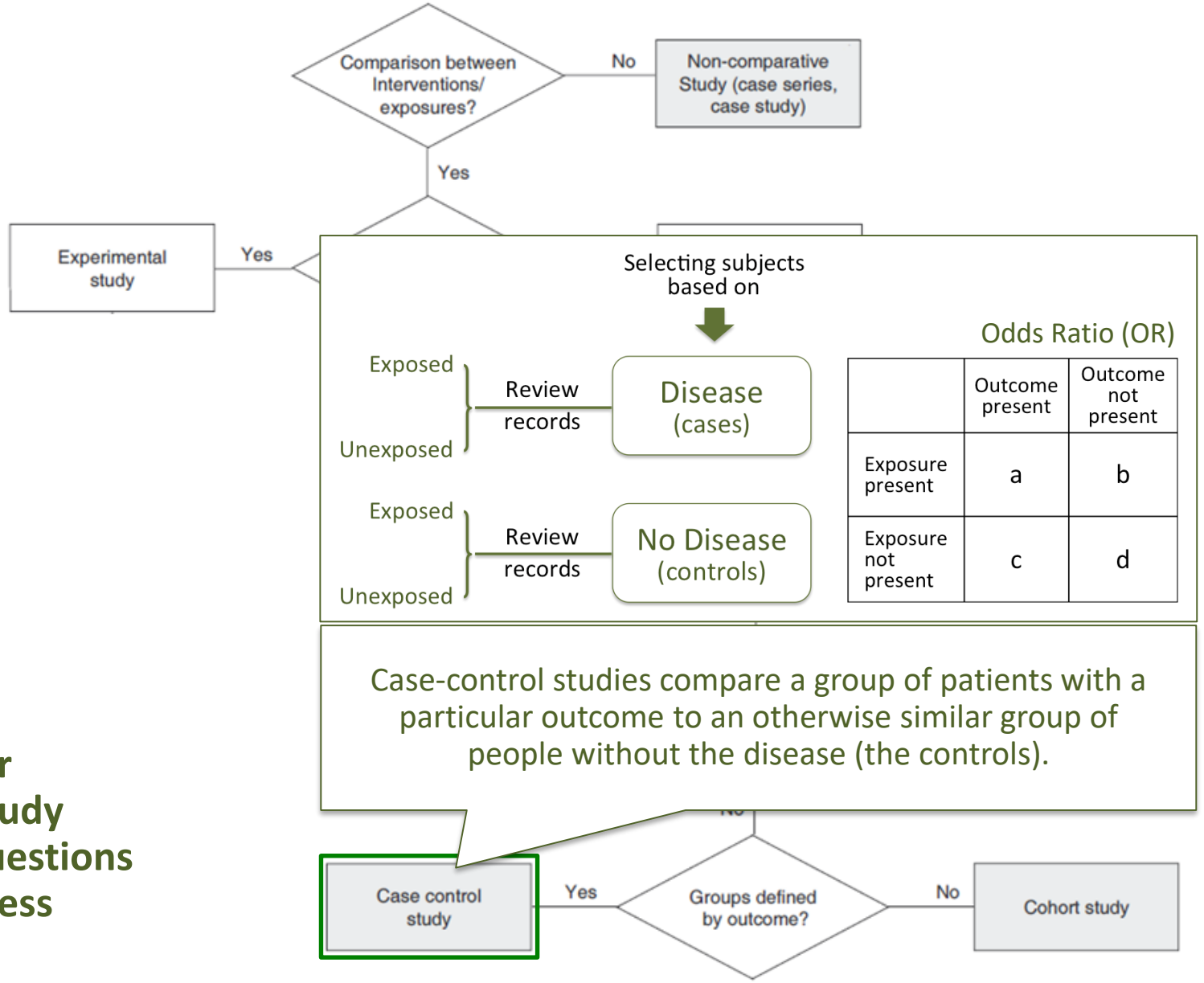
Non-comparative case series

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

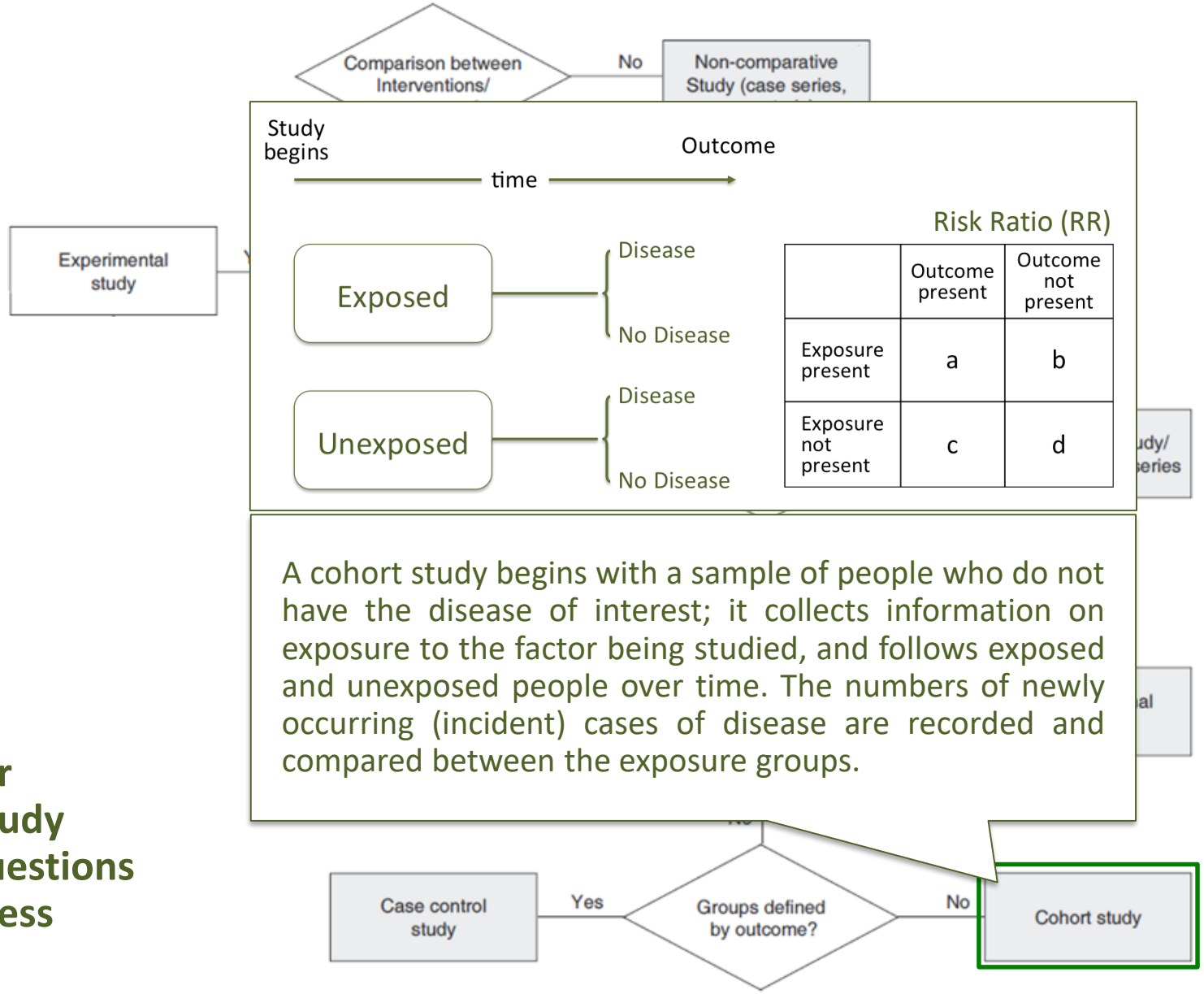
Non-comparative case series

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 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

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

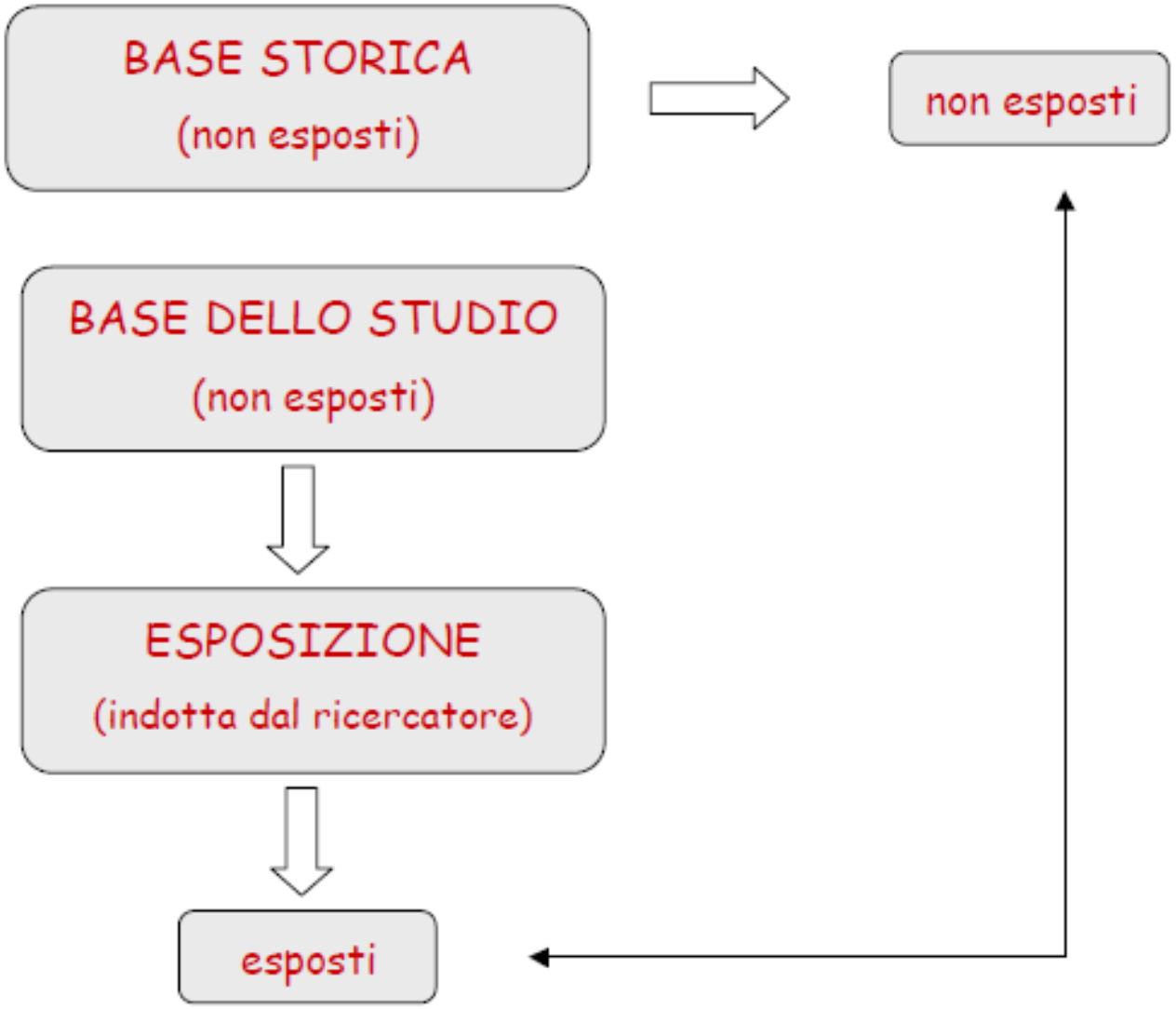
Report outcomes of patients who received a specific intervention

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.

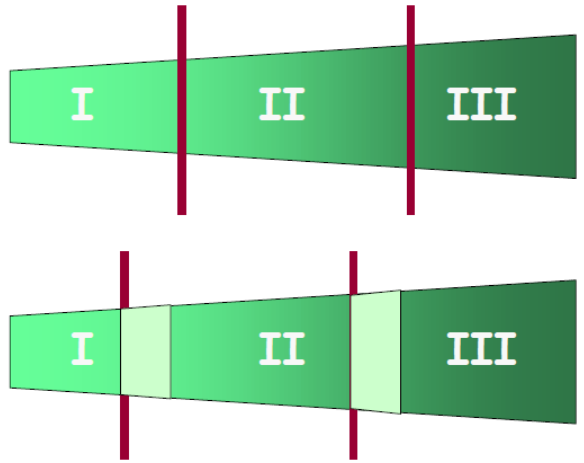
Studi non controllati (*es. controlli storici*)



Problemi

- Variabilità del decorso
- Selezione della popolazione
- Effetto “placebo”

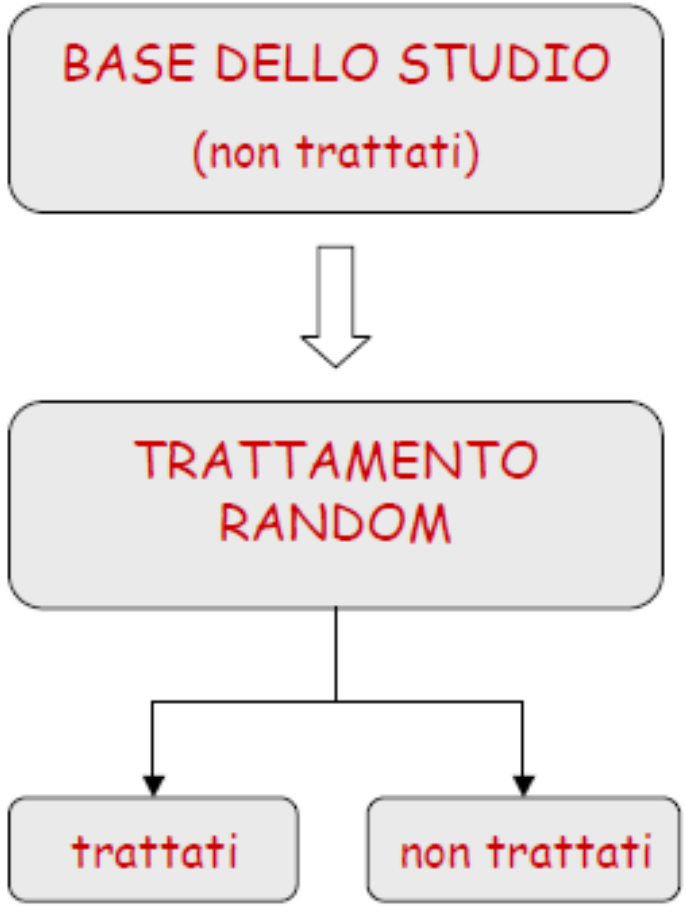
La migrazione di stadio: il fenomeno di Will Rogers



Studi Controllati

Studi con controlli paralleli

Studi randomizzati



Dall'Ipotesi al Disegno di una Sperimentazione Clinica

- **Premessa: la buona ricerca clinica**
- **L'importanza del disegno di studio**
  - **Descrizione dei dati e inferenza statistica**
  - **<u>L'importanza della numerosità campionaria</u>**
  - **La rappresentatività del campione**
  - **La tentazione delle analisi per sottogruppi**

La Precisione della Stima

Numeri assoluti	Proporzione Osservata	Intervallo di confidenza 95%
2/10	20%	0.1% - 44.8%
4/20	20%	2.5% - 37.5%
6/30	20%	5.7% - 34.3%
8/40	20%	7.6% - 32.4%
10/50	20%	8.9% - 31.1%
12/60	20%	9.9% - 30.1%
14/70	20%	10.6% - 29.4%
16/80	20%	11.2% - 28.8%
18/90	20%	11.7% - 28.3%
20/100	20%	12.2% - 27.8%
50/250	20%	15.0% - 25.0%

Fattori che influenzano la numerosità del campione in una sperimentazione clinica

Importanza della Numerosità Campionaria

- Immaginiamo di voler verificare la superiorità di un trattamento sperimentale rispetto allo standard
- Fissati il rischio di risultato falso positivo (5%) e la potenza (80%), il numero di eventi necessari dipende dall'entità dell'impatto prognostico che ipotizziamo (**Hazard Ratio**)

Hazard Ratio (HR)	Numero di eventi necessari
0.80	631
0.70	247
0.60	121
0.50	66
0.40	38
0.30	22
0.20	13
0.10	6

Numero di eventi	Potenza
20	20%
30	28%
40	36%
50	44%
60	51%
70	57%
80	63%
90	68%
100	72%
110	76%
120	80%

SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

- Etica e integrità della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- **Misure di effetto relativo e assoluto**
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

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

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

VARIABILE DI RISPOSTA

- di tipo **qualitativo (nominale)**
  - esprime categorie di risposta del tipo successo / insuccesso (di un trattamento somministrato).
- di tipo **quantitativo**
  - 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.

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

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

Risk and Odds

- **Risk** (proportion of persons with disease = cumulative incidence)
  - Risk Ratio = ratio of 2 cumulative incidence estimates = tive Risk
  - bsolute difference of 2 cumulative
  - event will or will not be observed)
  - Odds Ratio = ratio of 2 odds

$$\frac{\text{number of new events during the specified period}}{\text{number of persons at risk during the specified period}}$$

Risk and Odds

- **Risk** (proportion of persons with disease = cumulative incidence)
  - **Risk Ratio** = ratio of 2 cumulative incidence estimates = Relat[illegible] Risk
  - [illegible]ute difference of 2 cumulative
- Odd[illegible]ent will or will not be observed)
  - [illegible]s

$$\frac{\text{Risk of event within experimental arm}}{\text{Risk of event within control arm}}$$

Risk and Odds

number of new events during the specified period

number of non events during the specified period

with disease = cumulative incidence)

cumulative incidence estimates =

olute difference of 2 cumulative

dence estimates

- **Odds** (the likelihood that an event will or will not be observed)
  - Odds Ratio = ratio of 2 odds

Risk and Odds

- **Risk** (proportion of persons with disease = cumulative incidence)
  - R… …umulative incidence estimates =
    R…
  - R… …te difference of 2 cumulative
    i…

$$\frac{\text{Odds of event within experimental arm}}{\text{Odds of event within control arm}}$$

- **Odds** (th… …kelihood that an event will or will not be observed)
  - **Odds Ratio** = ratio of 2 odds

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

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

Risk and Odds

- **Risk** (proportion of persons with disease = cumulative incidence)
  - **Risk Ratio** = ratio of 2 cumulative incidence estimates = Relative Risk
  - **Risk Difference** = absolute difference of 2 cumulative incid e estimates
- C l not be observed)

Risk of event within experimental arm – **Risk of event within control arm**

Riassumendo…

Misure di effetto relativo

Risk 70 risposte / 100 pazienti = 0.70

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

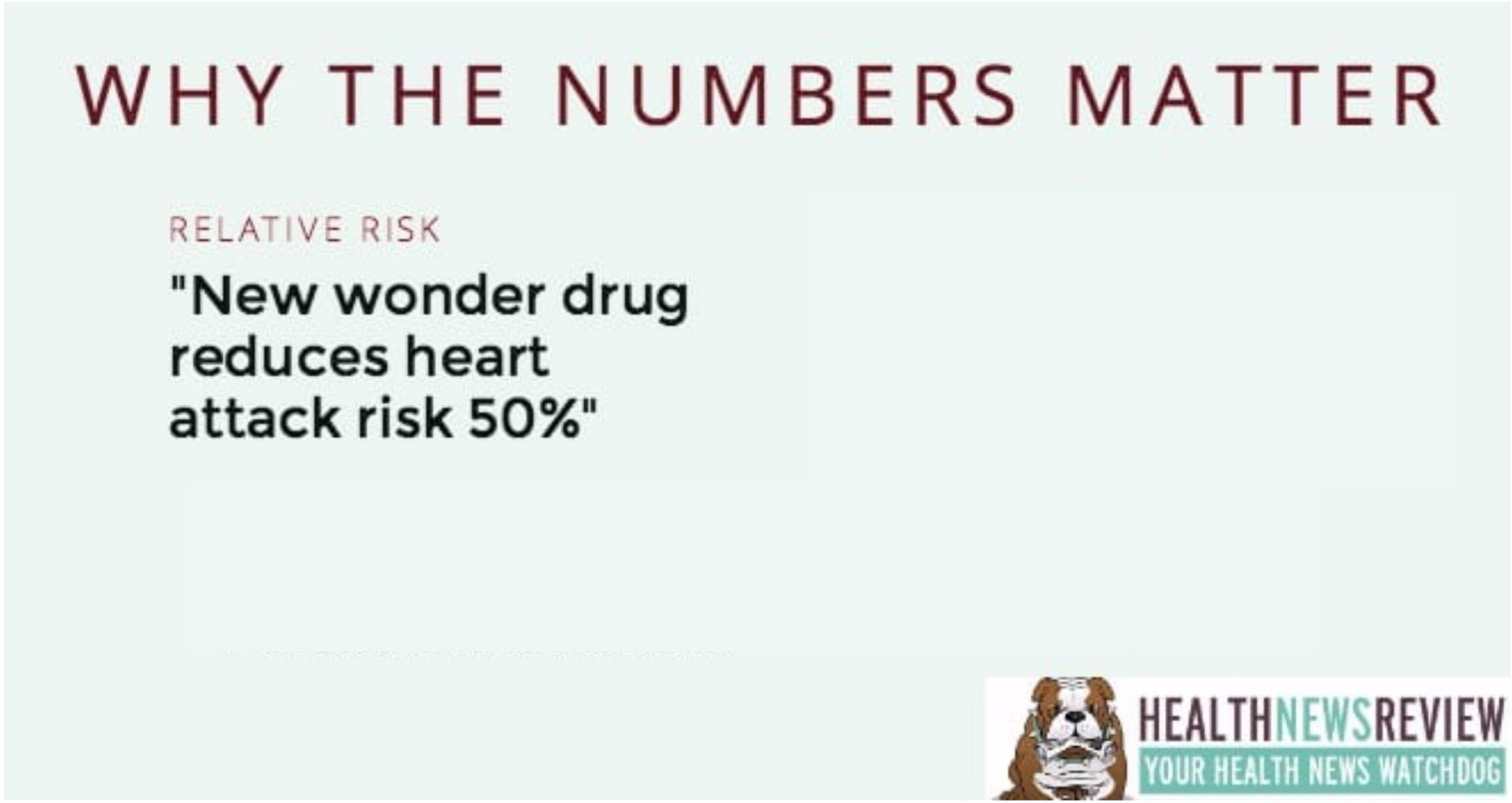
Odds 65 risposte / 35 non risposte = 1.86

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

Misura di effetto assoluto

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

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



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

WHY THE NUMBERS MATTER

RELATIVE RISK

"New wonder drug reduces heart attack risk 50%"

ABSOLUTE RISK

"New wonder drug reduced heart attacks from from 2 per 100 to 1 per 100"

The **absolute risk** is more useful at conveying the true impact of an intervention, yet is often under-reported in the research and the news.

HEALTHNEWSREVIEW
YOUR HEALTH NEWS WATCHDOG

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.

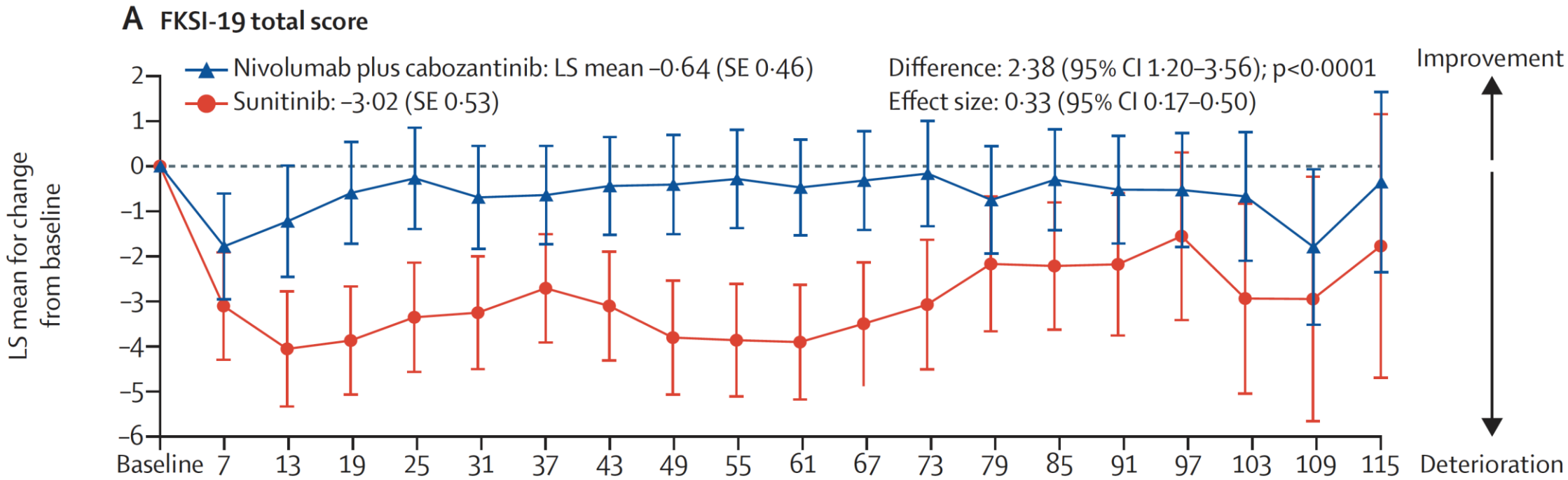
Variabili di comune riscontro negli studi clinici e indicatori comparativi di effetto

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

Patient-reported outcomes with first-line nivolumab plus cabozantinib versus sunitinib in patients with advanced renal cell carcinoma treated in CheckMate 9ER: an open-label, randomised, phase 3 trial

David Cella, Robert J Motzer*, Cristina Suarez, Steven I Blum, Flavia Ejzykowicz, Melissa Hamilton, Joel F Wallace, Burcin Simsek, Joshua Zhang, Cristina Ivanescu, Andrea B Apolo, Toni K Choueiri*

Lancet Oncol **2022; 23: 292–303**



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.

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

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

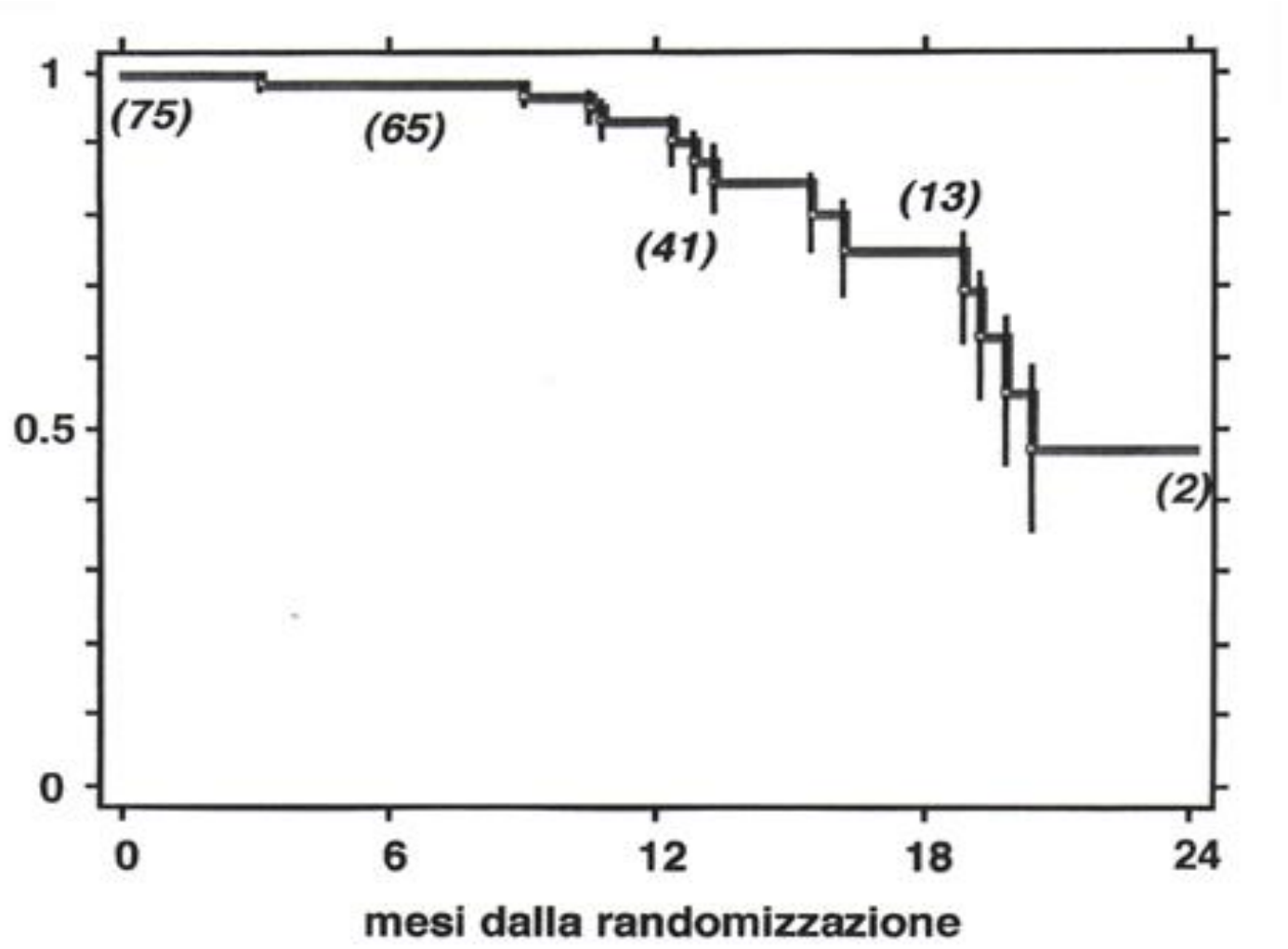
Variabile tempo-a-evento

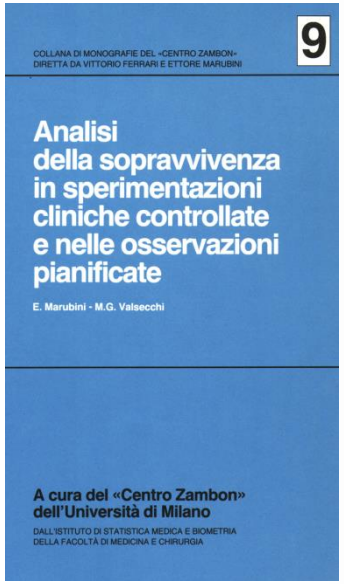
- Apparentemente assimilabile a una variabile di tipo quantitativo (intervallare)…
- …ma il verificarsi o meno di un evento la rende assimilabile a una variabile di tipo qualitativo (nominale)

Metodo di Kaplan-Meier

Stima della probabilità di sopravvivere in corrispondenza di ciascuno dei tempi in cui si verifica almeno un evento

CURVA DI SOPRAVVIVENZA

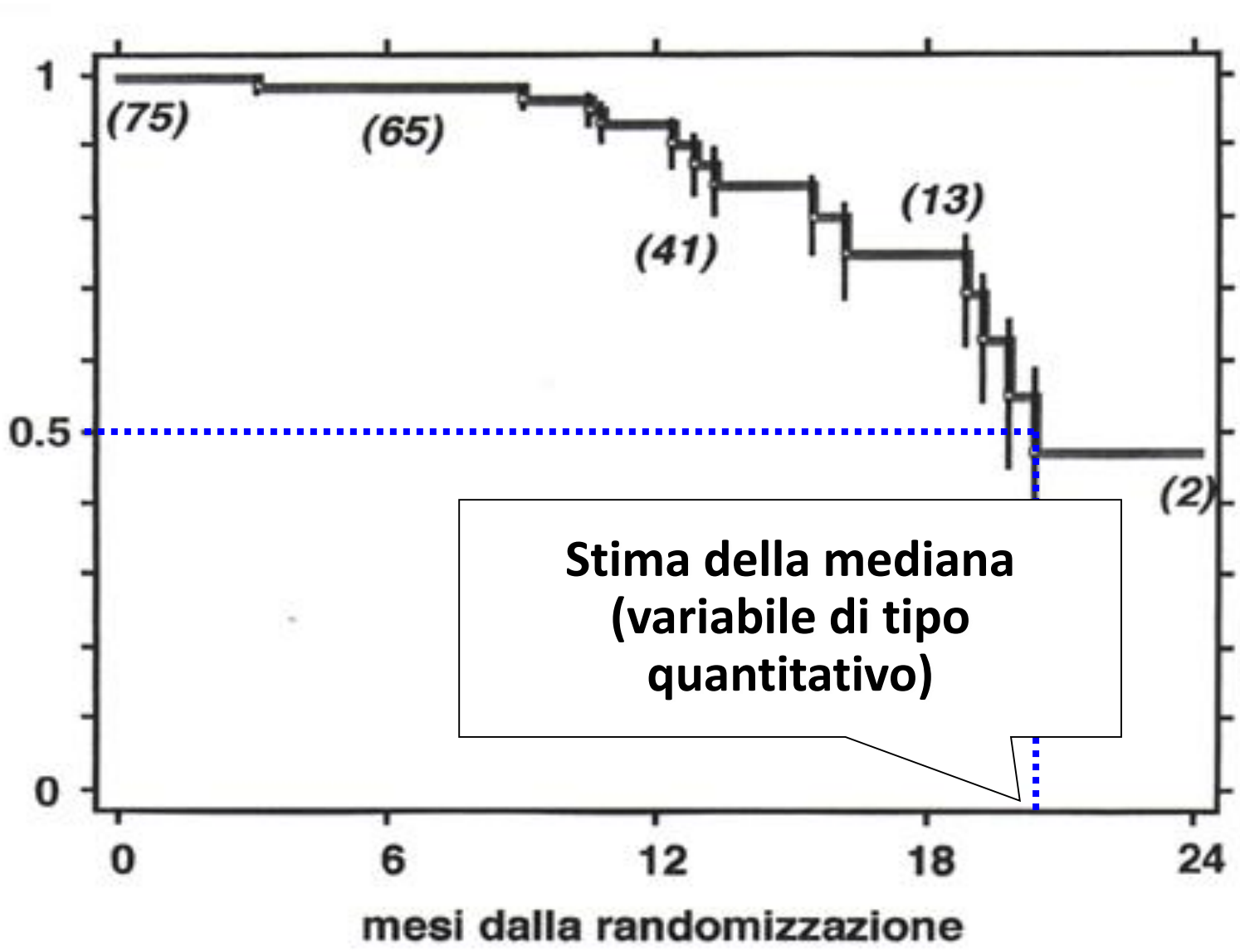


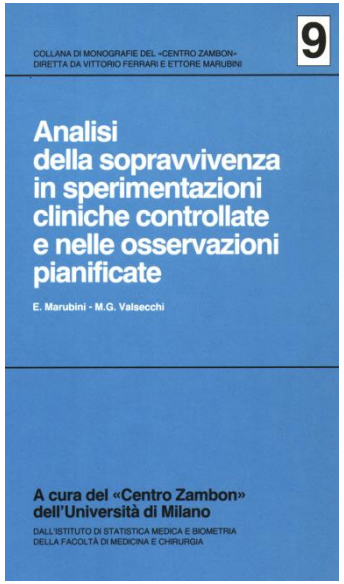


Tempi di risposta $t_{(j)}$	Tempi troncati* t^*	N° soggetti esposti a rischio n_j	N° eventi terminali d_j	Rischio istantaneo di "morte" $\hat{\lambda}(t_{(j)})$	Probabilità cumulativa di sopravvivere $t_{(j)}$ $\hat{P}_j$
9		20	1	1/20 = .050	(1 – 1/20) × 1 = .9500
13		19	1	1/19 = .053	(1 – 1/19) × .9500 = .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.

CURVA DI SOPRAVVIVENZA

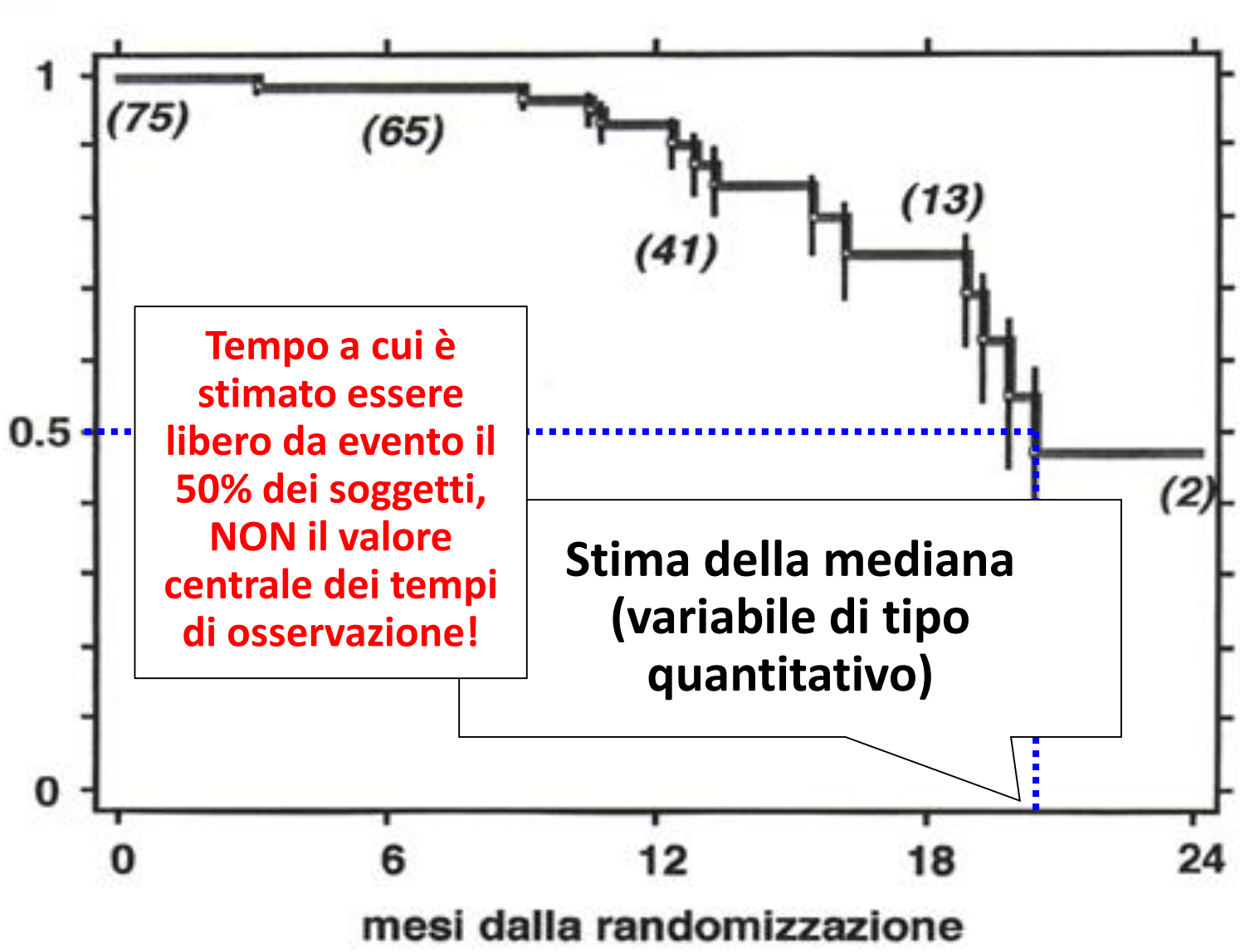




Tempi di risposta $t_{(j)}$	Tempi troncati* t^*	N° soggetti esposti a rischio n_j	N° eventi terminali d_j	Rischio istantaneo di "morte" $\hat{\lambda}(t_{(j)})$	Probabilità cumulativa di sopravvivere $t_{(j)}$ $\hat{P}_j$
9		20	1	1/20 = .050	(1 – 1/20) × 1 = .9500
13		19	1	1/19 = .053	(1 – 1/19) × .9500 = .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.

CURVA DI SOPRAVVIVENZA



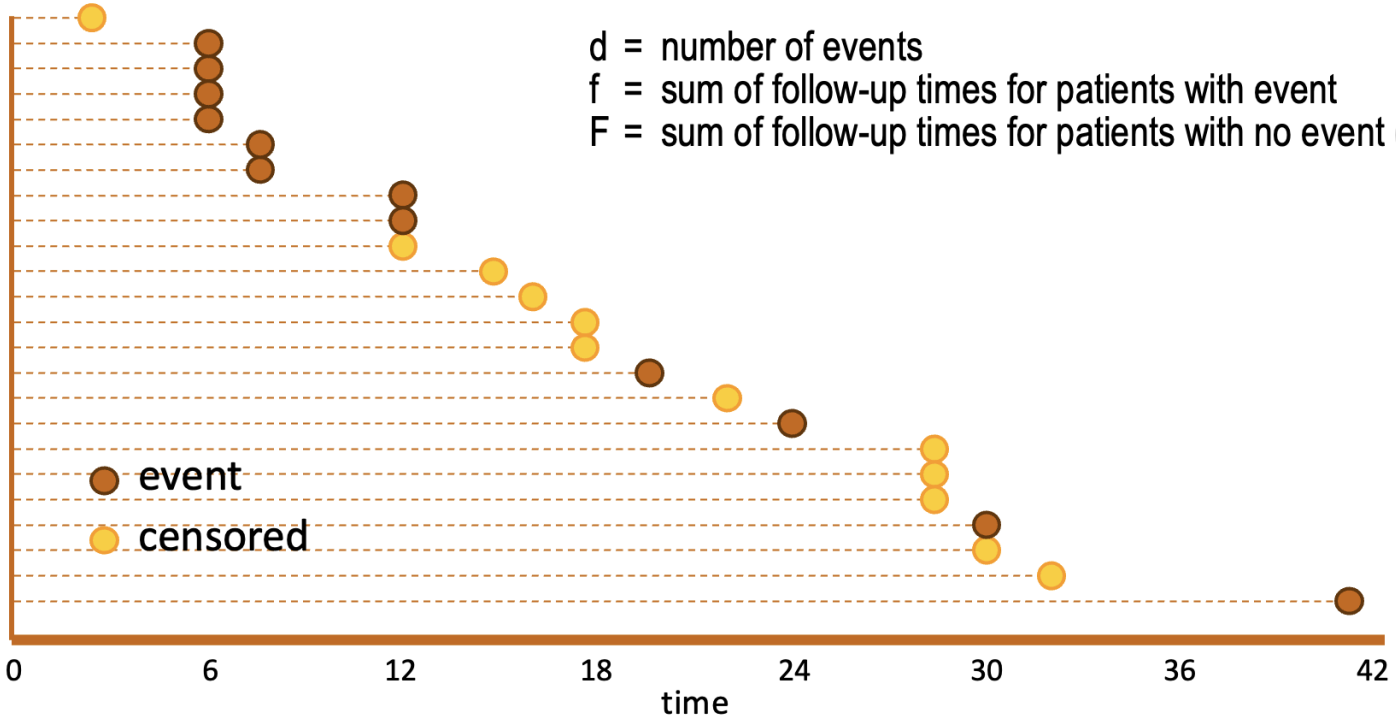
Variabili di comune riscontro negli studi clinici e indicatori comparativi di effetto

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

Defining a Hazard Ratio (HR)

- Compares risk of event in two populations or samples
- The ratio of risk (*hazard rate*) in experimental group to risk (*hazard rate*) in control group
  - *The Hazard Rate (λ)* is the rate at which events happen
- Assumption: *proportional hazards (PH),* i.e. the risk does not depend on time, that is, "risk is constant over time"

hazard rate



d = 12

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

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

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

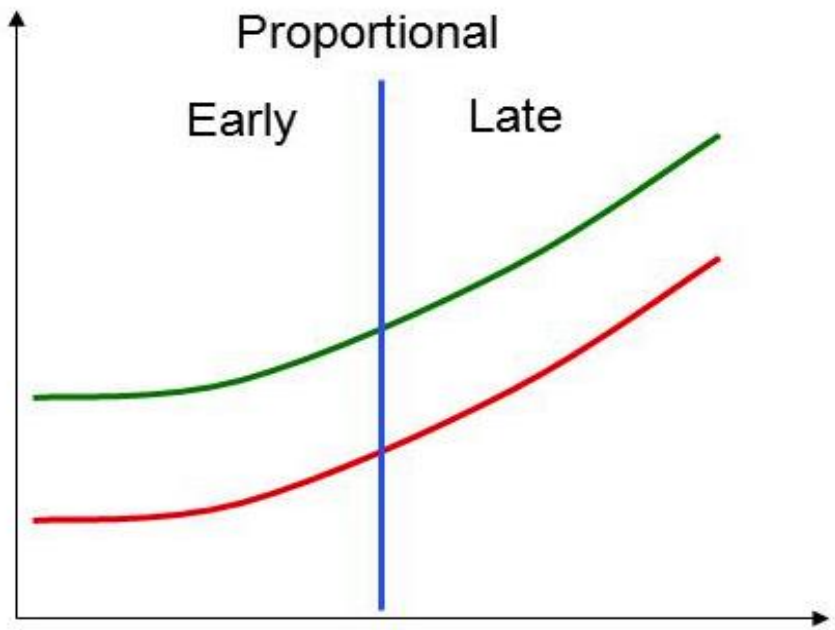
$$\textbf{Hazard Ratio} = \frac{\textit{hazard rate}\ \text{trattamento sperimentale}}{\textit{hazard rate}\ \text{trattamento di controllo}}$$

Defining a Hazard Ratio (HR)

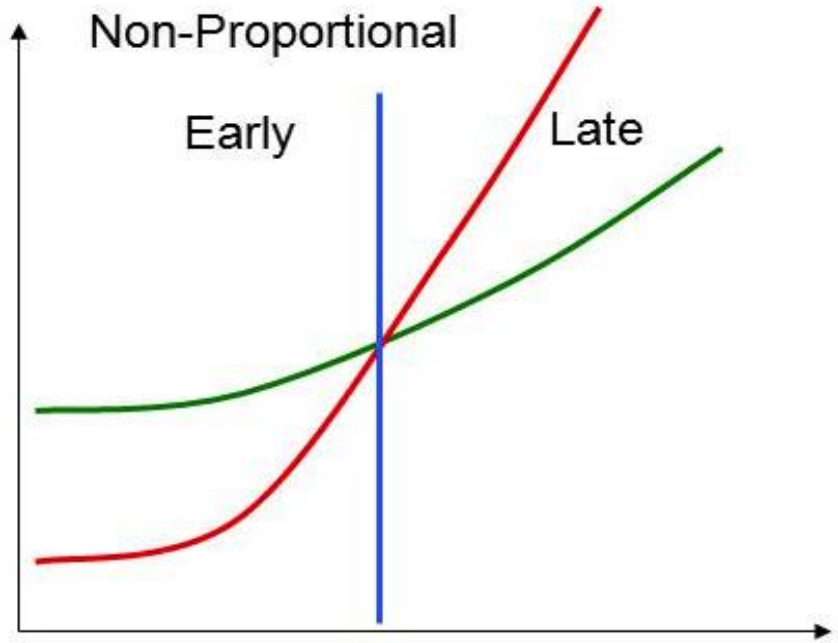
- Compares risk of event in two populations or samples
- The ratio of risk (*hazard rate*) in experimental group to risk (*hazard rate*) in control group
  - *The Hazard Rate (λ)* is the rate at which events happen
- Assumption: *proportional hazards (PH),* i.e. the risk does not depend on time, that is, "risk is constant over time"

Proportional Hazard Assumption

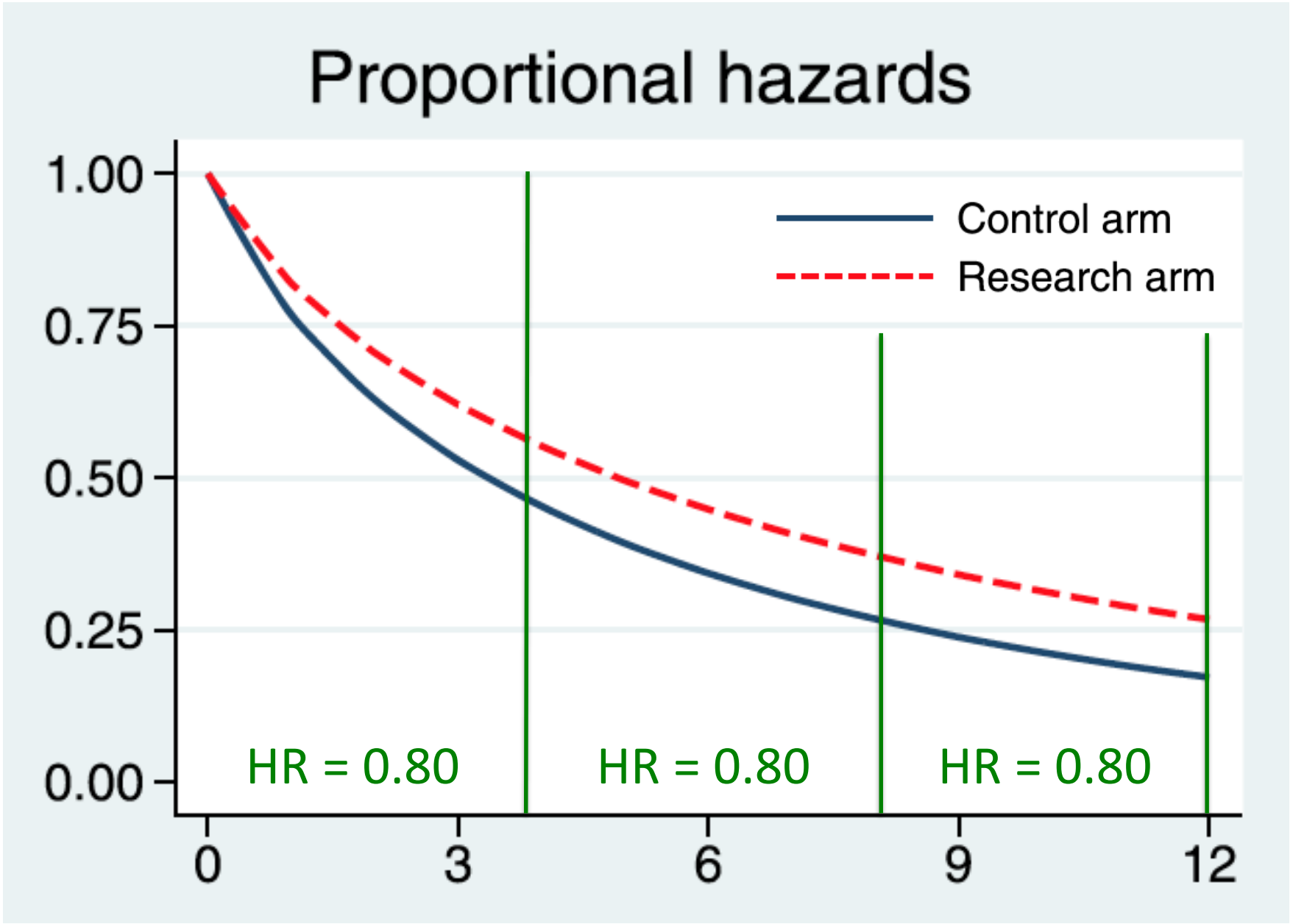
If we are comparing a new treatment with the standard treatment, it is assumed that the ratio of the hazard for an individual on a new treatment to that for an individual on the standard treatment remains constant over time

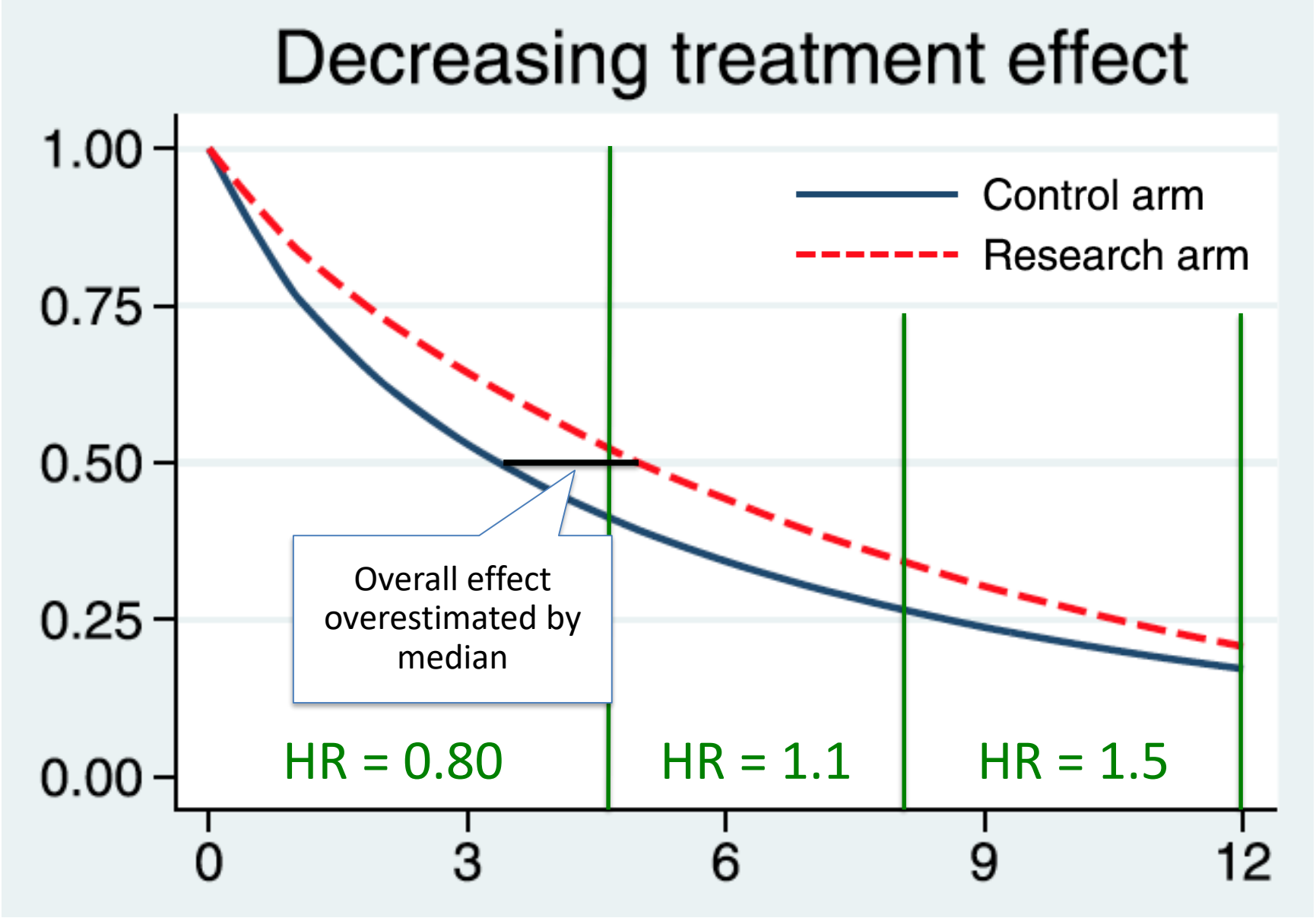


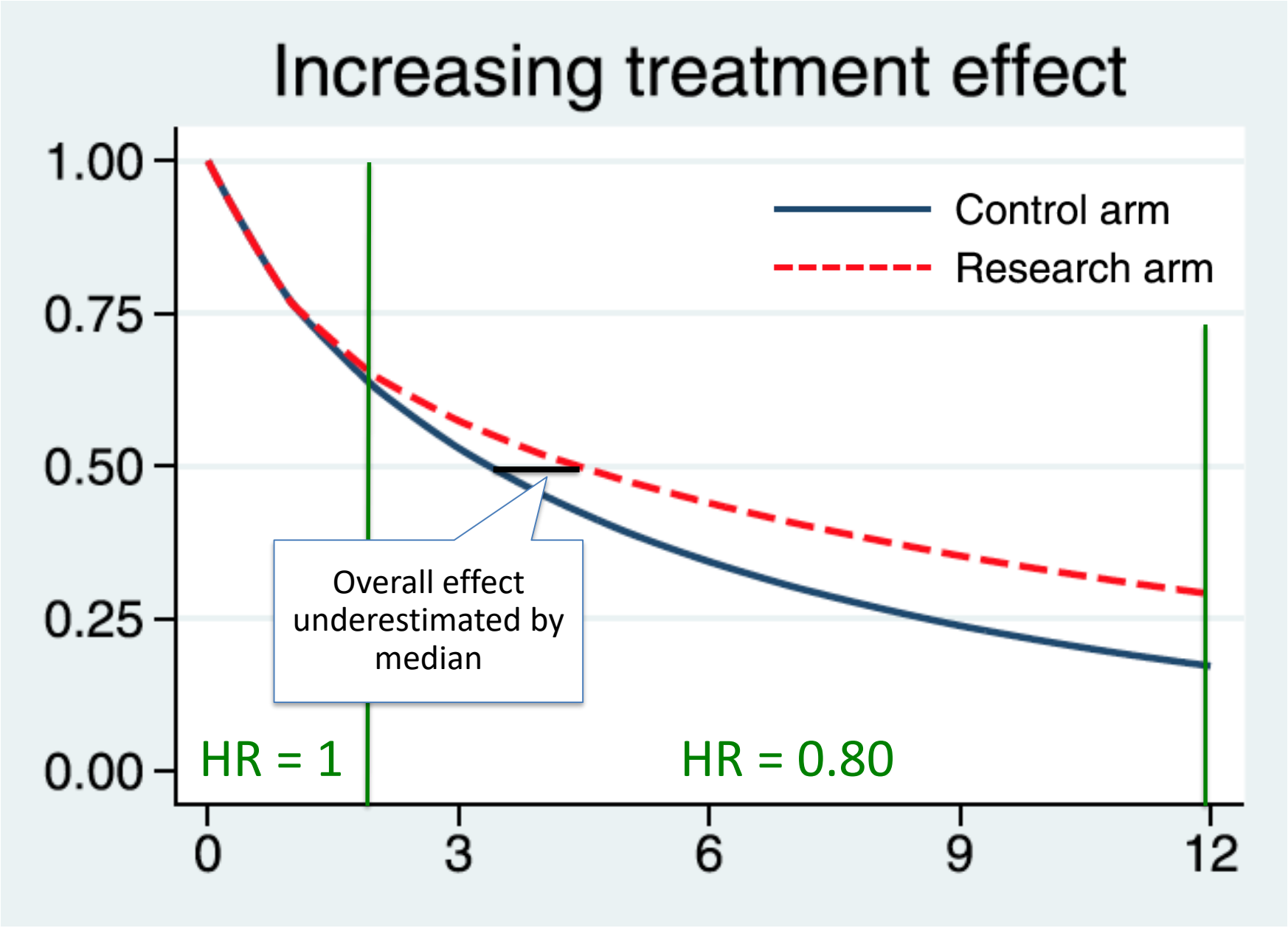
Here, the effect is the same in both time periods



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

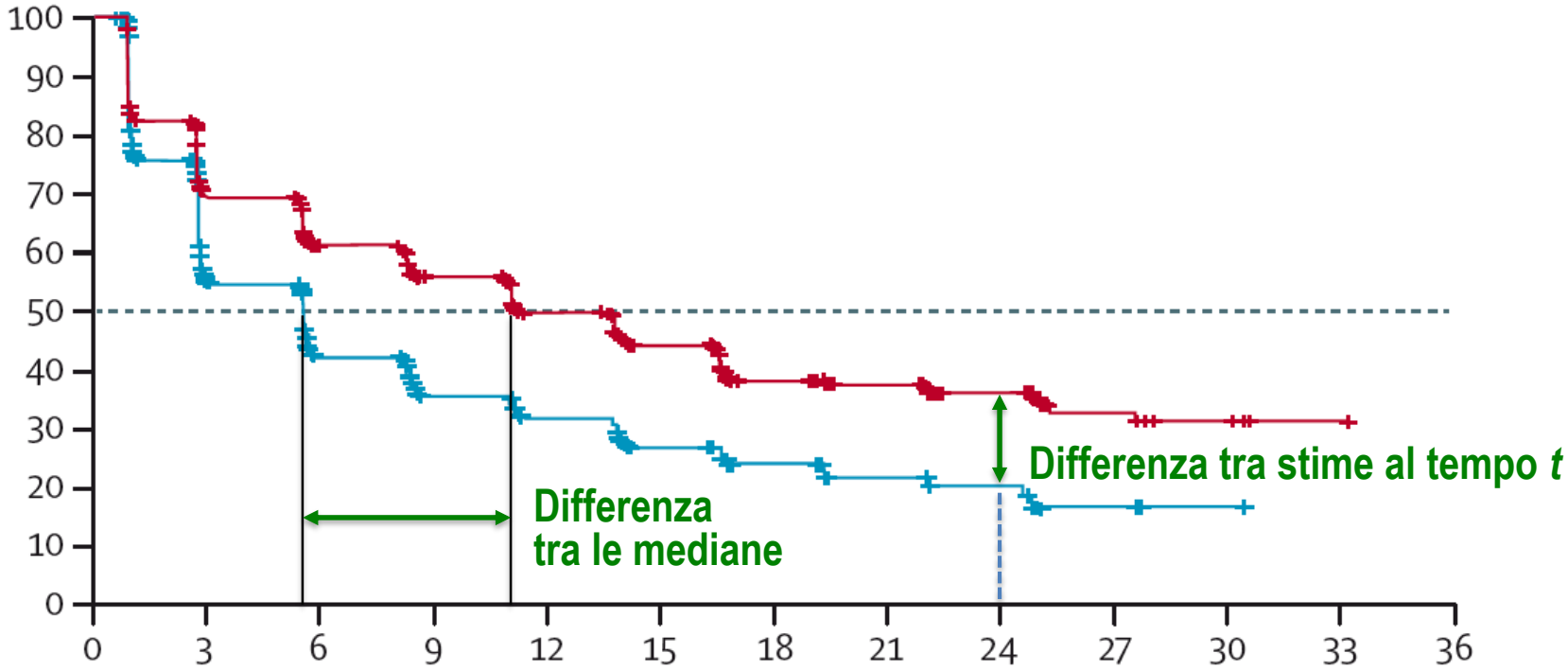

Proportional hazards
1.00
0.75
0.50
0.25
0.00
Control arm
Research arm
HR = 0.80
HR = 0.80
HR = 0.80
0
3
6
9
12

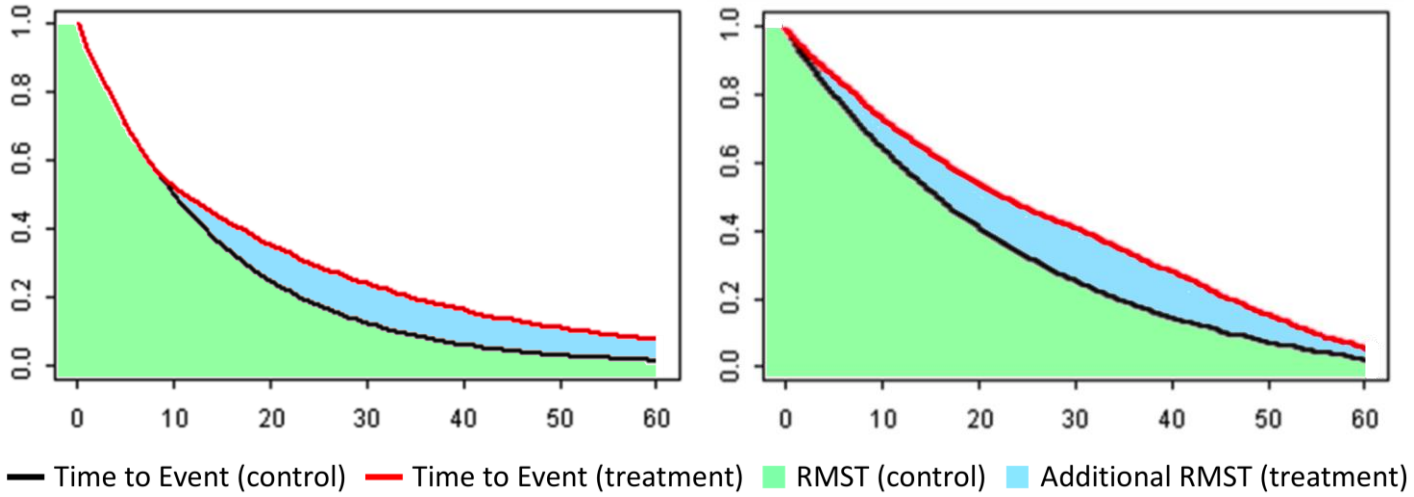

Decreasing treatment effect
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Control arm
Research arm
Overall effect overestimated by median
HR = 0.80
HR = 1.1
HR = 1.5


Increasing treatment effect
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Control arm
Research arm
Overall effect underestimated by median
HR = 1
HR = 0.80

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

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



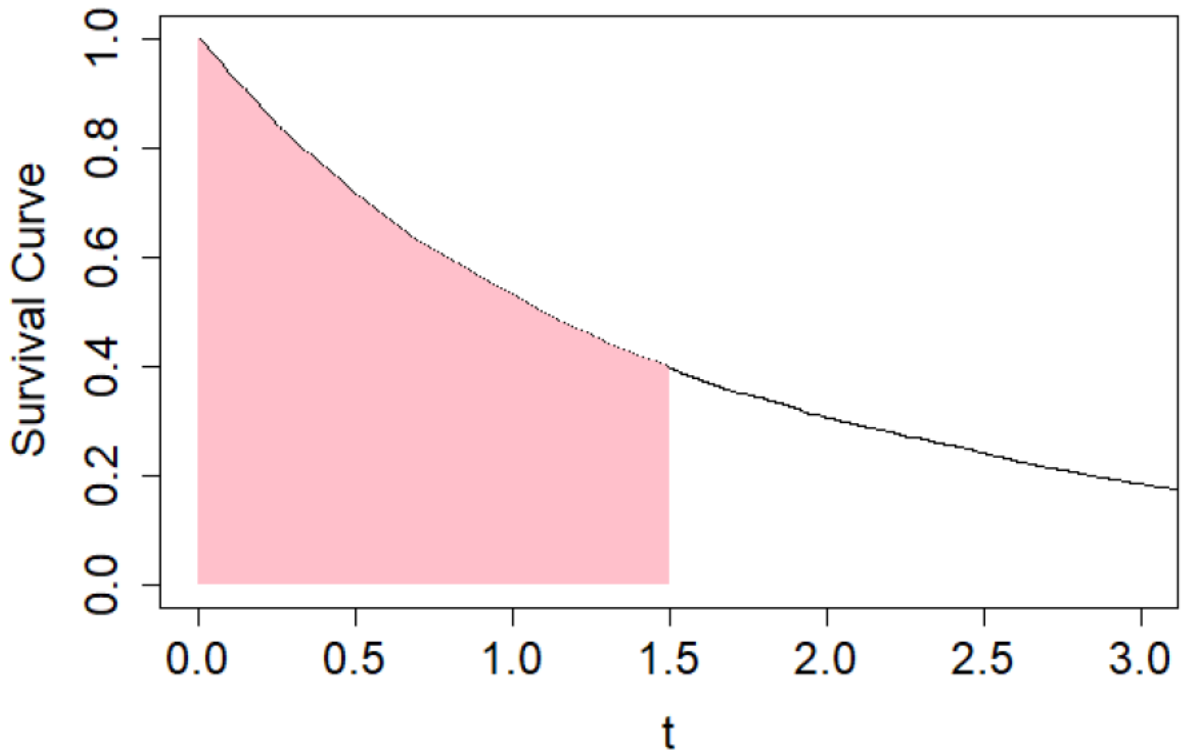


Restricted Mean Survival Time (differenza tra le aree sotto la curva)

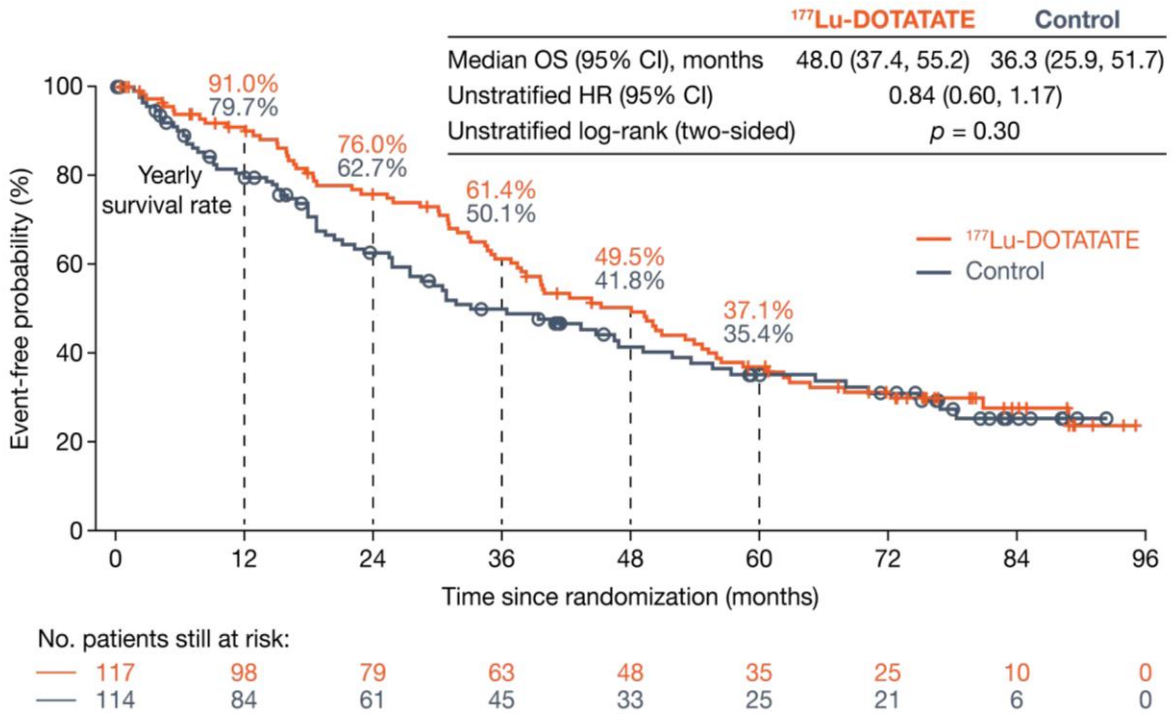
Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome

Patrick Royston* and Mahesh KB Parmar

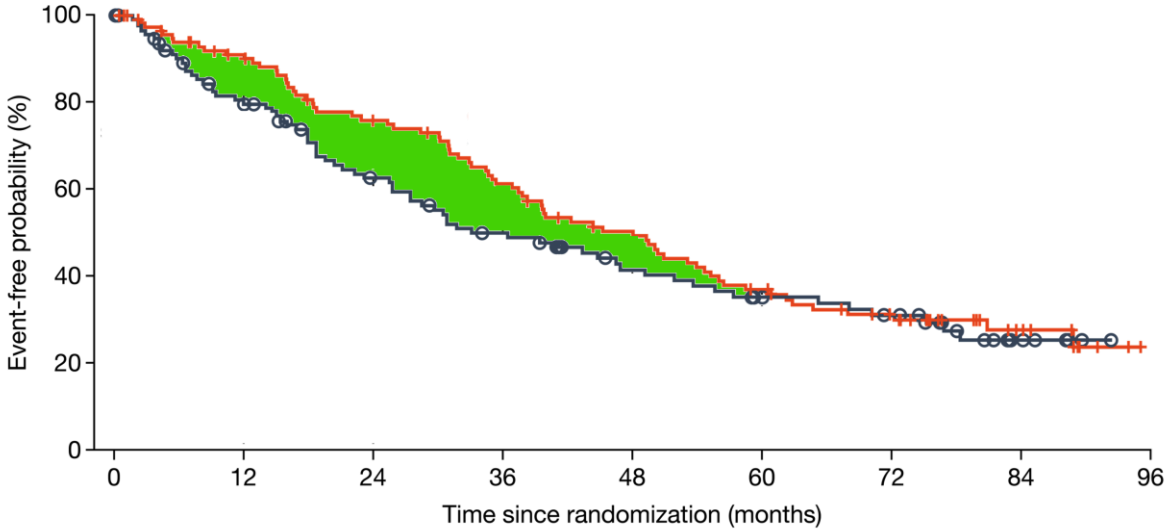
BMC Medical Research Methodology 2013, **13**:152



The phase 3 NETTER-1 study of 177Lu-DOTATATE in patients with midgut neuroendocrine tumours: further survival analyses



Deaths, n (%)	65 (55.6)	63 (55.3)
RMST, months (95% CI)	41.2 (37.6, 44.9)	36.1 (31.9, 40.4)
Difference, months (95% CI)	**5.1 (–0.5, 10.7)**	



Ruszniewski PB, et al. ESMO 2021

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

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

Journal of Clinical Epidemiology 118 (2020) 124–131

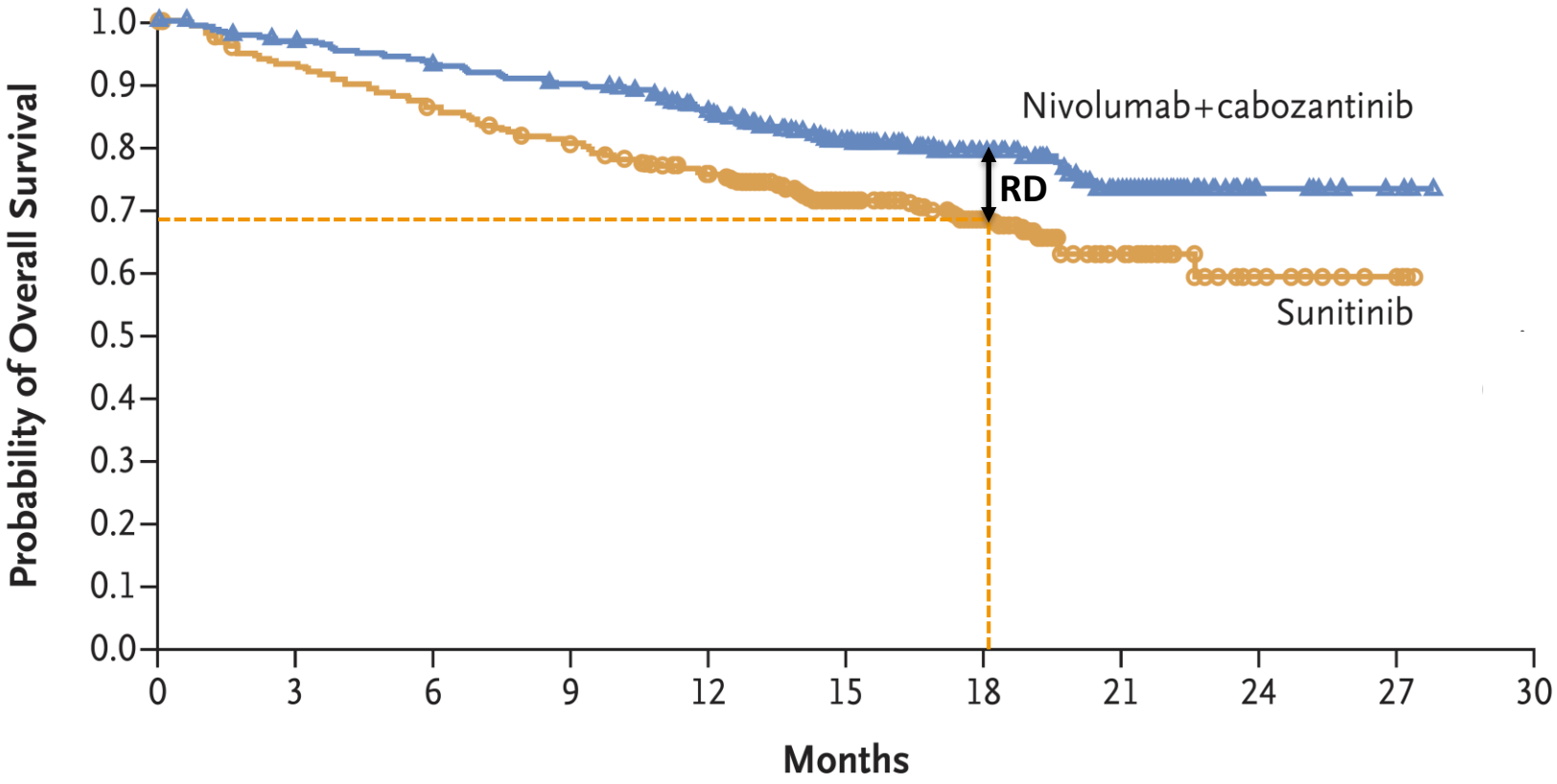
GRADE guidelines 27: how to calculate absolute effects for time-to-event outcomes in summary of findings tables and Evidence Profiles

Nicole Skoetz[a,*], Marius Goldkuhle[a], Elvira C. van Dalen[b], Elie A. Akl[c], Marialena Trivella[d], Reem A. Mustafa[e], Artur Nowak[f], Philipp Dahm[g], Holger Schünemann [h], Ralf Bender[i], GRADE Working Group

Absolute effect estimates (i.e., risk difference, the number needed to treat) provide important supplementary information to relative effect estimates by considering the control event rate over a given time period. As they take into account the underlying baseline risk for the event of interest in the study groups, absolute effect estimates are less vulnerable to exaggerated effect interpretation than relative effect estimates and allow a more appropriate assessment of the clinical relevance of effects.

…

Data from Kaplan-Meier survival curves from the control groups of the trials included in the analysis may be used to estimate the baseline risk.



Hazard ratio for death, 0.60
(98.89% CI, 0.40–0.89)

- Median f.u.: 18.1 months
  Baseline risk* at median f.u.: 31%
- **Risk Difference**: **11 events lower / 100 pts** (95%CI: 15 lower to 5 lower)

N Engl J Med 2021;384:829-41.

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

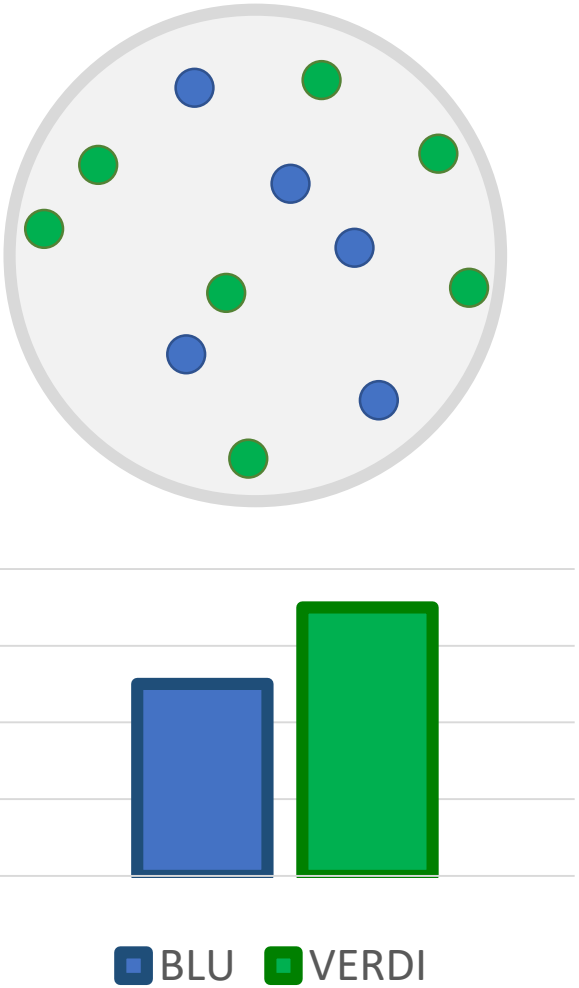
- Etica della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- **Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)**
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

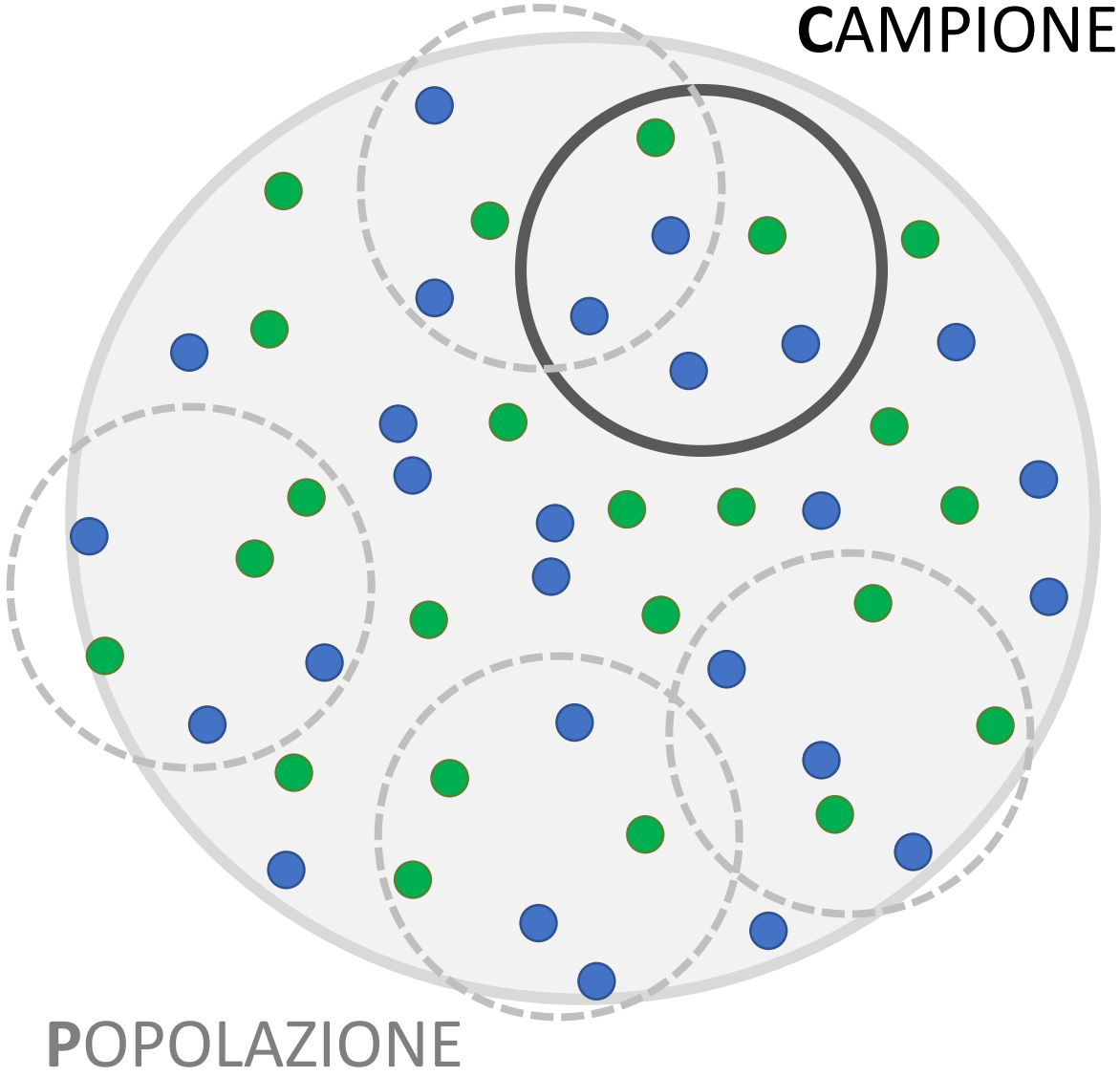
PRINCIPI DI STATISTICA MEDICA:
ERRORI STATISTICI,
VERIFICA DI IPOTESI,
DIMENSIONAMENTO CAMPIONARIO

“we can be blind to the obvious, and we are also blind to our blindness.”

— **Daniel Kahneman**

STATISTICA DESCRITTIVA VS. INFERENZIALE





STIMA PUNTUALE VS. STIMA INTERVALLARE

Stima puntuale
Singolo valore che mira al parametro della popolazione.

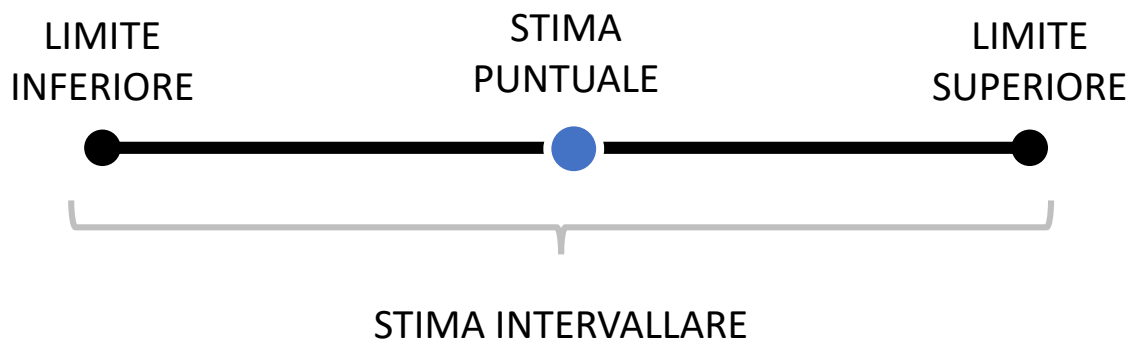
Stima intervallare
Specifica un intervallo di valori che contiene il parametro della popolazione. Solitamente viene costruito tenendo conto di una probabilità del 95%.

Se riportiamo un intervallo di valori plausibili, abbiamo buone probabilità di cogliere il parametro.

INTERVALLI DI CONFIDENZA

Qual è la cosa che ha il 95% di probabilità di accadere?
Circa **il 95% degli intervalli costruiti** seguendo la procedura di estrazione del campione e di stima conterrà il vero parametro di popolazione.

Il nostro intervallo costruito dal campione osservato conterrà o non conterrà il parametro, non lo sappiamo.



TEST D'IPOTESI

In un **tribunale**:
l'imputato è considerato innocente se non diversamente dimostrato.

L'obiettivo del pubblico ministero è di trovare delle **evidenze** che dimostrino la colpevolezza dell'imputato.

Ma quanto siamo disposti a considerare certe le evidenze che definiranno la colpevolezza dell'innocente?

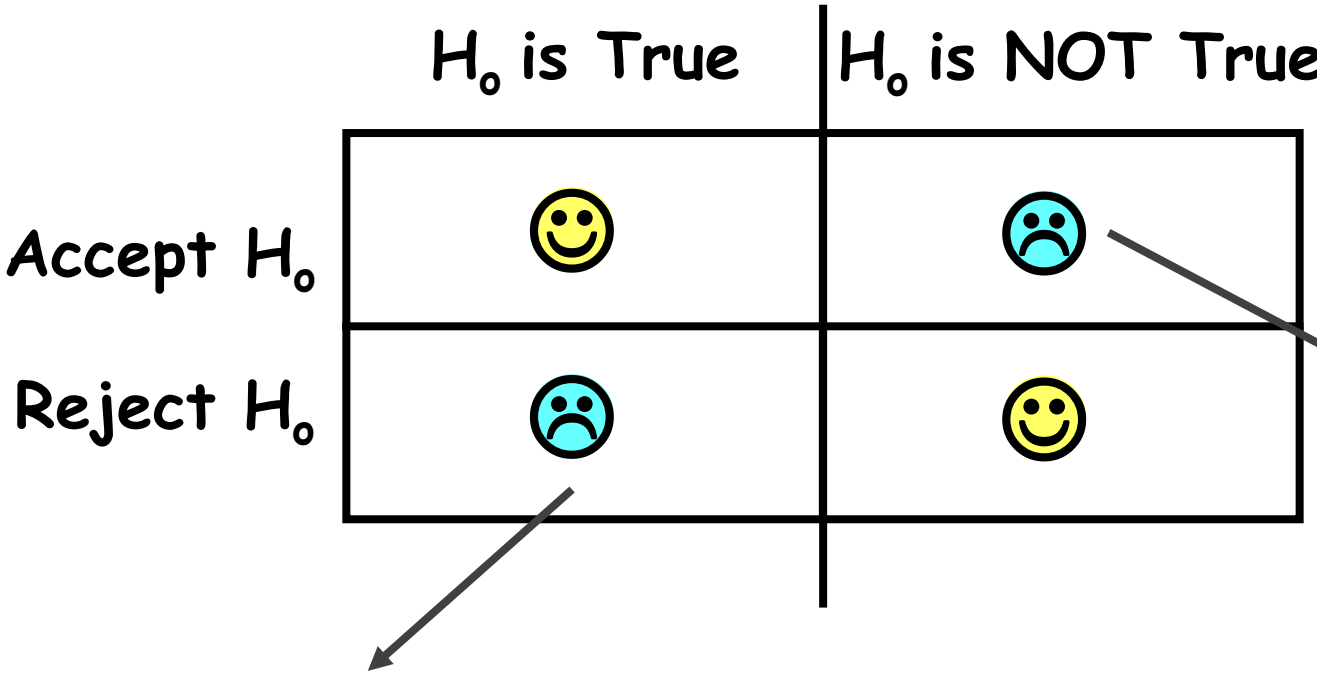
		Verità	
		INNOCENTE	COLPEVOLE
Verdetto	INNOCENTE	Corretto	Errore
	COLPEVOLE	Errore	Corretto

TEST D'IPOTESI

In tribunale	In uno studio clinico
Innocente	Non esiste differenza tra i trattamenti (H0)
Colpevole	Esiste una differenza tra i trattamenti (H1)

G.L. Pappagallo - S. Hilsenbeck, SABCS 2007 (modif. CM)

TEST D'IPOTESI



Errore di II tipo (β)
Accetto un'ipotesi che non è vera

Errore di I tipo (α)
Rifiuto un'ipotesi vera

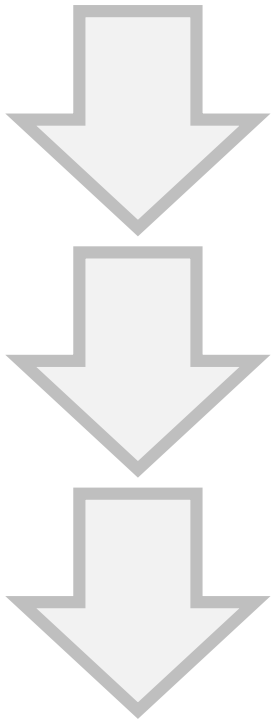
G.L. Pappagallo - S. Hilsenbeck, SABCS 2007 (modif. CM)

TEST D'IPOTESI

In tribunale	In uno studio clinico
Innocente	Non esiste differenza tra i trattamenti (H0)
Colpevole	Esiste una differenza tra i trattamenti (H1)
Assolvo l'innocente	Corretta accettazione di H0
Condanno il colpevole	Corretto rifiuto di H0
Condanno l'innocente	Errore di I tipo
Assolvo il colpevole	Errore di II tipo

G.L. Pappagallo - S. Hilsenbeck, SABCS 2007 (modif. CM)

VERIFICA D'IPOTESI



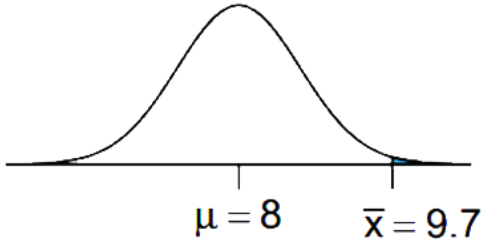
1. Impostare le ipotesi:
   - $H0 : \mu0 - \mu1 = 0$ (Non esiste differenza tra i trattamenti)
   - $H1 : \mu0 - \mu1 \neq 0$ (Esiste differenza tra i trattamenti)
2. Calcolare il valore del test e il corrispettivo **pvalue** (p)
3. Prendere una decisione:
   - Se il valore di $p < \alpha$, rifiutare H0
   - Se il valore di $p > \alpha$, non rifiutare H0

P-VALUE

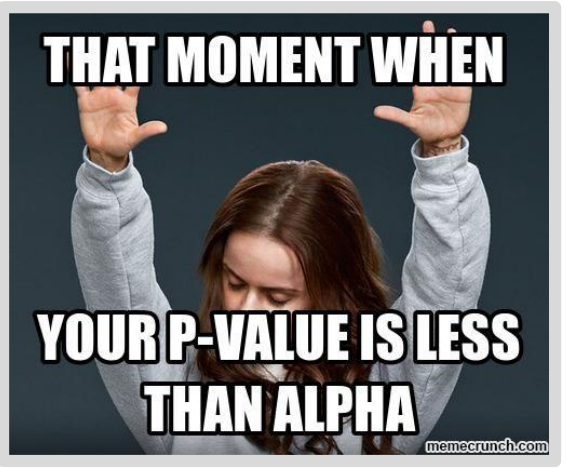
Il p-value è la probabilità che il parametro che voglio stimare, sotto l'ipotesi H0, abbia un valore uguale o più estremo di quello che ho osservato.

Ipotesi: H0: $\mu = 8$

Campione: $\bar{X} = 9.7$



$$p\text{-value} = P(\bar{x} > 9.7 \mid \mu = 8) = P(Z > 2.50) = 0.0062$$



P-VALUE

Non possiamo concludere da un P-value piccolo, se è presente un rapporto di causa effetto o se le nostre conclusioni possono essere generalizzate ad una popolazione più ampia.

Questo perché i p-value non possono dirci se il disegno di uno studio è stato impostato correttamente o se i dati sono stati raccolti in modo appropriato

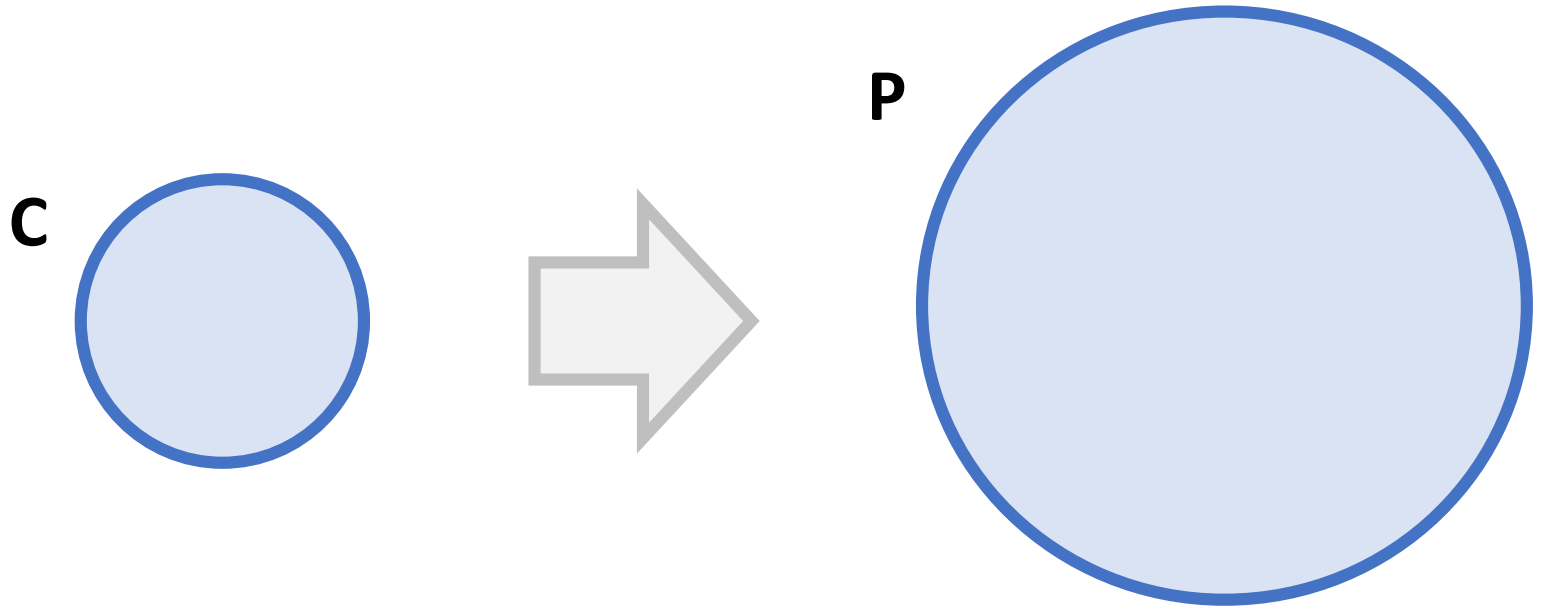
Garbage In → Garbage Out

Inoltre, il p-value non ci da alcuna informazione sulla rilevanza degli effetti che abbiamo osservato. **Rilevanze statistica non implica rilevanza clinica**.

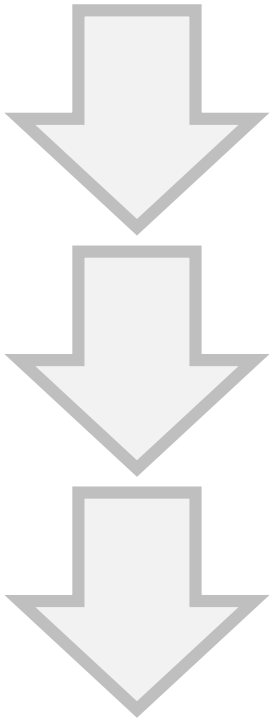
P-VALUE

Il campione deve essere **rappresentativo** per permettere l'inferenza

e **casuale** per essere rappresentativo.



DALL'IPOTESI ALLA NUMEROSITÀ

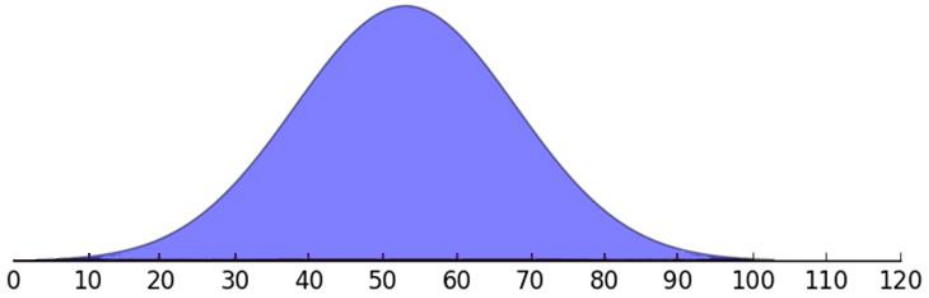


- **Obiettivo primario.**
- Definizione del **disegno di studio.**
- Scelta dei **criteri di inclusione** dei soggetti e metodi di selezione del campione.
- Definizione dell'**endpoint** corrispondente.
- Calcolo della **numerosità campionaria.**

ENDPOINT

Gli endpoints o **outcome** sono **misure di efficacia** scelte per valutare il raggiungimento degli obiettivi.

Devono essere validi, appropriati e clinicamente rilevanti.



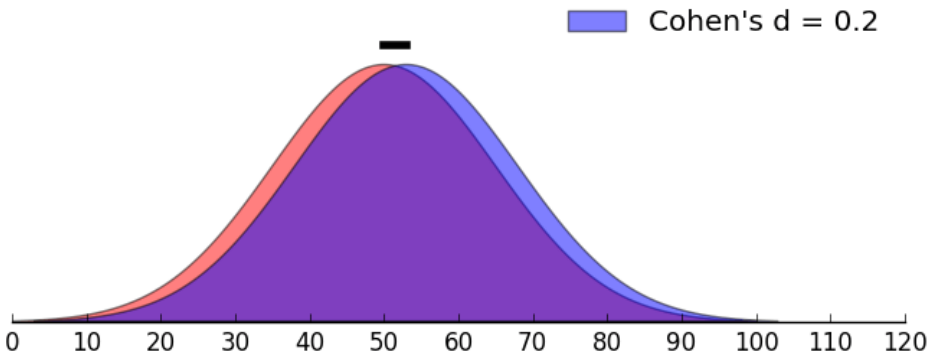
EFFECT SIZE

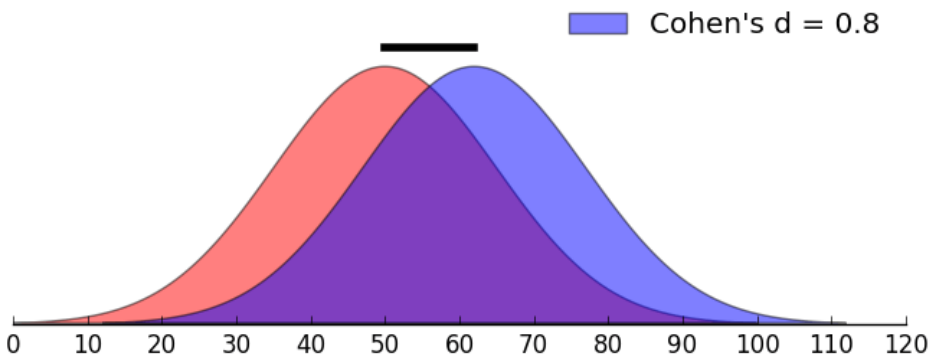
L'effect size è la misura quantitativa della forza di un fenomeno.

Esprime la **grandezza** (size) della differenza o relazione indagata.

Esempio:

$$\boldsymbol{d}_{Cohen} = \frac{differenza\ media}{standard\ deviation}$$





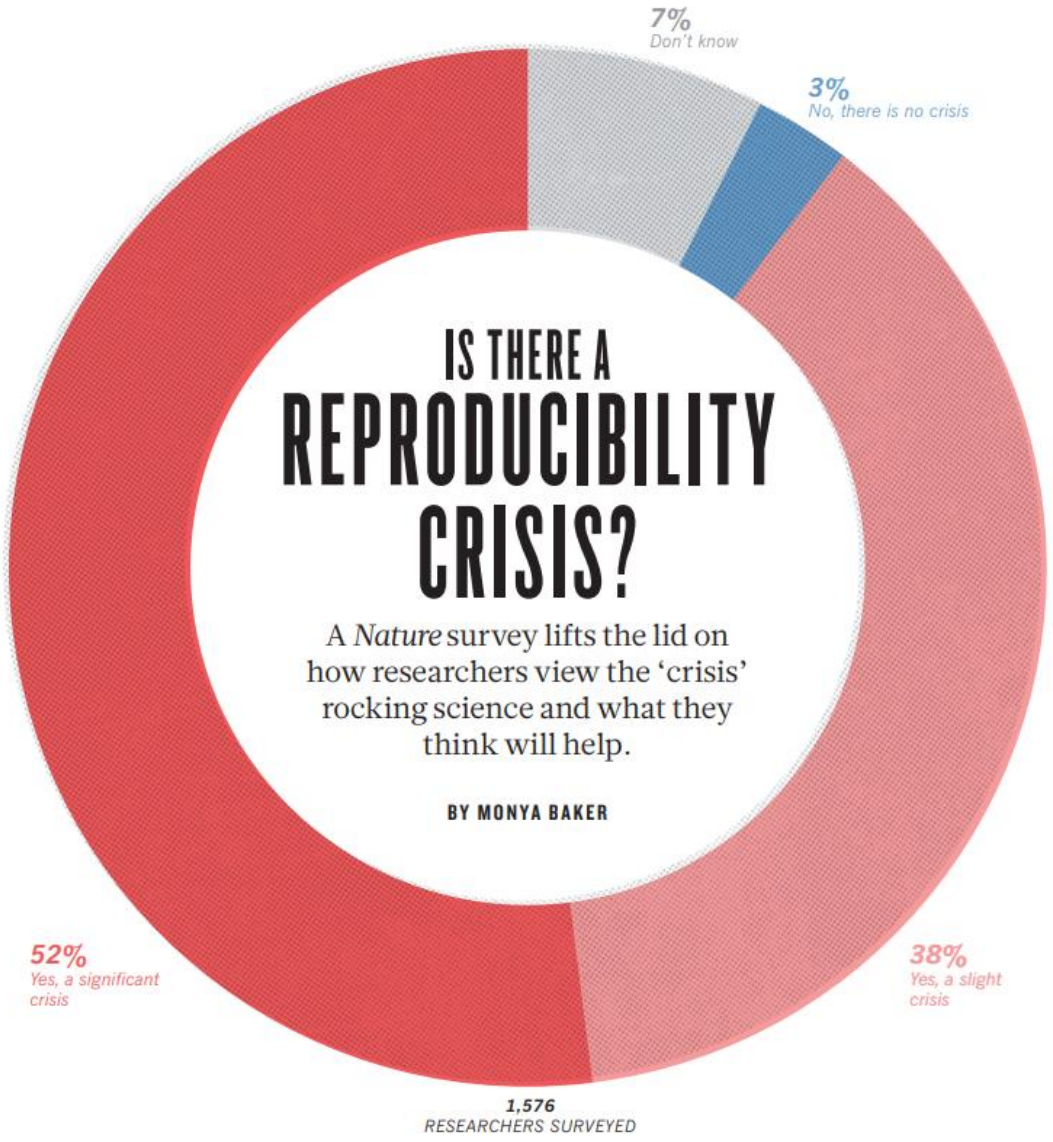
SAMPLE SIZE

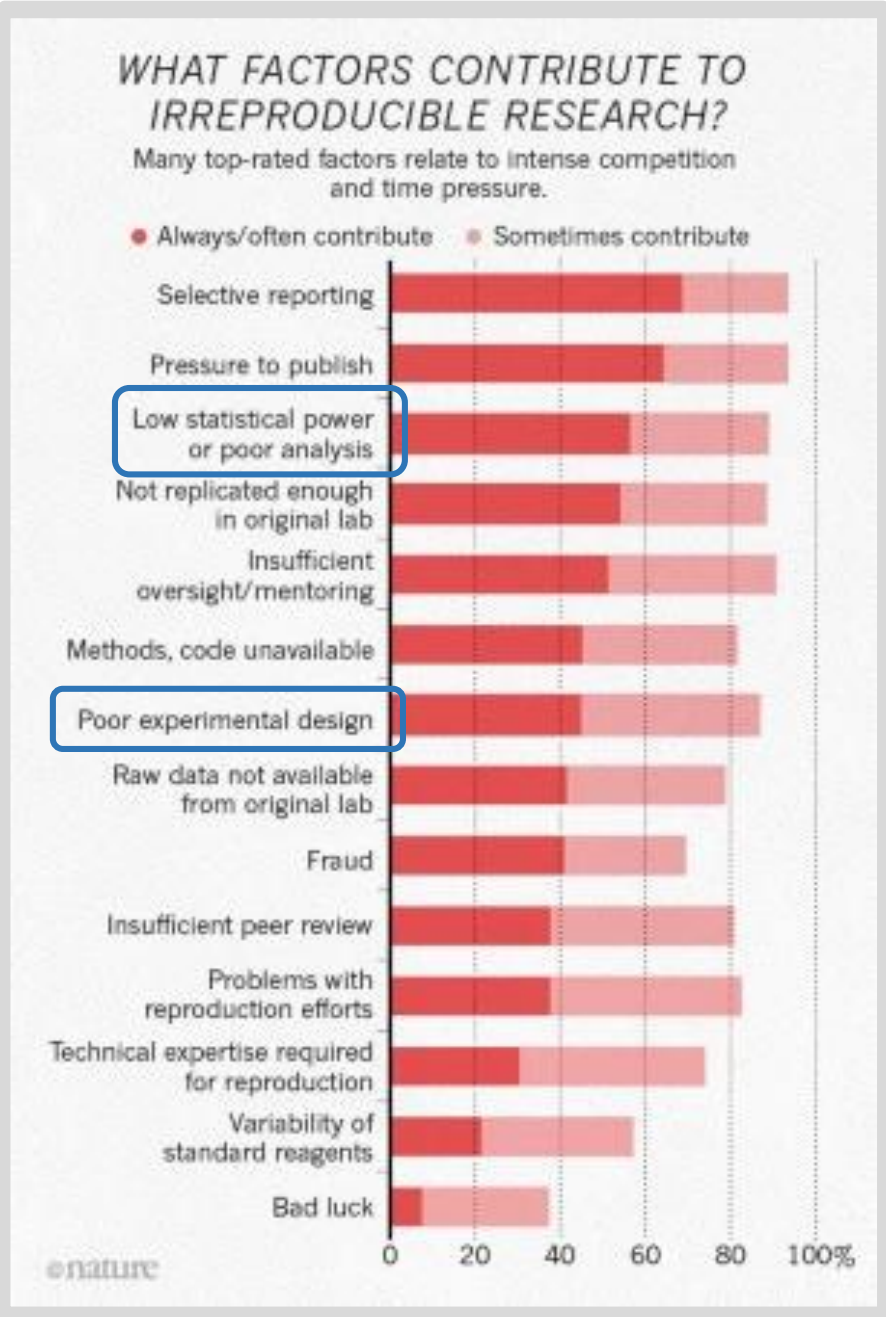
Un trial clinico è un esperimento che coinvolge persone. L'arruolamento di un numero adeguato di soggetti ha anche una rilevanza **etica**. Il numero di soggetti dovrebbe essere sufficiente affinché si riesca a rispondere adeguatamente agli obiettivi prefissati.

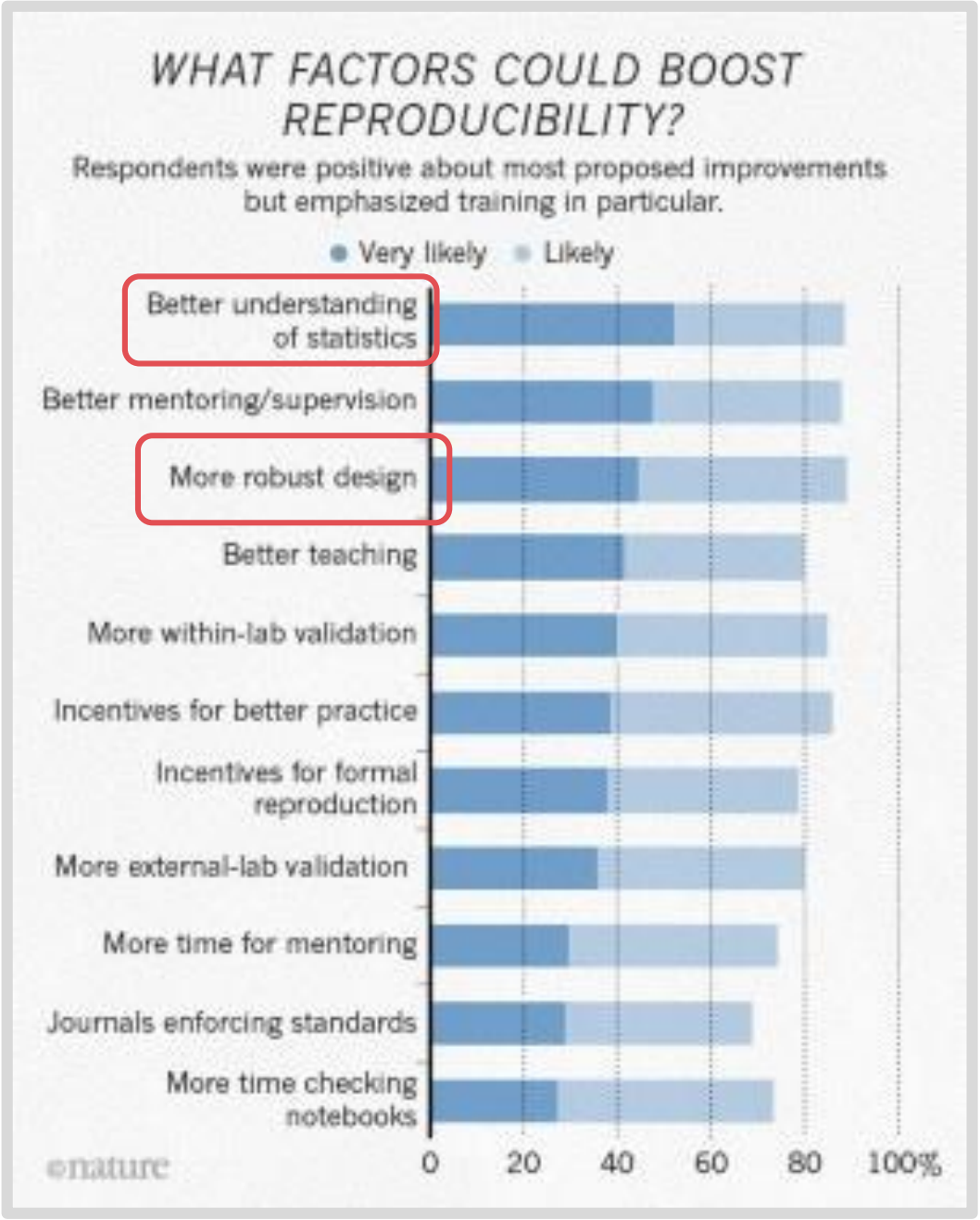
Sottodimensionare o sovradimensionare uno studio ha anche implicazioni **economiche**.

1,500 scientists lift the lid on reproducibility.

Baker, Nature, 2016




WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?
Many top-rated factors relate to intense competition and time pressure.
Always/often contribute
Sometimes contribute
Selective reporting
Pressure to publish
Low statistical power or poor analysis
Not replicated enough in original lab
Insufficient oversight/mentoring
Methods, code unavailable
Poor experimental design
Raw data not available from original lab
Fraud
Insufficient peer review
Problems with reproduction efforts
Technical expertise required for reproduction
Variability of standard reagents
Bad luck
0
20
40
60
80
100%
nature


WHAT FACTORS COULD BOOST REPRODUCIBILITY?
Respondents were positive about most proposed improvements but emphasized training in particular.
Very likely
Likely
Better understanding of statistics
Better mentoring/supervision
More robust design
Better teaching
More within-lab validation
Incentives for better practice
Incentives for formal reproduction
More external-lab validation
More time for mentoring
Journals enforcing standards
More time checking notebooks
0
20
40
60
80
100%
nature

- Etica e integrità della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- **Rilevanza clinica Vs significatività statistica**
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

Statistical Vs Clinical Significance

- **Statistical Significance**

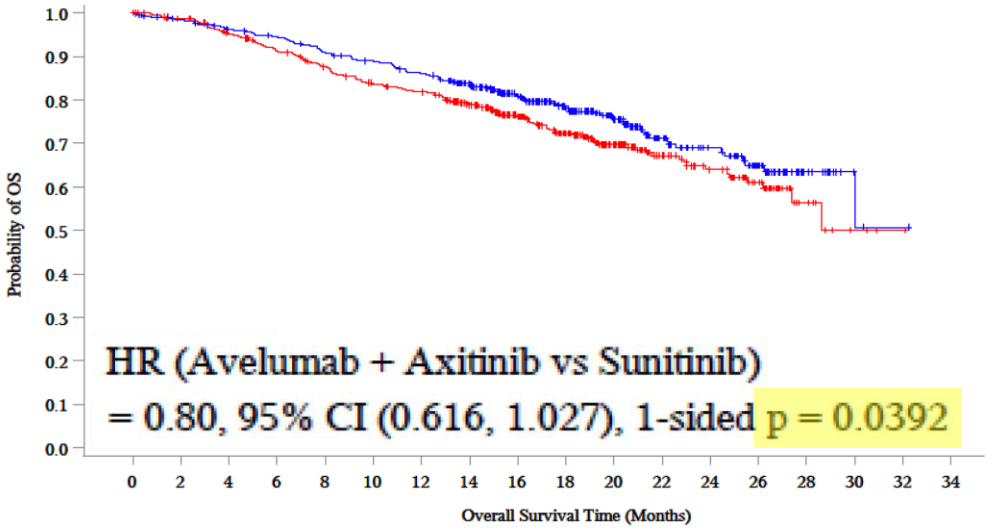
  "Is an observed difference **likely to be real**"

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

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

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

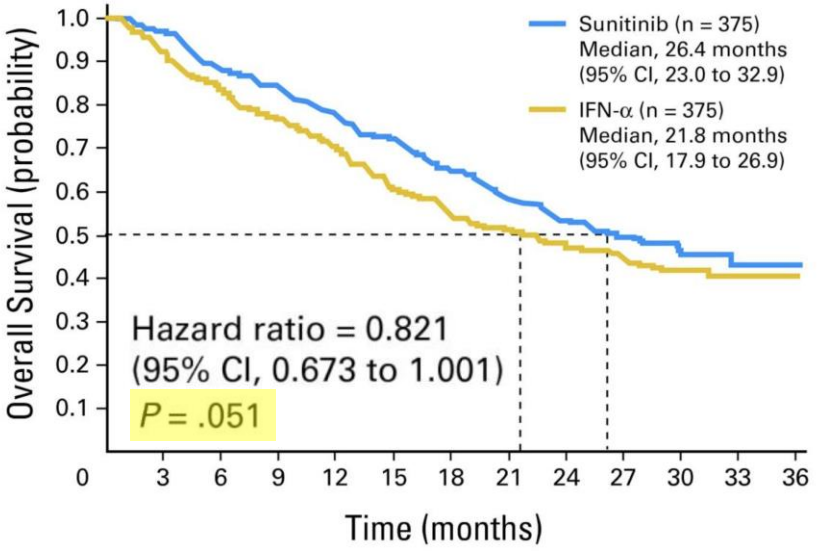
- Don't base your conclusions solely on whether an association or effect was found to be "statistically significant".



Motzer R, et al. N Engl J Med. 2019 Mar 21;380(12):1103-1115.

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

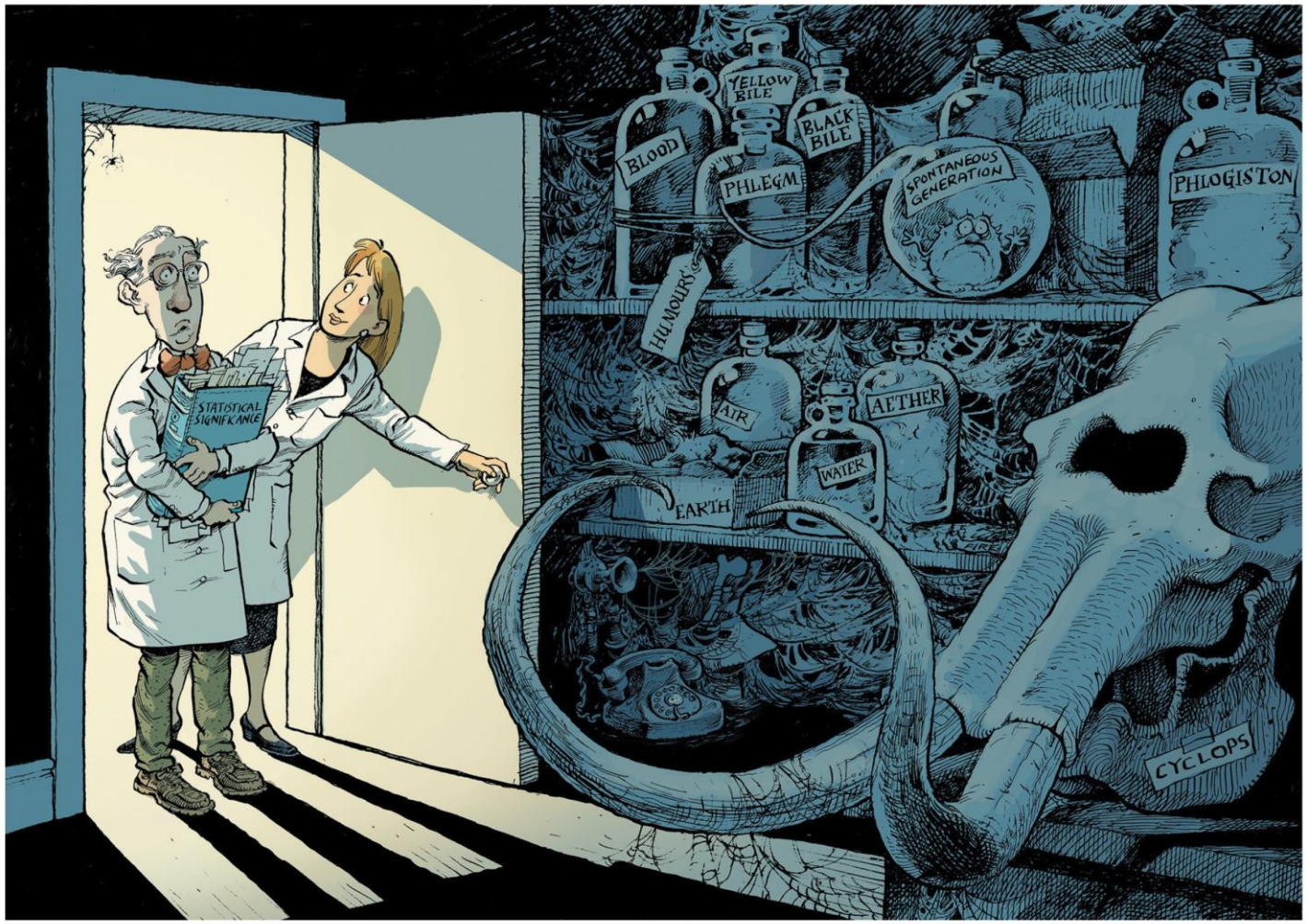
- Don't base your conclusions solely on whether an association or effect was found to be "statistically significant".
- Don't believe that an association or effect is absent just because it was not statistically significant.



Motzer R, et al. N Engl J Med. 2007 Jan 11;356(2):115

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

The fragility of phase 3 trials supporting FDA-approved anticancer medicines: a retrospective analysis

Joseph C Del Paggio, Ian F Tannock

***Lancet Oncol* 2019; 20: 1065–69**

Fragility Index (FI) = the minimum number of changes from non-events to events resulting in loss of statistical significance in a RCT

FI≤2 in 59% of trials

	Control group (n=362)	Experimental group (n=360)	
Number of events	157	129	p=0·040
Number of non-events	205	231	
Number of events	157	129+1=130	p=0·048
Number of non-events	205	231–1=230	
Number of events	157	129+2=131	p=0·058
Number of non-events	205	231–2=229	

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?

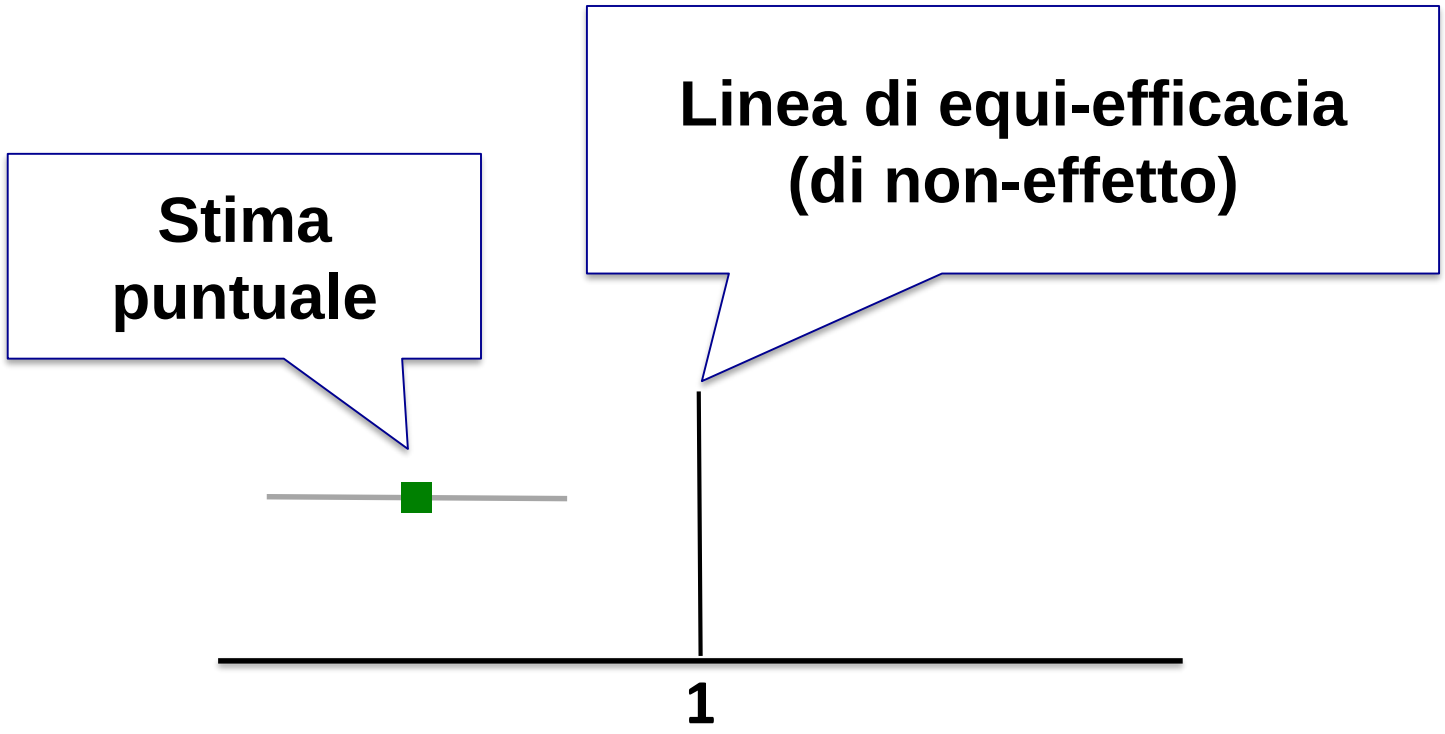
We would define a positive trial as one in which the predefined value of δ represents a clinically important difference in an end-point that directly reflects benefit (mainly OS or quality of life) to patients and for which the results provide a best estimate of the difference that exceeds that predefined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

Obiettivi di uno studio comparativo

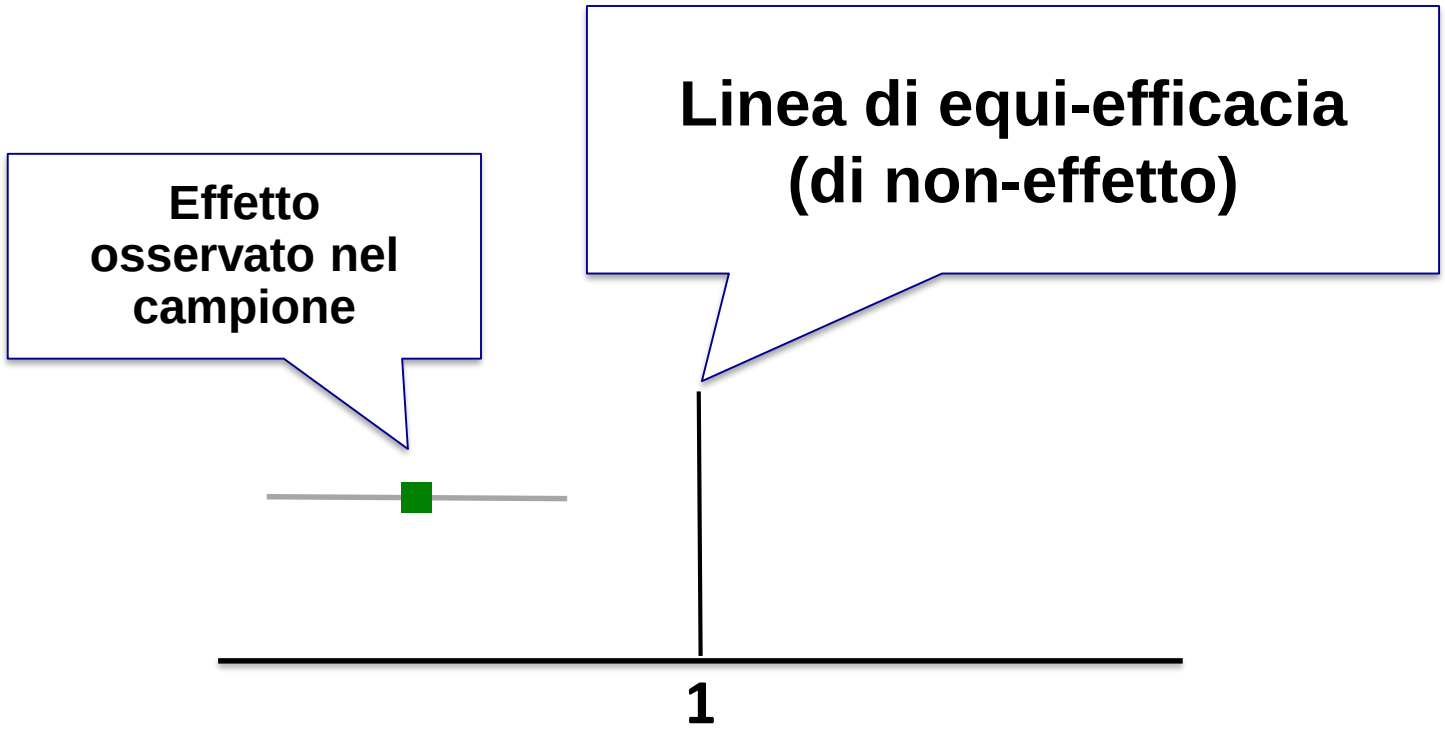
Si ritiene che il trattamento in esame "A" abbia le potenzialità per migliorare il trattamento standard "B" almeno di una quantità Δ

studio di superiorità	studio di non inferiorità
A > B di una quantità Δ di interesse clinico	A < B non oltre una quantità M di rilevanza clinica

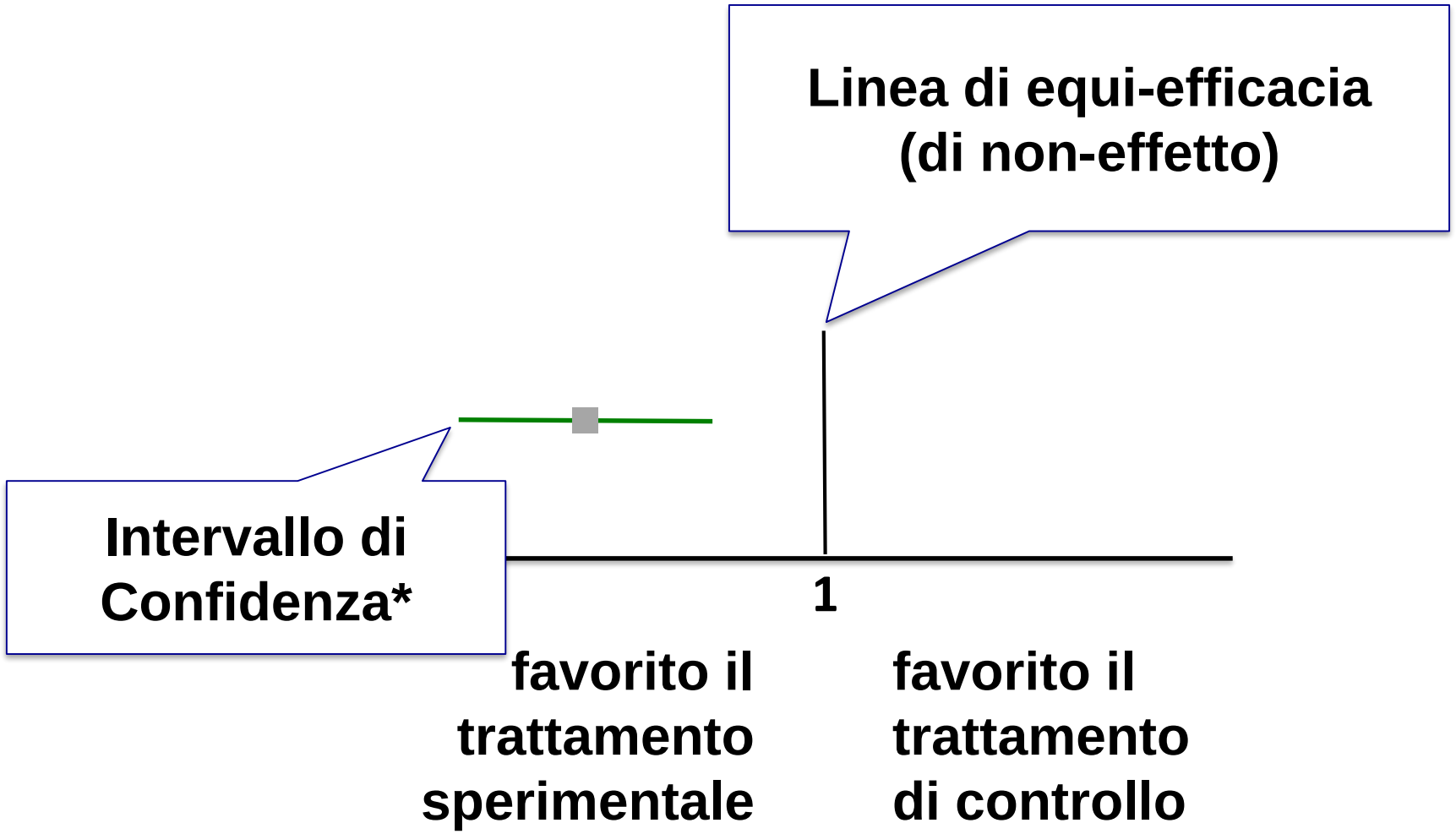
Interpretazione degli studi clinici mediante Forest Plot



Interpretazione degli studi clinici mediante Forest Plot

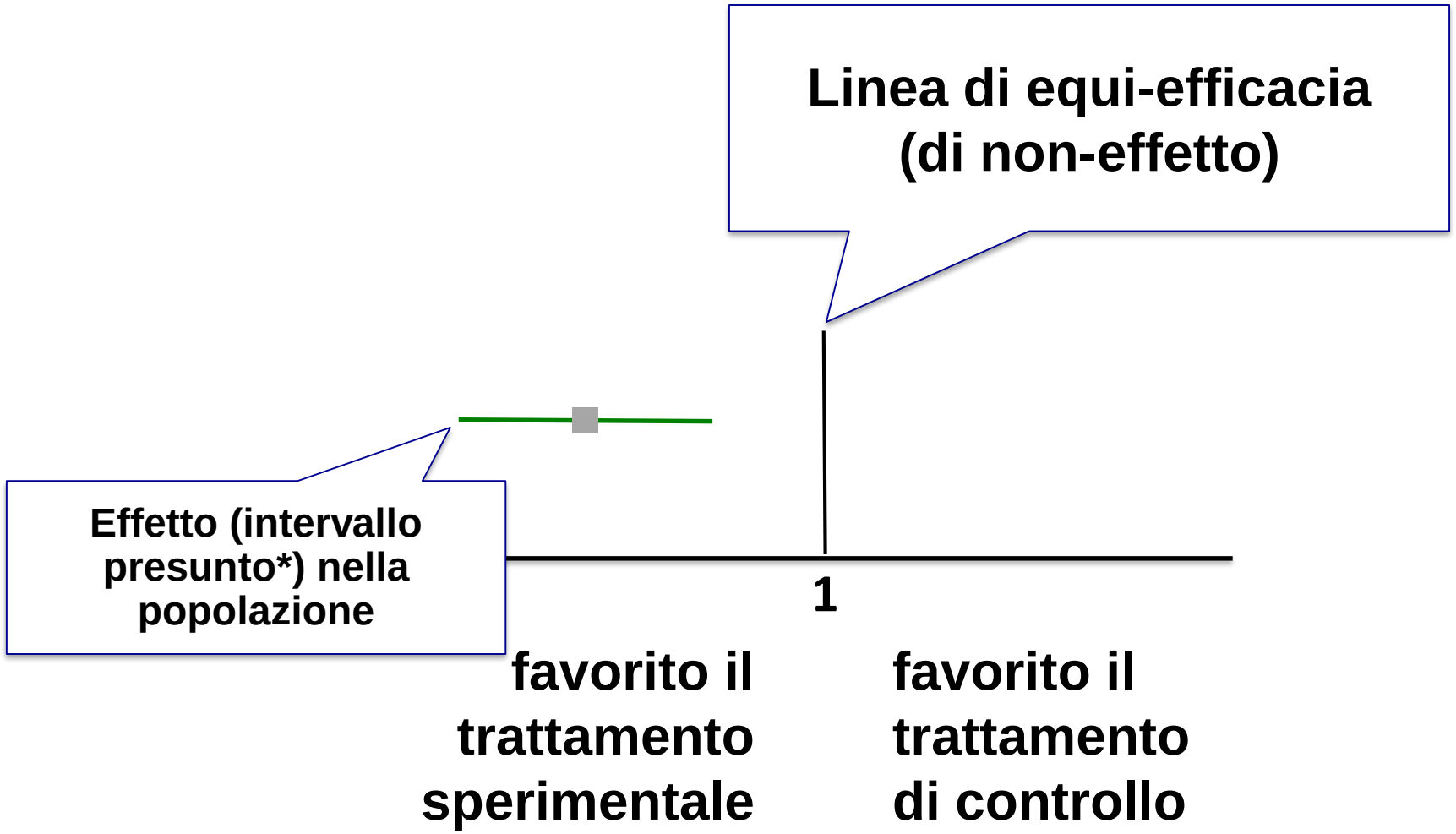


Interpretazione degli studi clinici mediante Forest Plot



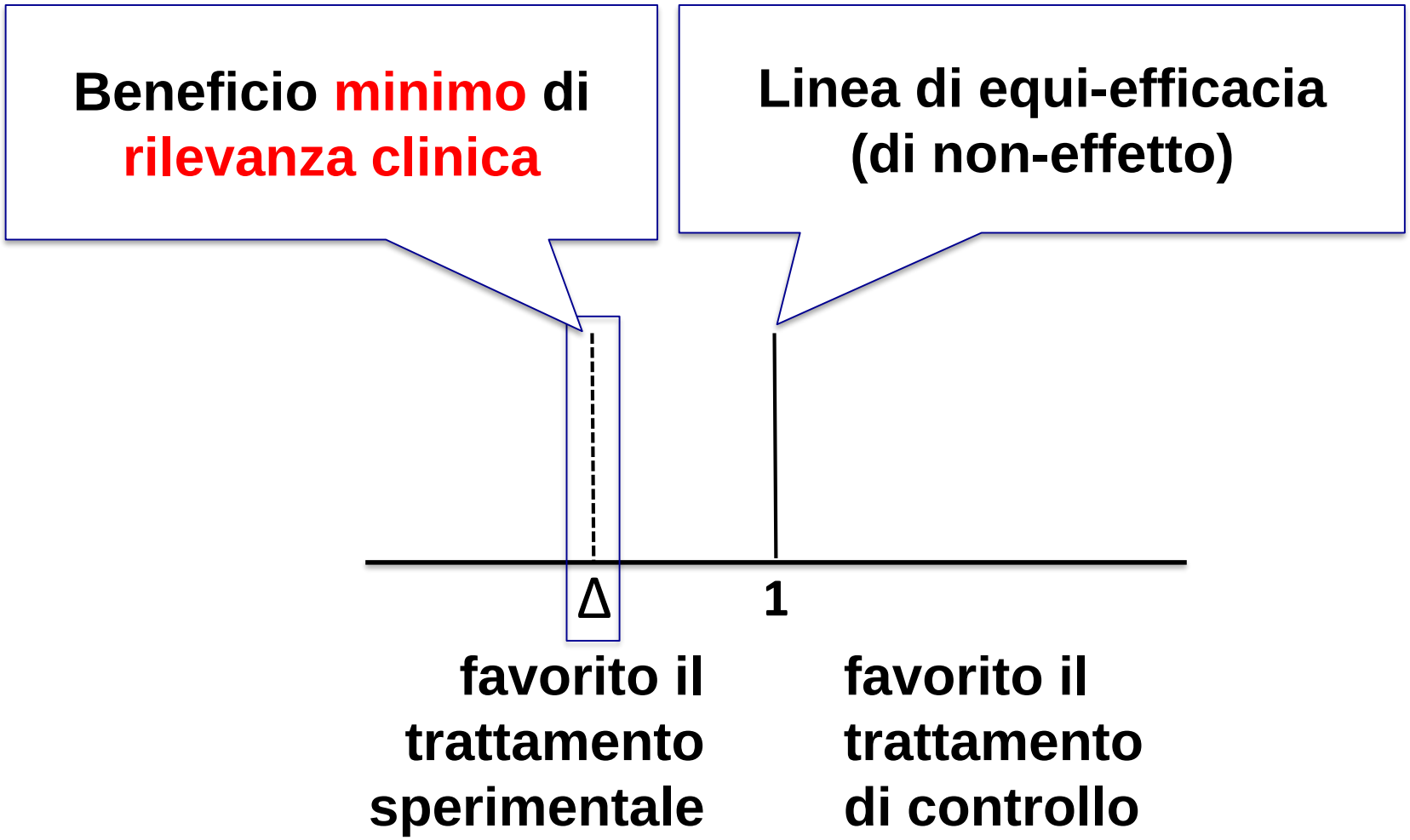
* convenzionalm. 95%

Interpretazione degli studi clinici mediante Forest Plot



* convenzionalm. 95%

Interpretazione degli studi clinici mediante Forest Plot

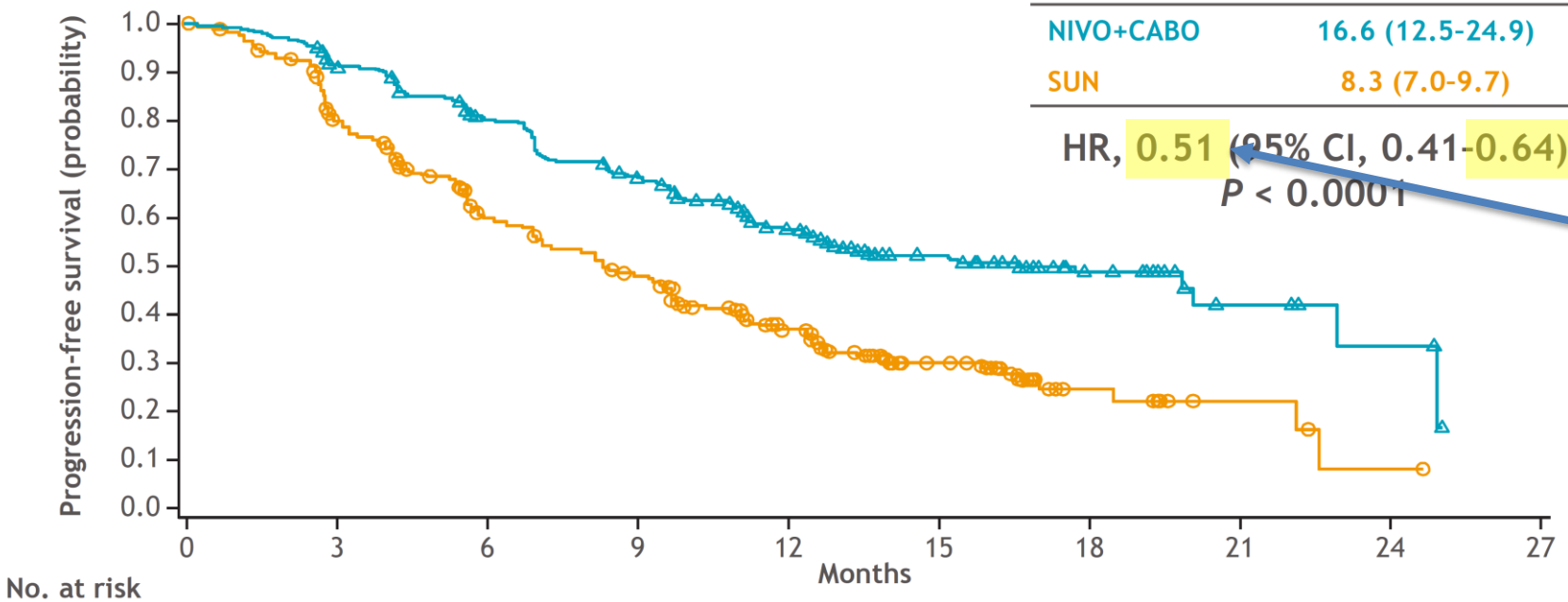


VIRTUAL 2020 ESMO congress

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



No. at risk

NIVO+CABO	323	279	234	196	144	77	35	11	4	0
SUN	328	228	159	122	79	31	10	4	1	0

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

Symptom Endpoints (Patient-Reported Outcomes)

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

Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma

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

N Engl J Med 2013;369:722-31

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

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

Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma

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

N Engl J Med 2013;369:722-31

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

Instrument	Pazopanib	Sunitinib	Difference in Mean Change from Baseline Score with Pazopanib vs. Sunitinib†	P Value§	Drug Favored According to Significant Difference¶	Effect Size‖
	number of patients					
FACIT-F**	377	403	2.32	<0.001	Pazopanib	0.24
FKSI-19**						
Treatment side effects	3			0.03	Pazopanib	0.14
Disease-related physical symptoms	3			0.03	Pazopanib	0.13
Disease-related emotional symptoms	3			0.41	Neither	–0.04
Functional well-being	378	403	0.31	0.10	Neither	0.09
Total score	377	408	1.41	0.02	Pazopanib	0.14

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

Minimal (Clinical) Interesting Difference (MID / MCID)

the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management

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

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

Kimberly Webster, David Cella* and Kathleen Yost

Health and Quality of Life Outcomes 2003, 1:79

Table 1: Minimally important differences for select FACIT scales

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

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

Parliamo di Maturità di uno Studio Clinico

$$\frac{\text{numero di eventi osservati}}{\text{numero di pazienti arruolati}} \quad Vs \quad \frac{\text{numero di eventi osservati}}{\text{numero degli eventi previsti}}$$

For trials with a well-designed interim-monitoring plan, stopping at 50% or greater information has a negligible impact on estimation.

Freidlin & Korn. Clinical Trials 2009; 6: 119–125

Nivolumab plus Cabozantinib versus Sunitinib for Advanced Renal-Cell Carcinoma

T.K. Choueiri, T. Powles, M. Burotto, B. Escudier, M.T. Bourlon, B. Zurawski, V.M. Oyervides Juárez, J.J. Hsieh, U. Basso, A.Y. Shah, C. Suárez, A. Hamzaj, J.C. Goh, C. Barrios, M. Richardet, C. Porta, R. Kowalyszyn, J.P. Feregrino, J. Żołnierek, D. Pook, E.R. Kessler, Y. Tomita, R. Mizuno, J. Bedke, J. Zhang, M.A. Maurer, B. Simsek, F. Ejzykowicz, G.M. Schwab, A.B. Apolo, and R.J. Motzer, for the CheckMate 9ER Investigators*

N Engl J Med 2021;384:829-41.

	Nivo+Cabo N = 323	Sunitinib N = 328
PFS per BICR		
Events, n (%)	144 (44.6)	191 (58.2)

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

335 events observed / 350 planned: I.F. 95.7%

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

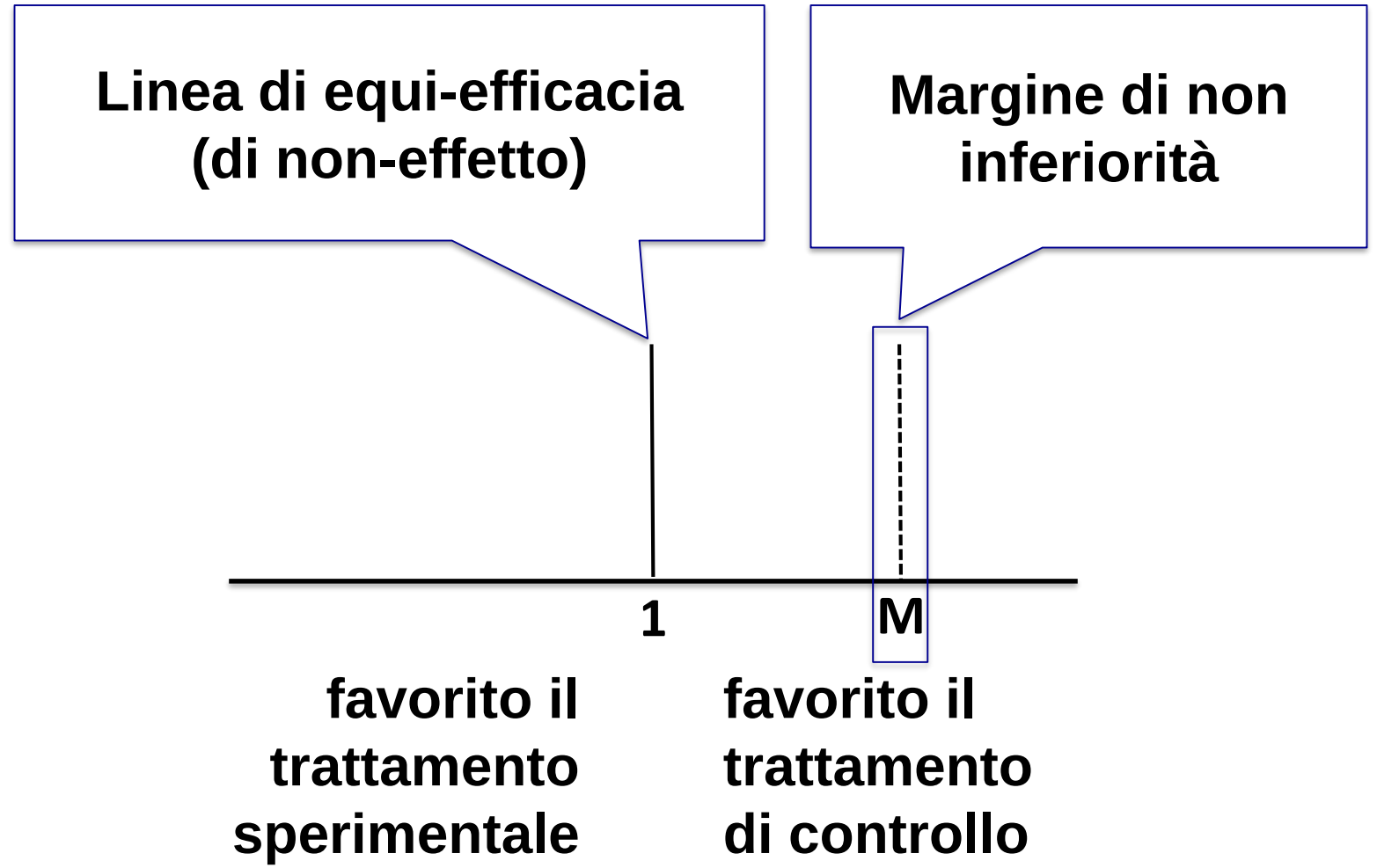
studio di superiorità

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 Plot



Through the looking glass: understanding non-inferiority

Jennifer Schumi* and Janet T Wittes

Trials 2011, **12**:106

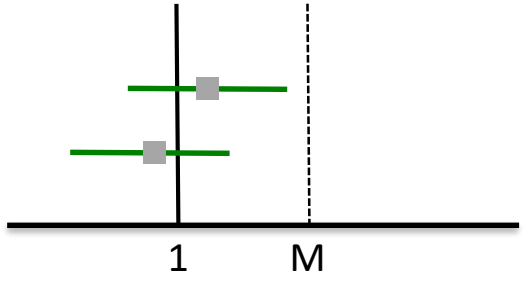
One focuses on the upper bound for this non-inferiority comparison; what happens at the lower end of the CI is not the primary concern.

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

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

Interpretazione clinica di uno Studio di Non-Inferiorità

(dato uno specifico M di interesse)



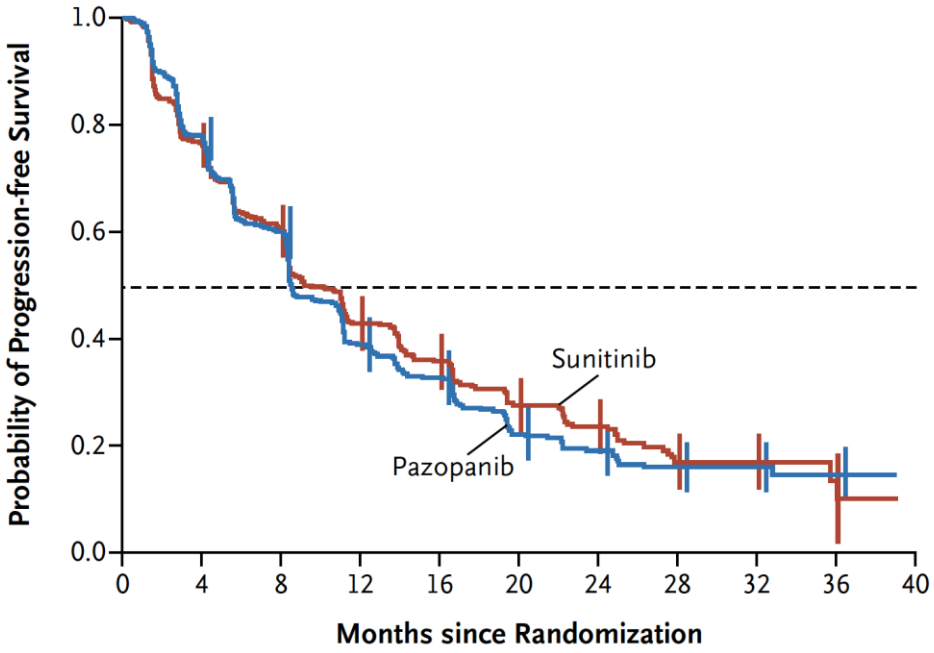
Dimostrazione di Non-Inferiorità

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

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)

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

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

IV.2.3 Choice of analysis set

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

Guidance for Industry
Non-Inferiority Clinical Trials

March 2010
Clinical/Medical

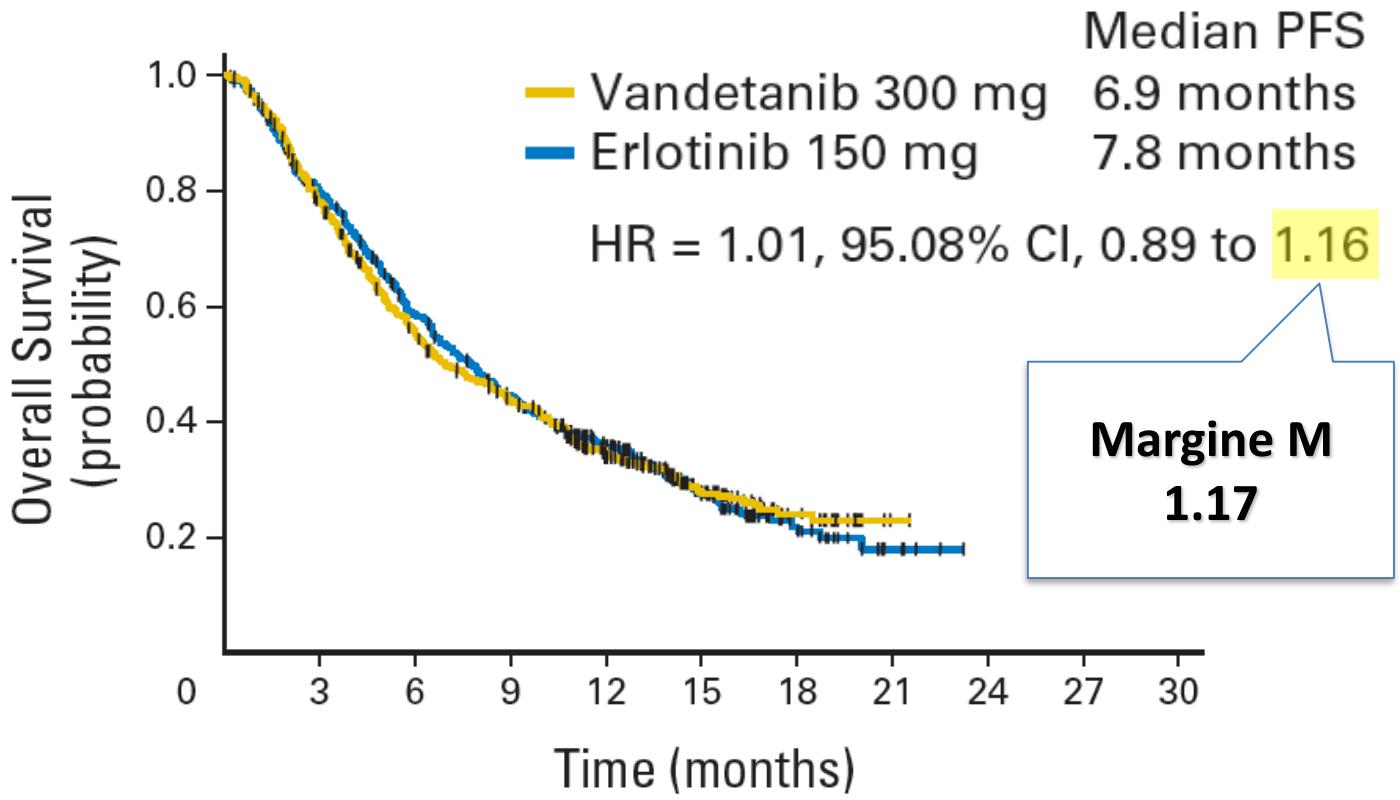
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

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

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

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

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



Margine M 1.17

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

SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

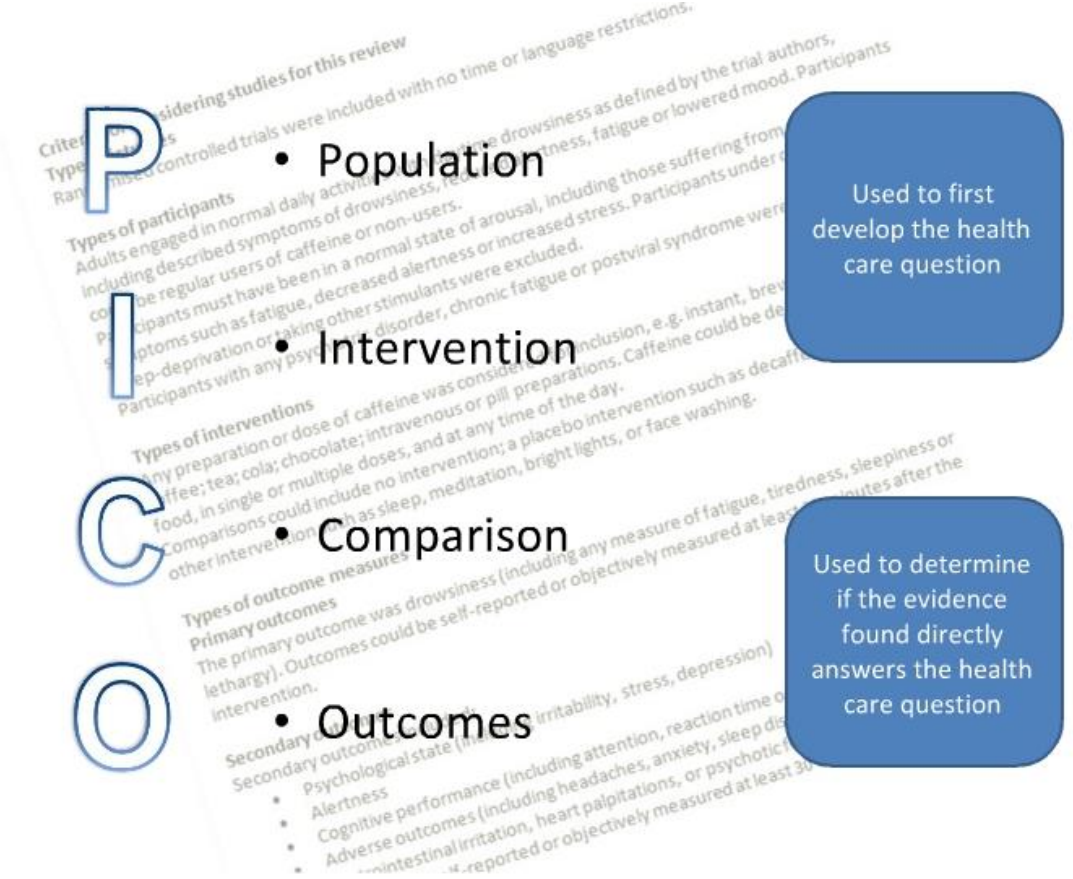
NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

- Etica e integrità della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- **Trasferibilità delle evidenze al quesito clinico**
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

PREMESSA

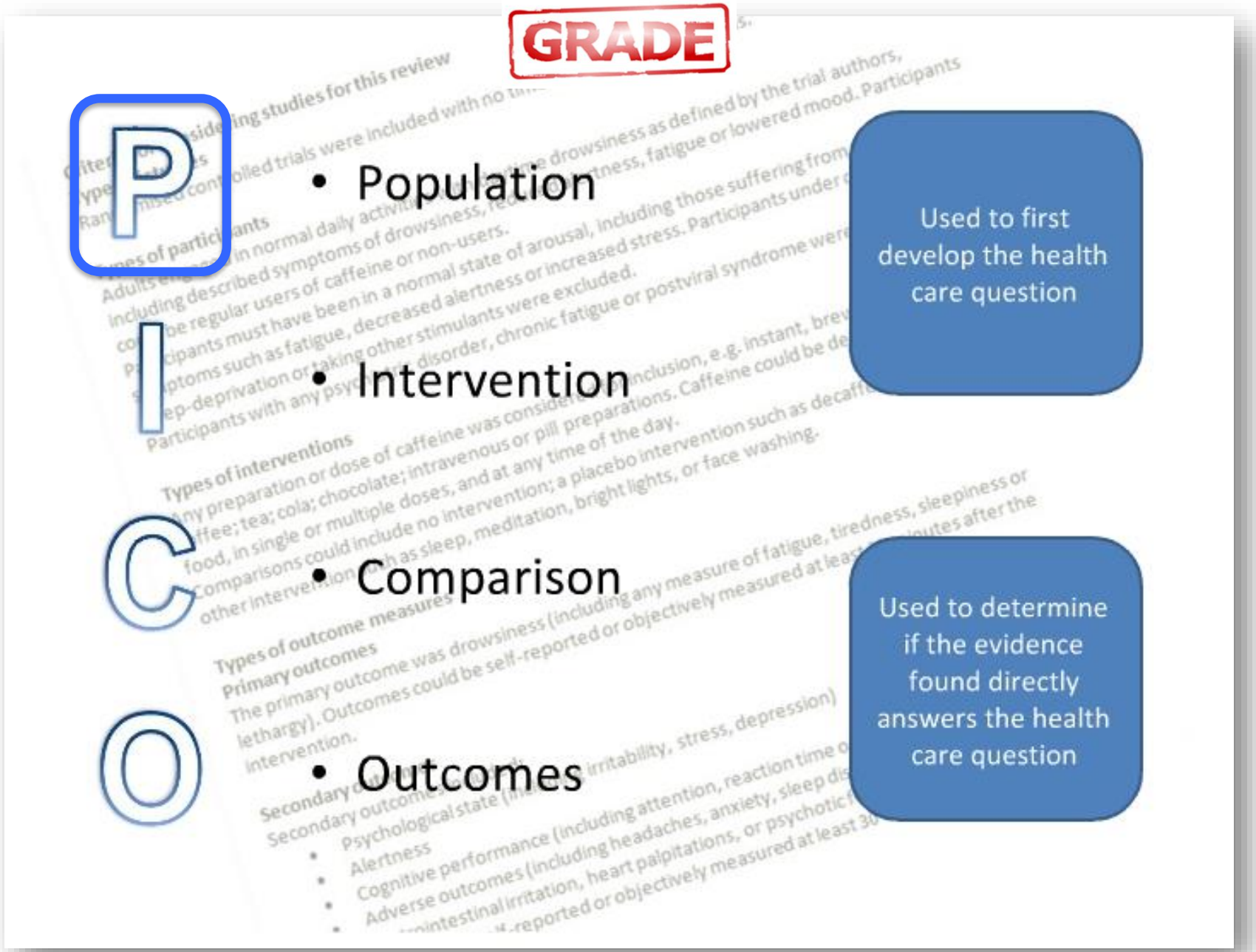
Un trial clinico non dovrebbe essere *letto così com'è*, ma avendo come riferimento uno specifico quesito di particolare interesse.



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

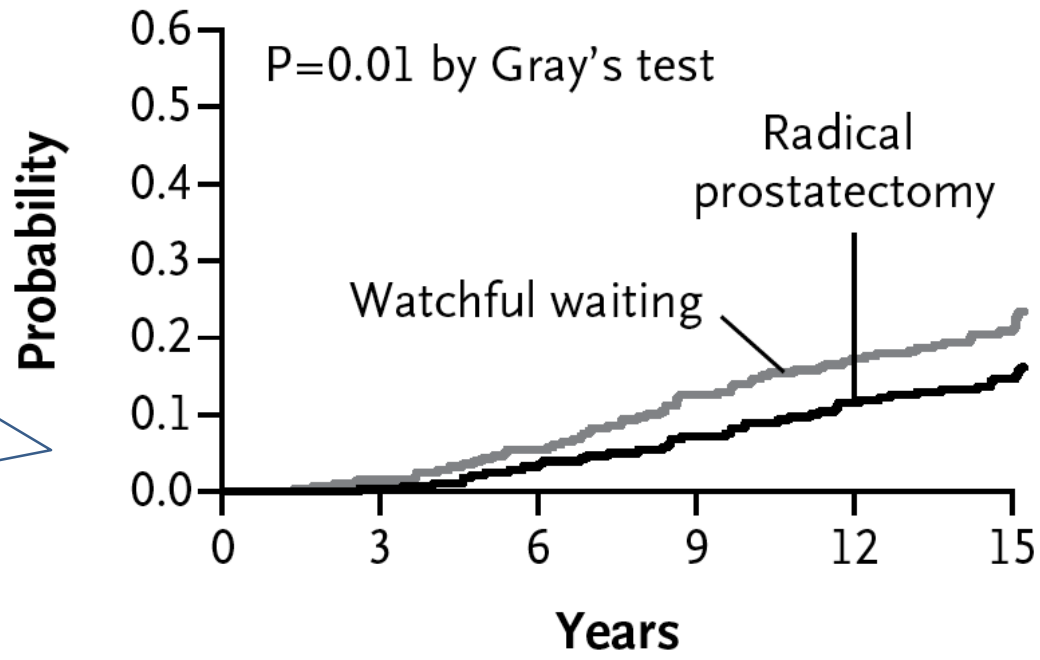

GRADE
P
• Population
I
• Intervention
C
• Comparison
O
• Outcomes
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question

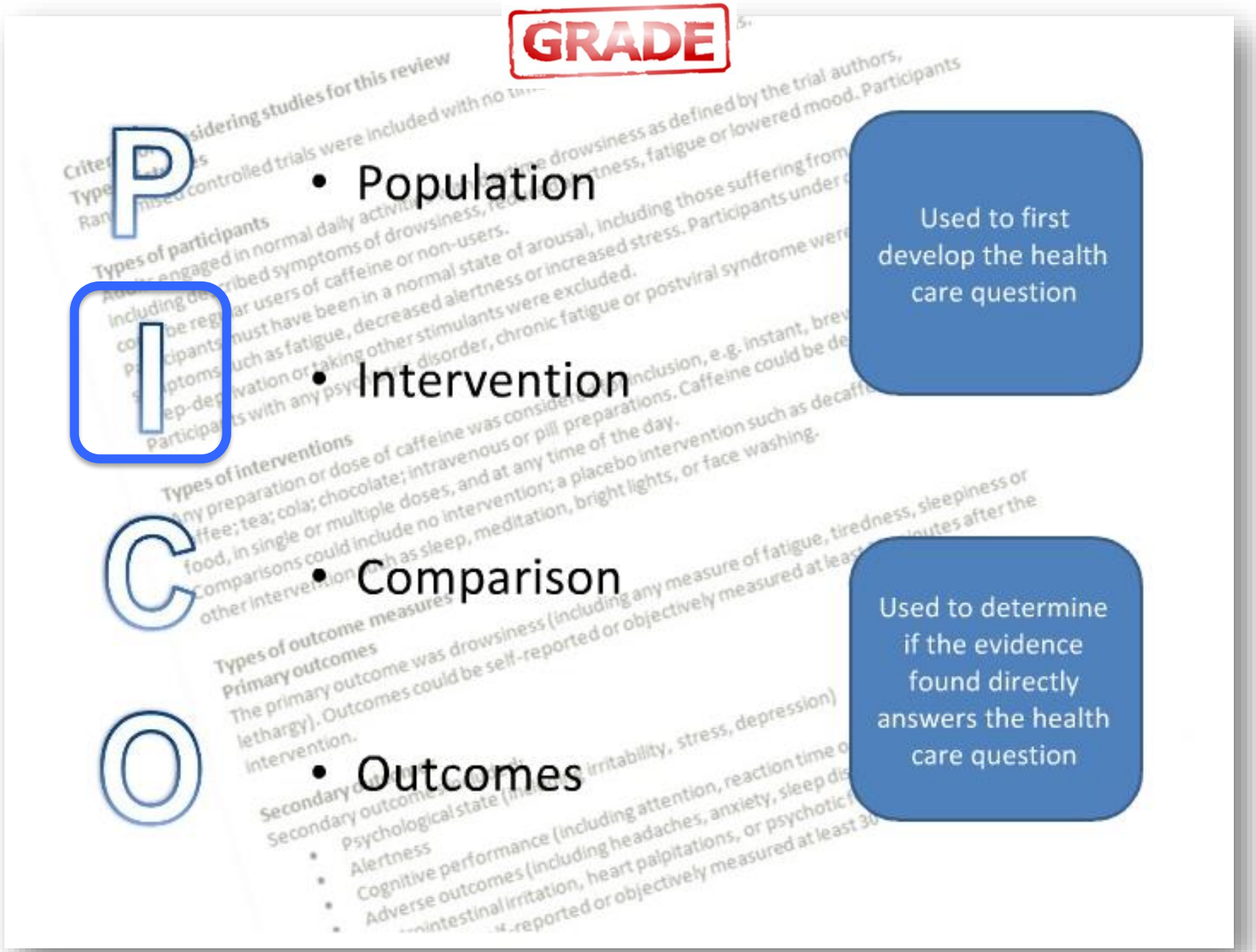
Radical Prostatectomy versus Watchful Waiting in Early Prostate Cancer

Anna Bill-Axelson, M.D., Ph.D., Lars Holmberg, M.D., Ph.D., Mirja Ruutu, M.D., Ph.D., Hans Garmo, Ph.D., Jennifer R. Stark, Sc.D., Christer Busch, M.D., Ph.D., Stig Nordling, M.D., Ph.D., Michael Häggman, M.D., Ph.D., Swen-Olof Andersson, M.D., Ph.D., Stefan Bratell, M.D., Ph.D., Anders Spångberg, M.D., Ph.D., Juni Palmgren, Ph.D., Gunnar Steineck, M.D., Ph.D., Hans-Olov Adami, M.D., Ph.D., and Jan-Erik Johansson, M.D., Ph.D., for the SPCG-4 Investigators*

N Engl J Med 2011;364:1708-17.

Pazienti in era pre-PSA




GRADE
P
• Population
I
• Intervention
C
• Comparison
O
• Outcomes
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question

SACRO CUORE DON CALABRIA I R C C S

Long-term oncologic outcomes of postoperative adjuvant versus salvage radiotherapy in prostate cancer: Systemic review and meta-analysis of 5-year and 10-year follow-up data

Ja Yoon Ku[1], Chan Ho Lee[1], Hong Koo Ha[1,2]

Korean J Urol 2015;56:735-741.

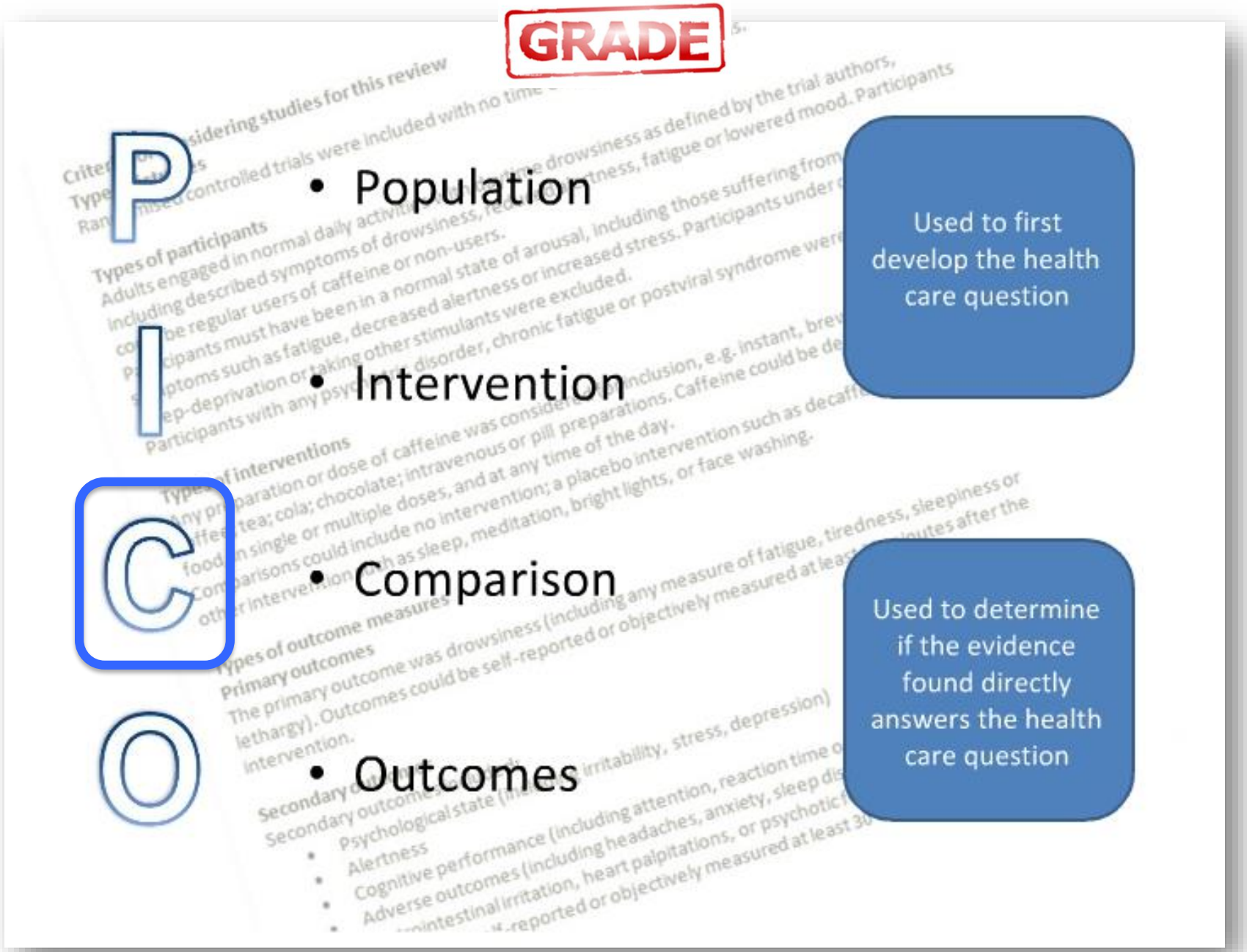
The present systemic review has the following limitations that must be taken into account.

The first limitation is that we heterogeneously recruit randomized controlled studies and retrospective studies: in retrospective studies, the initiation timing of radiotherapy is somewhat different in each study.

The second limitation is that there is a difference in dose and modality (2-dimensional, 3-dimensional, or intensity modulated radiation therapy) of radiation compared to recent practice, which may alter oncologic outcomes.

The third limitation [illegible] were different in eac[illegible] of long-term outcome be[illegible]

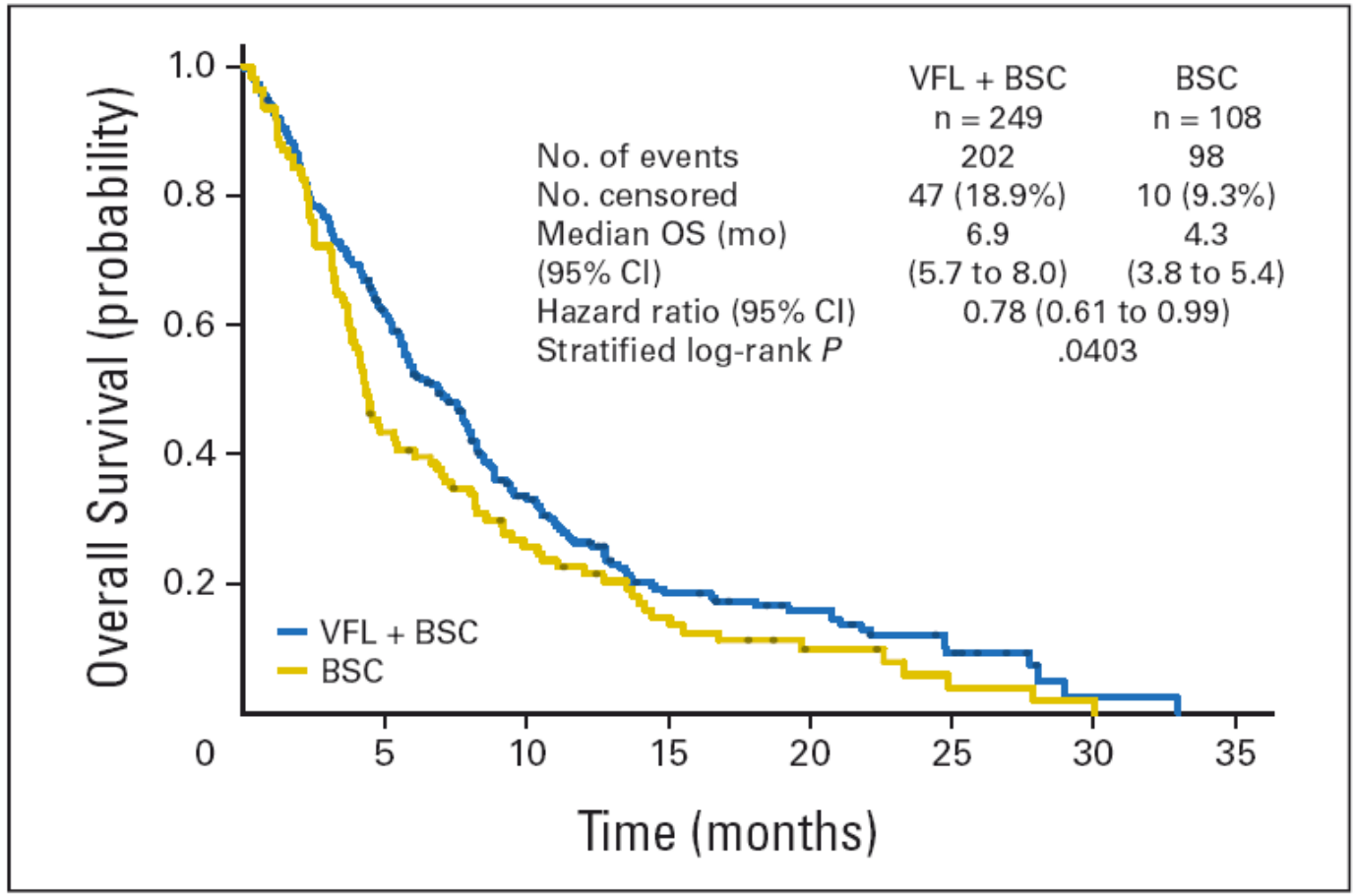
Indirectness per I. (di P.I.C.O.)

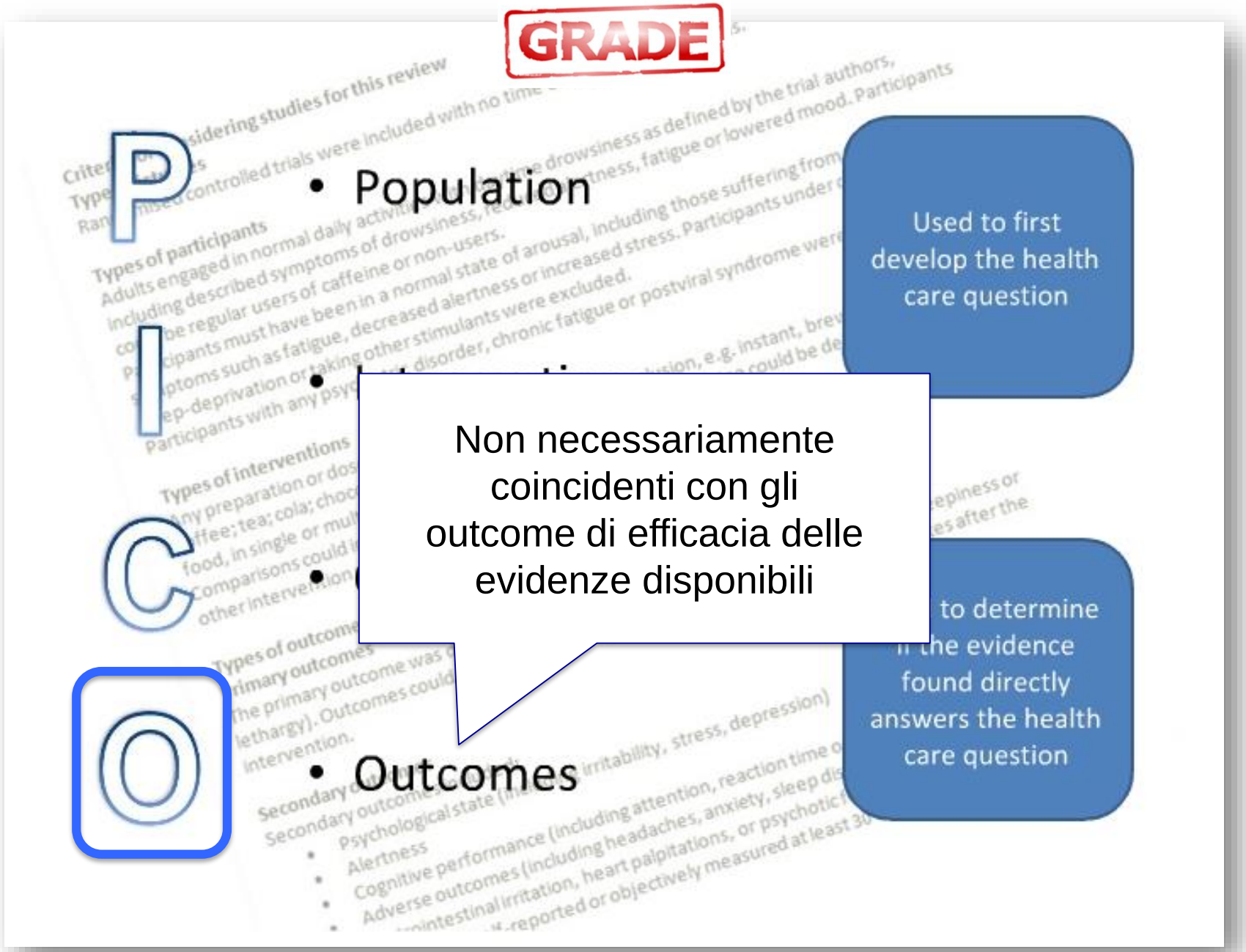

GRADE
P
Population
I
Intervention
C
Comparison
O
Outcomes
Used to first develop the health care question
Used to determine if the evidence found directly answers the health care question

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 by American Society of Clinical Oncology




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

Surrogate outcome markers in research and clinical practice

Scott Twaddell

(Aust Prescr 2009;32:47–50)

Table 1

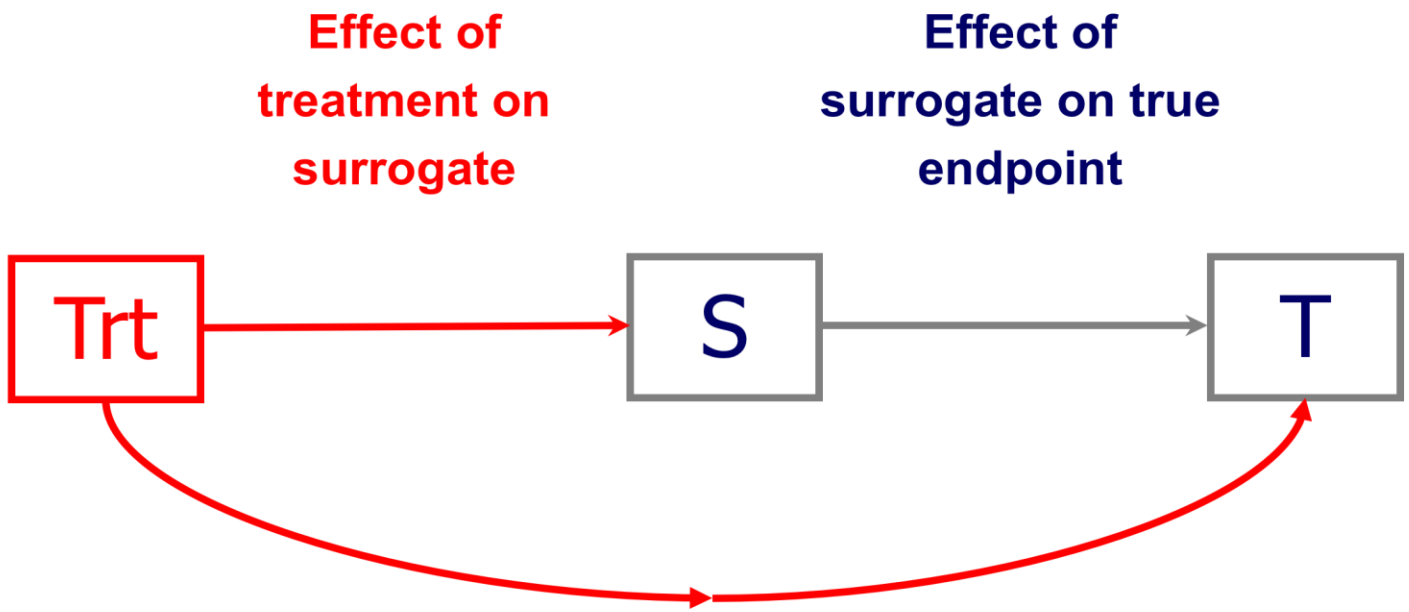
Surrogate markers often used in clinical practice

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

VALIDATION OF SURROGATE ENDPOINTS: "FULL CAPTURE OF EFFECT"



Effect of treatment on true endpoint must be fully captured by surrogate

*Prentice, Statist Med 1989;**8**:431.*

THE IDOLATRY OF THE SURROGATE

Easier to measure surrogate outcomes are often used instead of patient important outcomes such as death, quality of life, or functional capacity when assessing treatments. **John Yudkin, Kasia Lipska, and Victor Montori** argue that our obsession with surrogates is damaging patient care

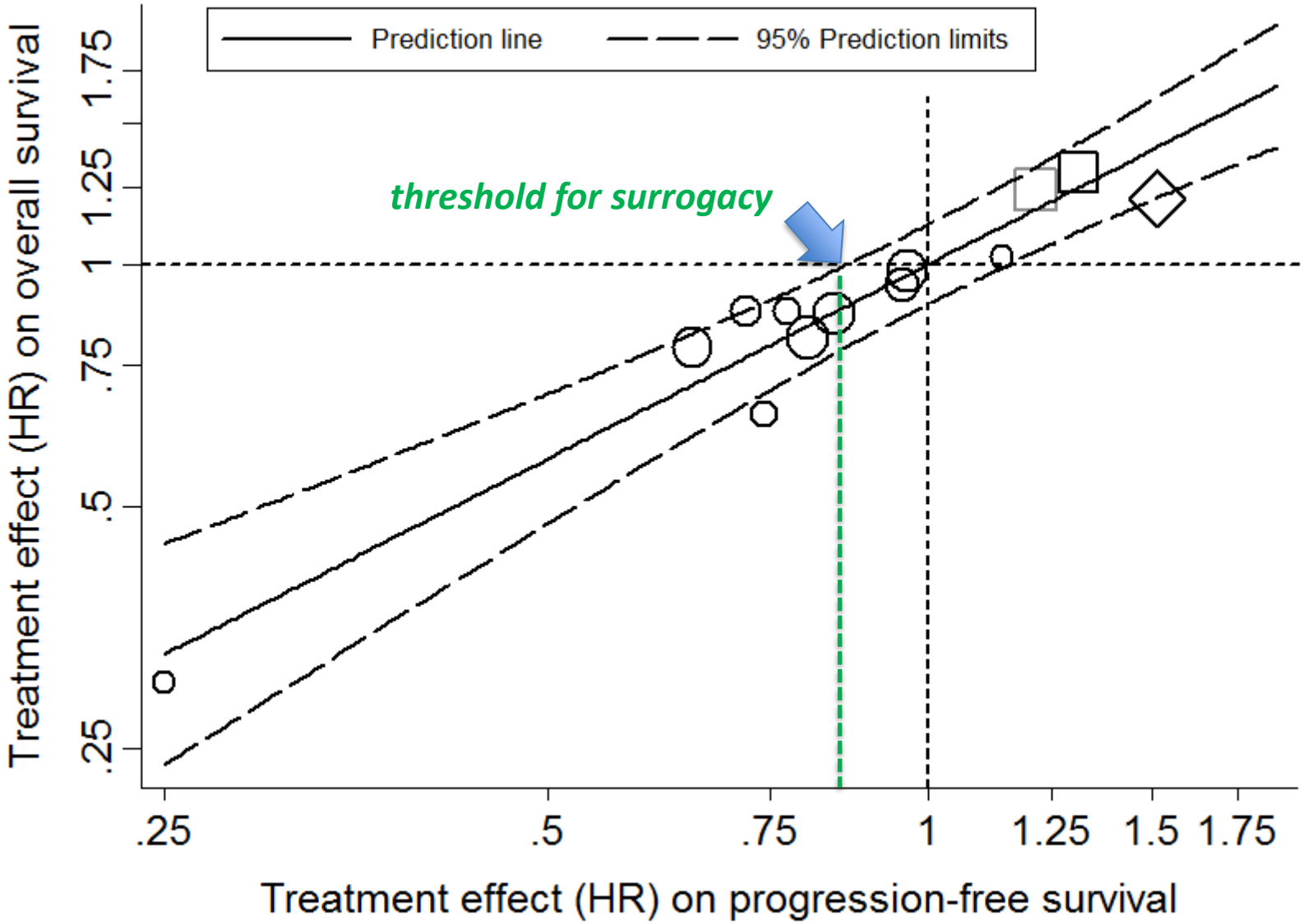
BMJ | 14 JANUARY 2012 | VOLUME 344

Since they respond sooner than outcomes that are important to patients, surrogates are better suited as end points in clinical trials that need to be completed quickly and at low cost

Validation issues

- Comprehensive data are required to assess the validity of a surrogate, preferably a meta-analysis of several randomized trials showing sufficient certainty of results.
- It is not readily possible to transfer conclusions about the validity of surrogates between different diseases or disease grades or between different interventions.
- There is neither a universally applicable measure for surrogate validation, nor a general best estimation method, nor a generally accepted threshold which, if exceeded, would mean proof of validity.

TRIAL LEVEL CORRELATION BETWEEN EFFECTS

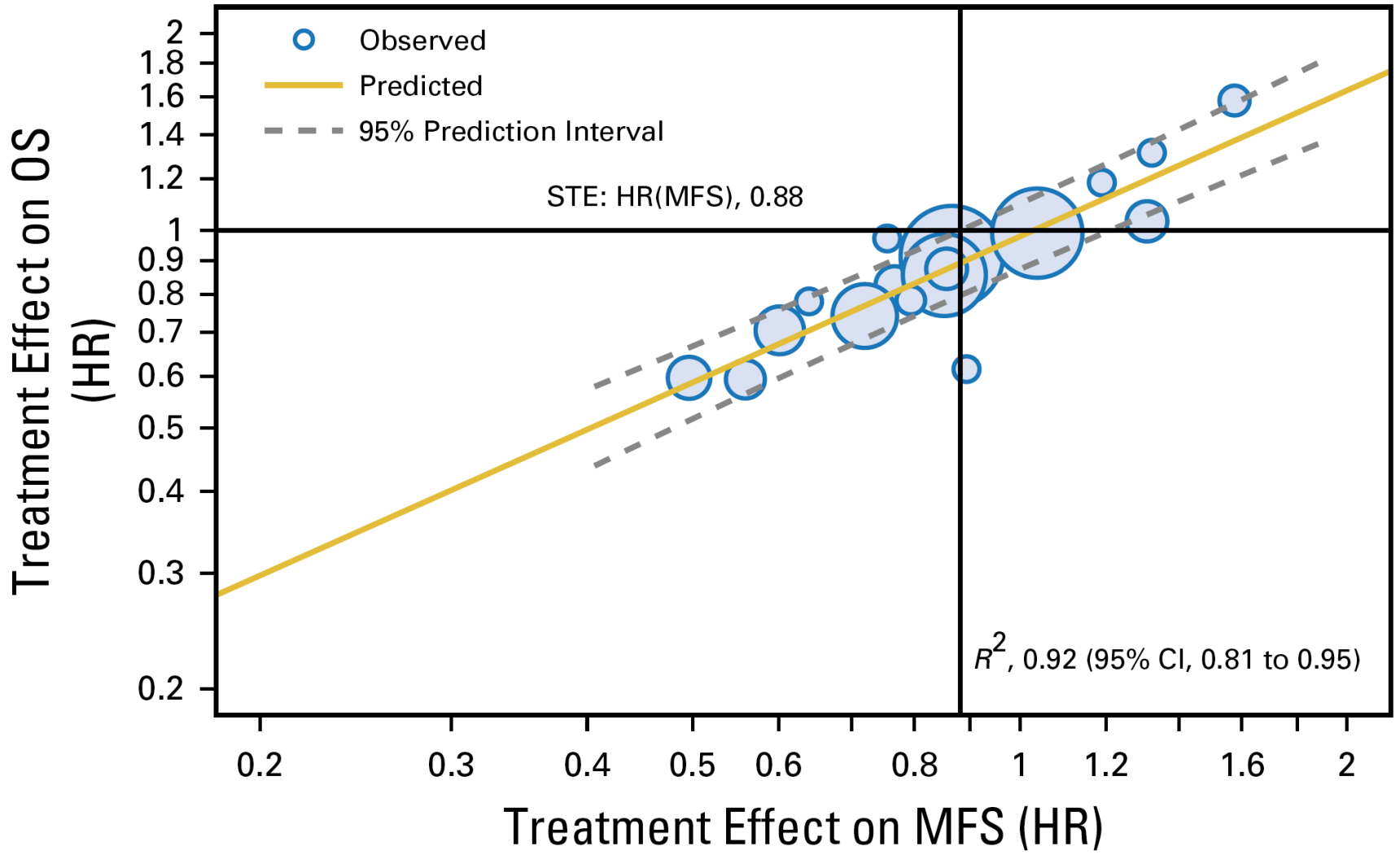


Burzykowski and Buyse, Pharmaceutical Statist 2006;5:173

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



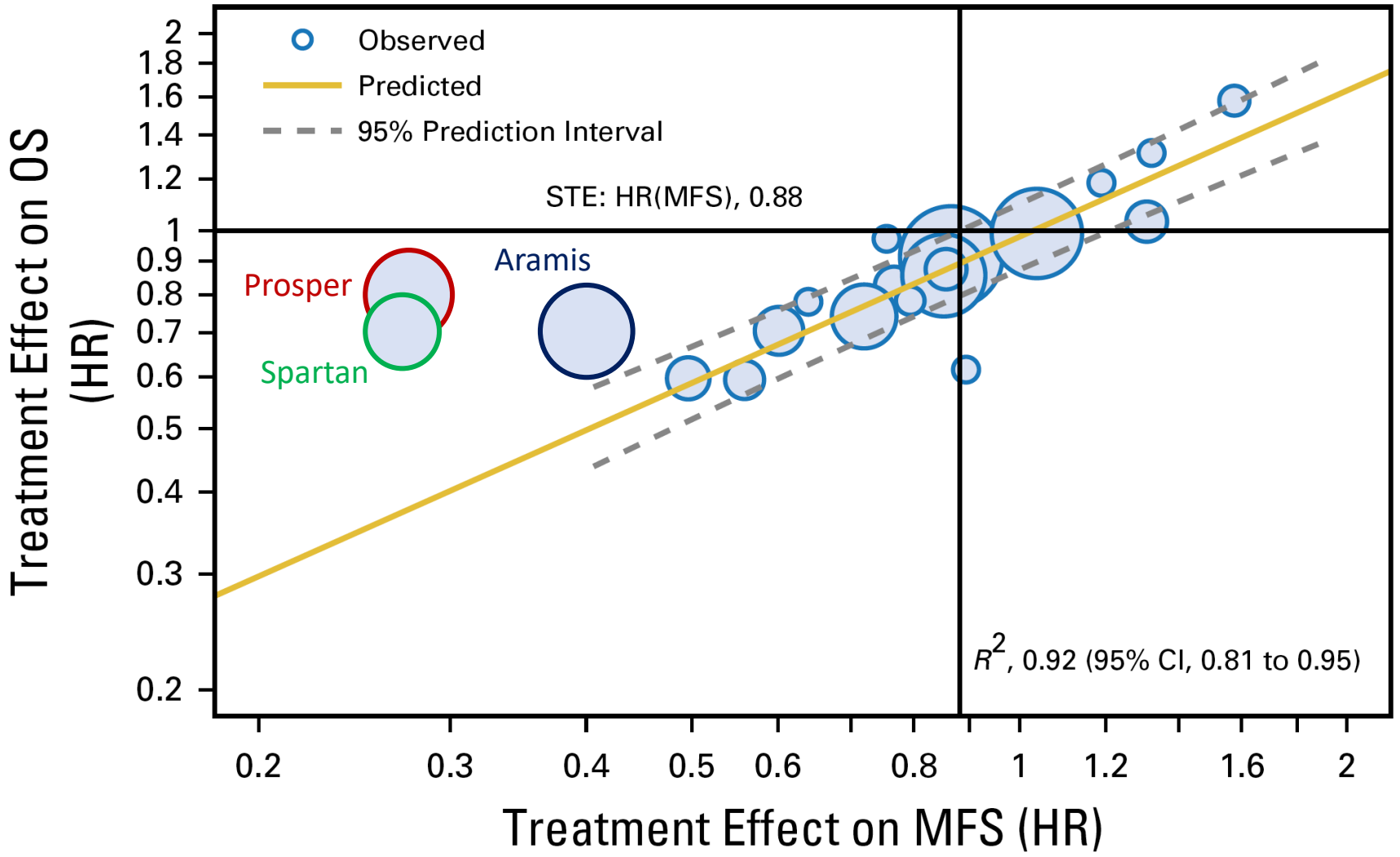
Validation issues

- Comprehensive data are required to assess the validity of a surrogate, preferably a meta-analysis of several randomized trials showing sufficient certainty of results.
- It is not readily possible to transfer conclusions about the validity of surrogates between different diseases or disease grades or between different interventions.
- There is neither a universally applicable measure for surrogate validation, nor a general best estimation method, nor a generally accepted threshold which, if exceeded, would mean proof of validity.

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

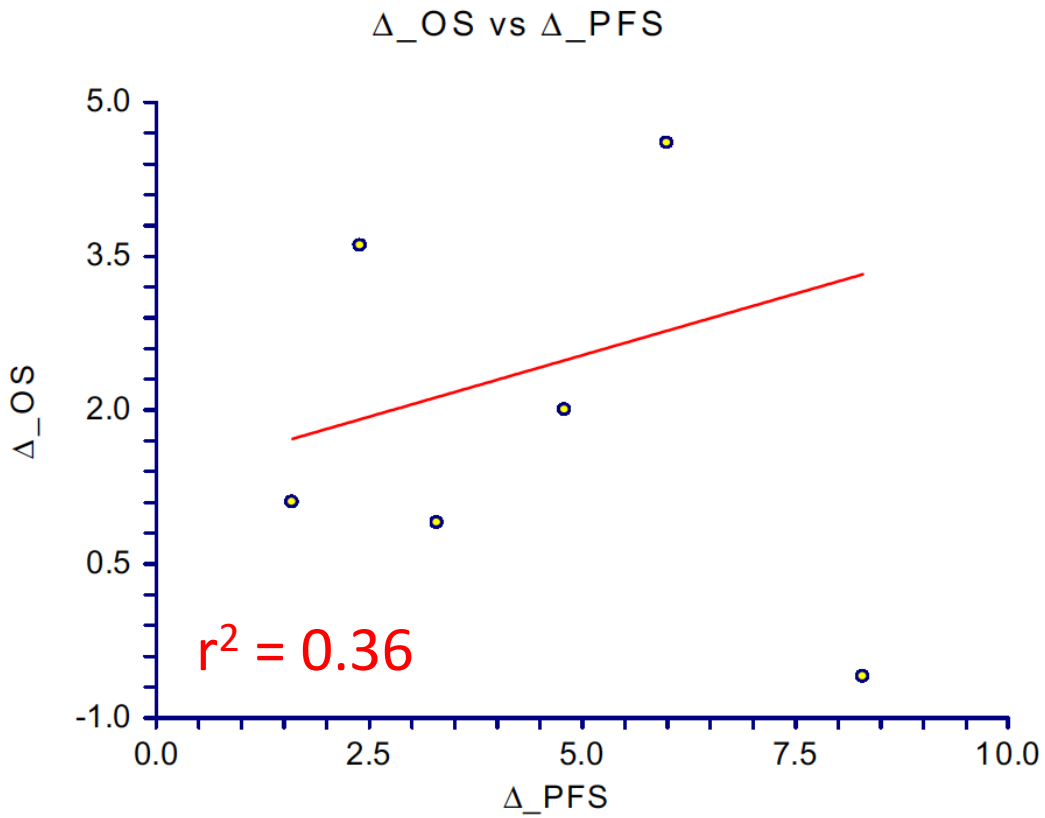
J Clin Oncol 35:3097-3104. © 2017 by American Society of Clinical Oncology

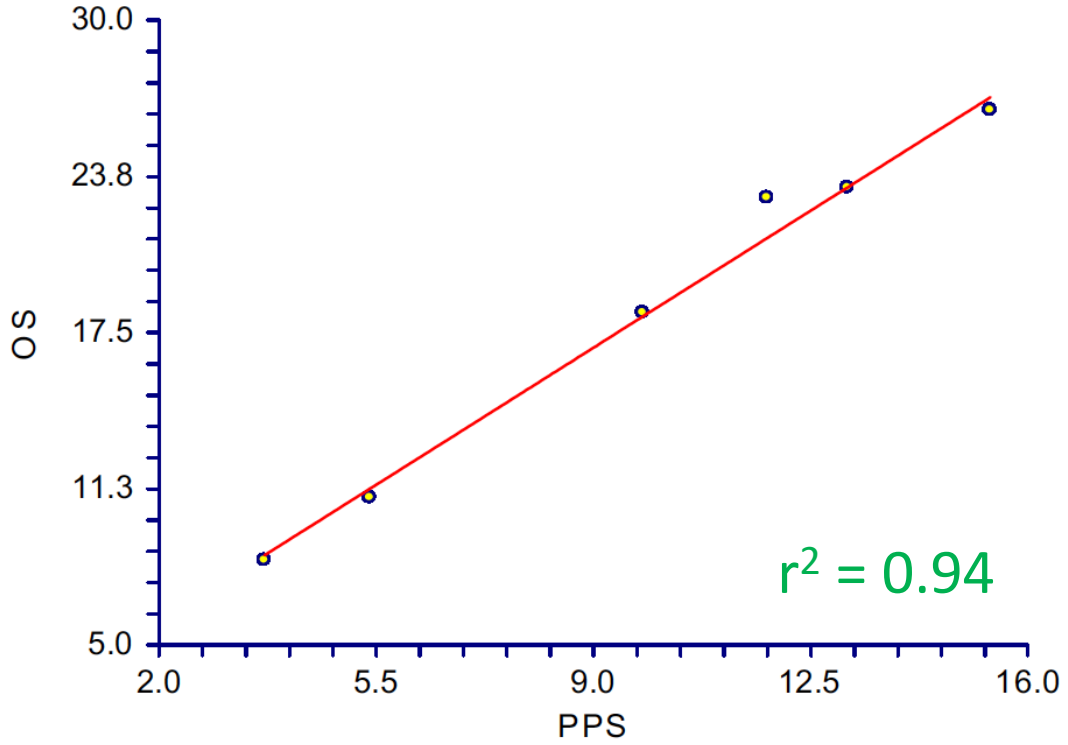


Surrogate End Points and Postprogression Survival in Renal Cell Carcinoma: An Analysis of First-Line Trials With Targeted Therapies

Fausto Petrelli, Sandro Barni

ICI + TKI ?
ICI + ICI ?



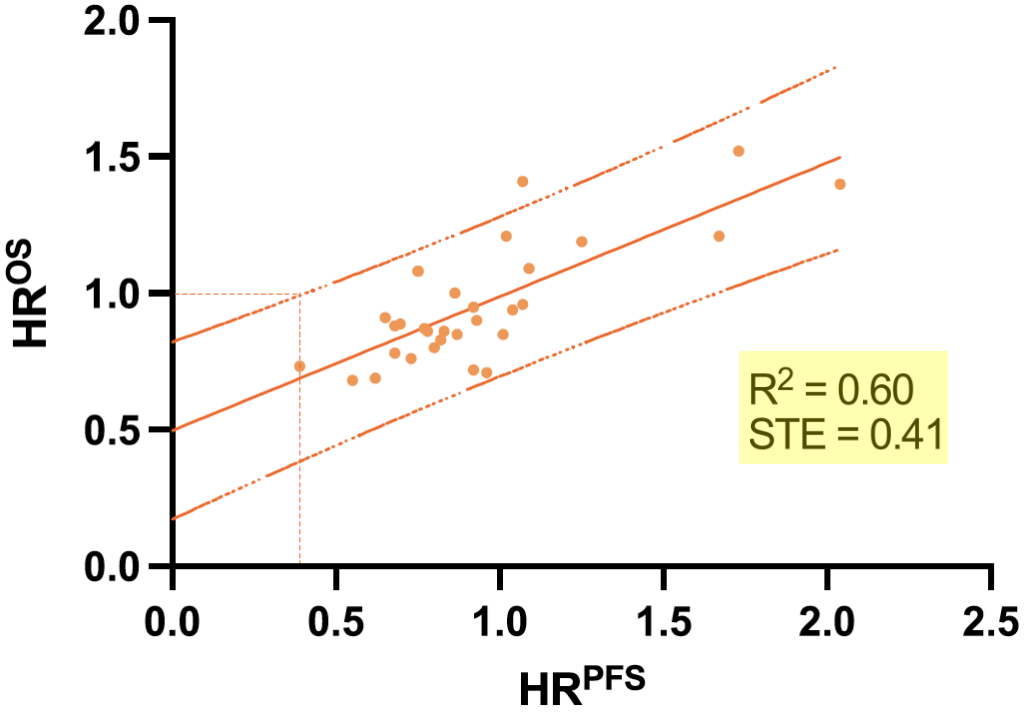


Clinical Genitourinary Cancer, Vol. 11, No. 4, 385-9 © 2013 Elsevier Inc.

EUROPEAN UROLOGY OPEN SCIENCE 47 (2023) 58–64

Surrogate Endpoints as Predictors of Overall Survival in Metastatic Urothelial Cancer: A Trial-level Analysis

Fady Ghali[a,*], *Yibai Zhao*[b], *Devin Patel*[c], *Teresa Jewell*[d], *Evan Y. Yu*[e], *Petros Grivas*[e], *R. Bruce Montgomery*[e], *John L. Gore*[a], *Ruth B. Etzioni*[b], *Jonathan L. Wright*[a]



PFS	R^2 HR^{PFS}	STE HR^{PFS}
All trials	0.60	0.41
No crossover	0.65	0.44
ICI	<0.01	NR
Non-ICI	0.63	0.33
Longer follow-up	0.76	0.59
First line	0.48	0.24
Non–first line	0.74	0.34

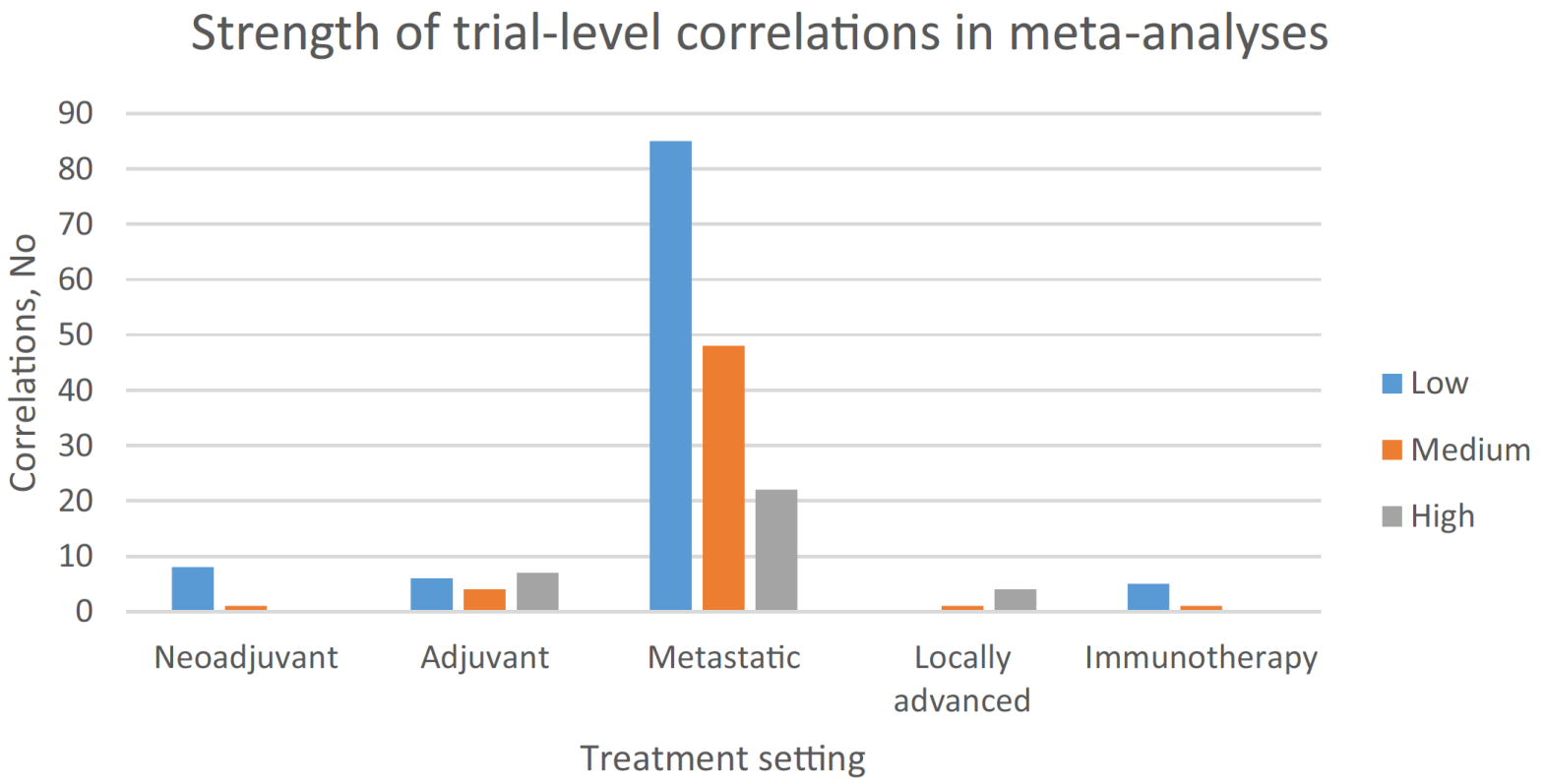
Validation issues

- Comprehensive data are required to assess the validity of a surrogate, preferably a meta-analysis of several randomized trials showing sufficient certainty of results.
- It is not readily possible to transfer conclusions about the validity of surrogates between different diseases or disease grades or between different interventions.
- There is neither a universally applicable measure for surrogate validation, nor a general best estimation method, nor a generally accepted threshold which, if exceeded, would mean proof of validity.

A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology

Alyson Haslam [a,*], Spencer P. Hey [b], Jennifer Gill [a], Vinay Prasad [c,d,e,f]

European Journal of Cancer 106 (2019) 196–211



trial-level correlation scored according to a modification to surrogate criteria pro-posed by the Institute of Quality and Efficiency in Health Care:

- low ($r^2 \leq 0.7$)
- medium-strength ($r^2 > 0.7$ to < 0.85)
- high ($r^2 \geq 0.85$)

MALCOLM WILLETT

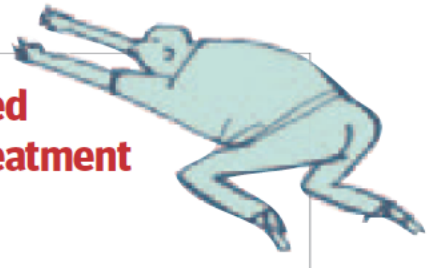
THE IDOLATRY OF THE SURROGATE

Easier to measure surrogate outcomes are often used instead of patient important outcomes such as death, quality of life, or functional capacity when assessing treatments. **John Yudkin, Kasia Lipska, and Victor Montori** argue that our obsession with surrogates is damaging patient care

BMJ | 14 JANUARY 2012 | VOLUME 344

Since they respond sooner than outcomes that are important to patients, surrogates are better suited as end points in clinical trials that need to be completed quickly and at low cost

Surrogate markers are not intrinsically flawed. When interpreted appropriately, they can be helpful in risk stratification and in treatment



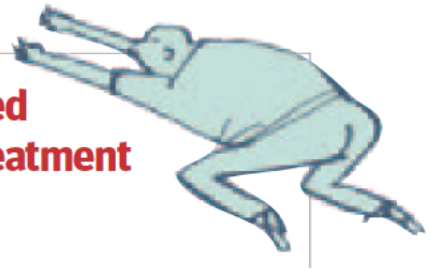
THE IDOLATRY OF THE SURROGATE

Easier to measure surrogate outcomes are often used instead of patient important outcomes such as death, quality of life, or functional capacity when assessing treatments. **John Yudkin, Kasia Lipska, and Victor Montori** argue that our obsession with surrogates is damaging patient care

BMJ | 14 JANUARY 2012 | VOLUME 344

Since they respond sooner than outcomes that are important to patients, surrogates are better suited as end points in clinical trials that need to be completed quickly and at low cost

Surrogate markers are not intrinsically flawed. When interpreted appropriately, they can be helpful in risk stratification and in treatment



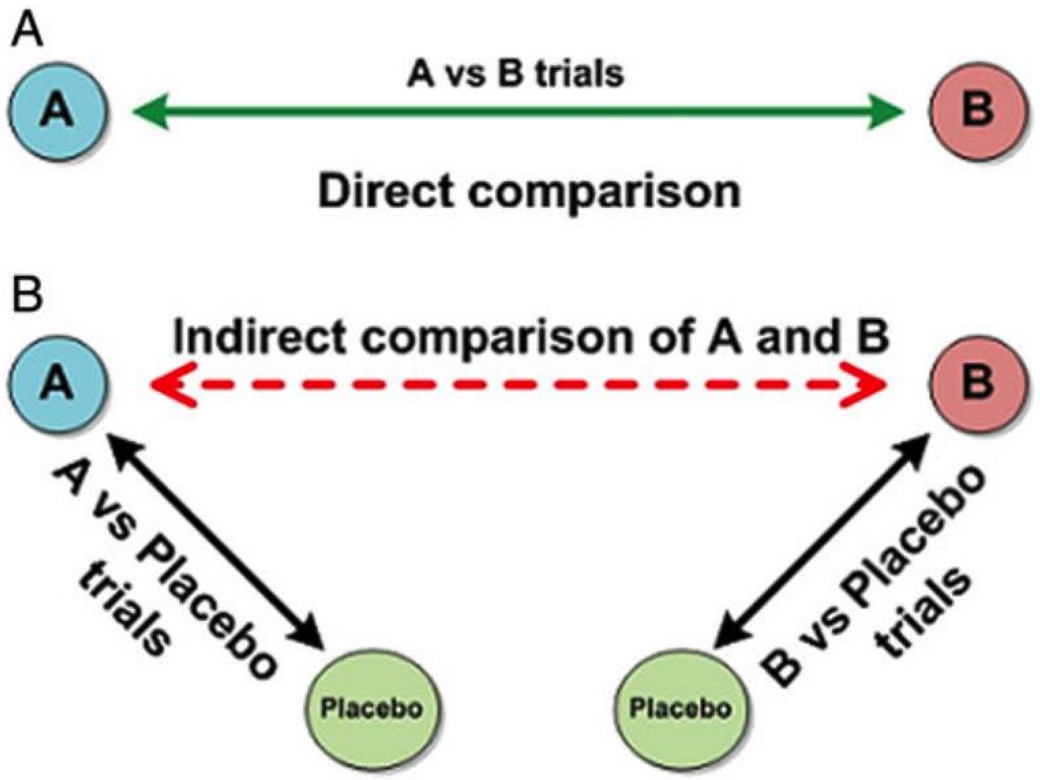
In order to fully engage our patients in treatment decisions, we must understand how therapies affect outcomes that are important to them. Surrogate end-points will not provide us with these answers.

Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.



through a

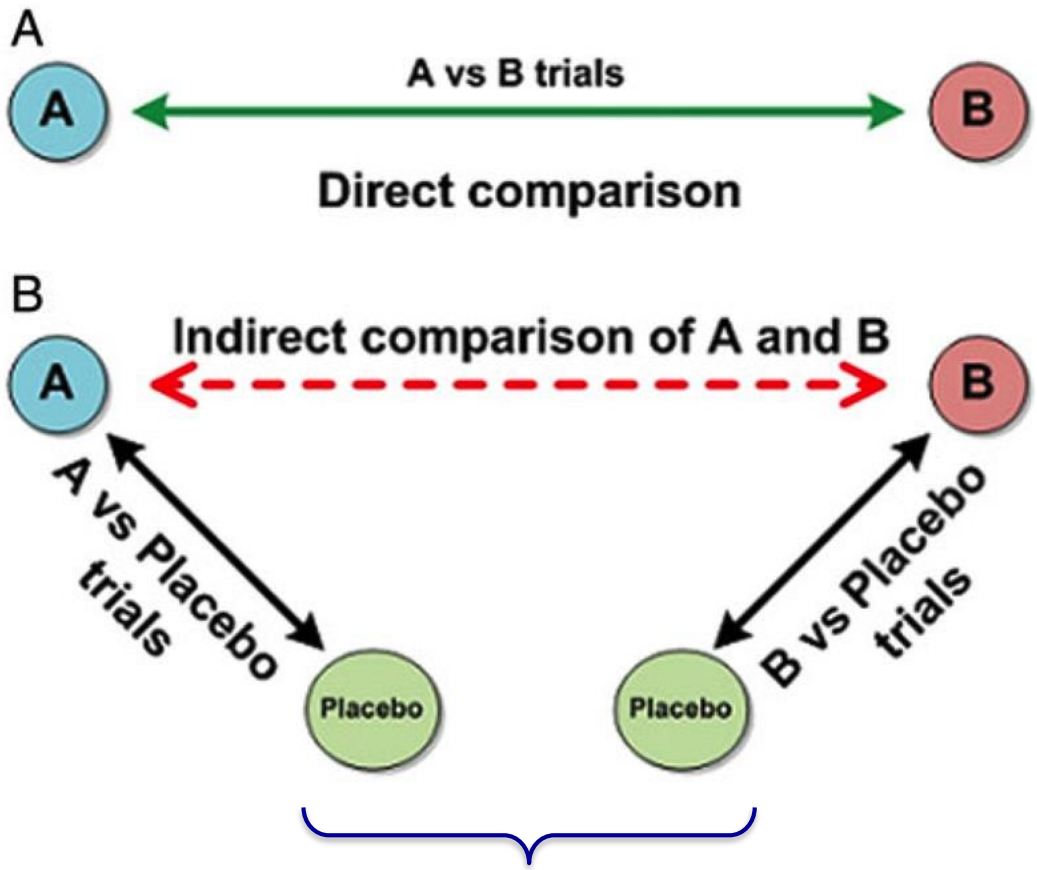
Common Comparator

Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.



Similarity Assumption

trials must be comparable on effect modifiers to obtain an unbiased pooled estimate.

Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.

SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

- Etica e integrità della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- **Affidabilità delle prove (imprecisione degli effetti, rischio di bias)**
- Dialogo tra clinico e metodologo (lettura di uno studio clinico)

Caso Vs Bias

CASO

↓

Errore Random

↓

Risultati Imprecisi

Errore in diminuzione con l'aumentare delle dimensioni del campione

BIAS

↓

Errore Sistematico

↓

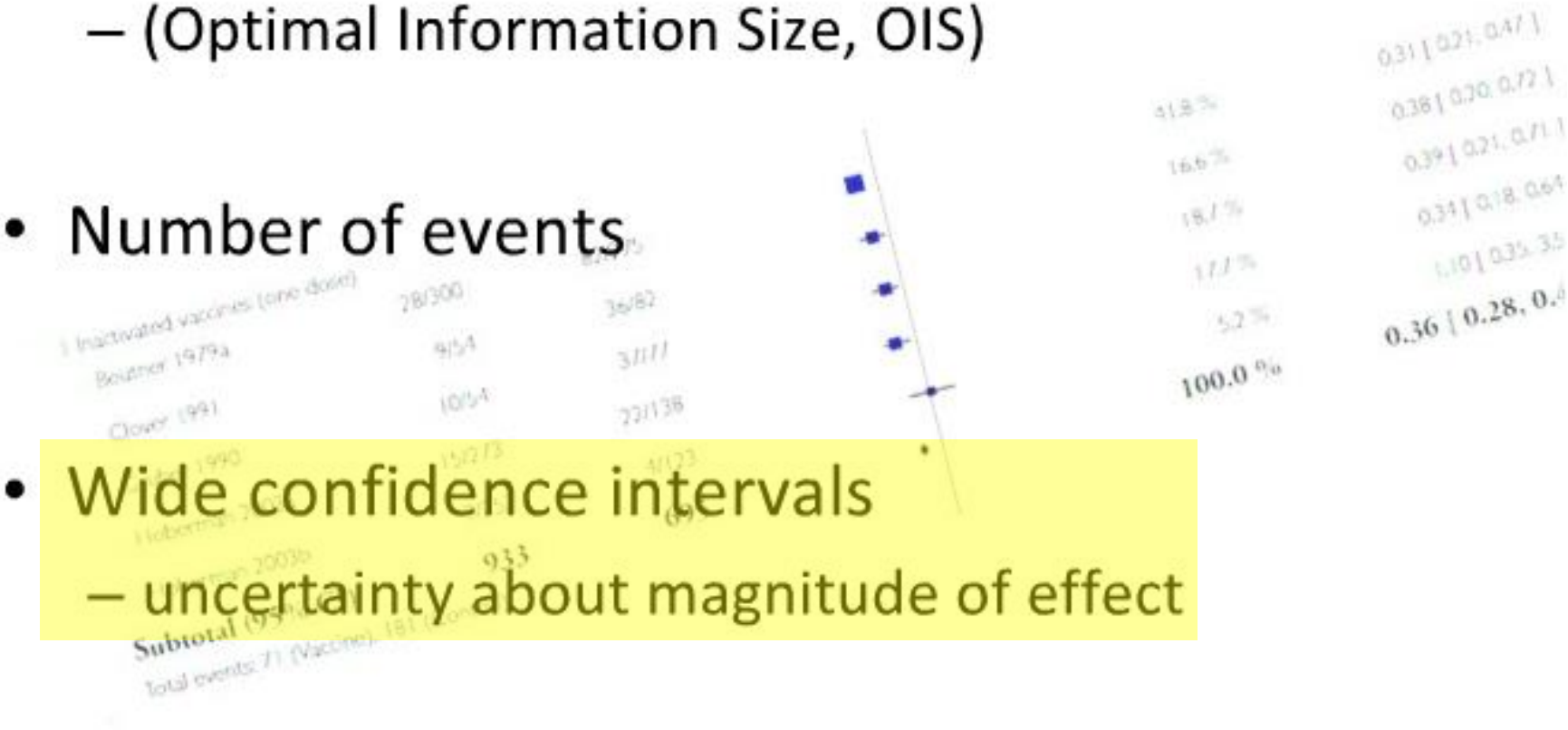
Risultati Inesatti

Errore non influenzato dalle dimensioni del campione

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.

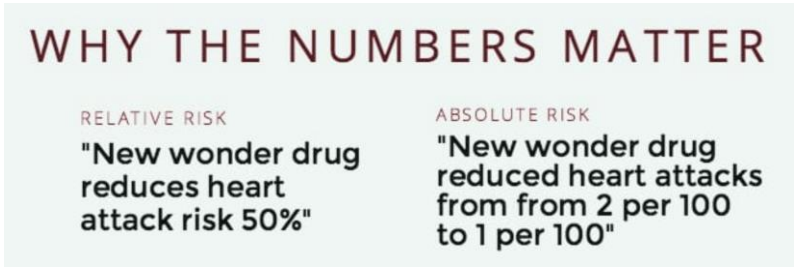
Journal of Clinical Epidemiology 64 (2011) 1283–1293

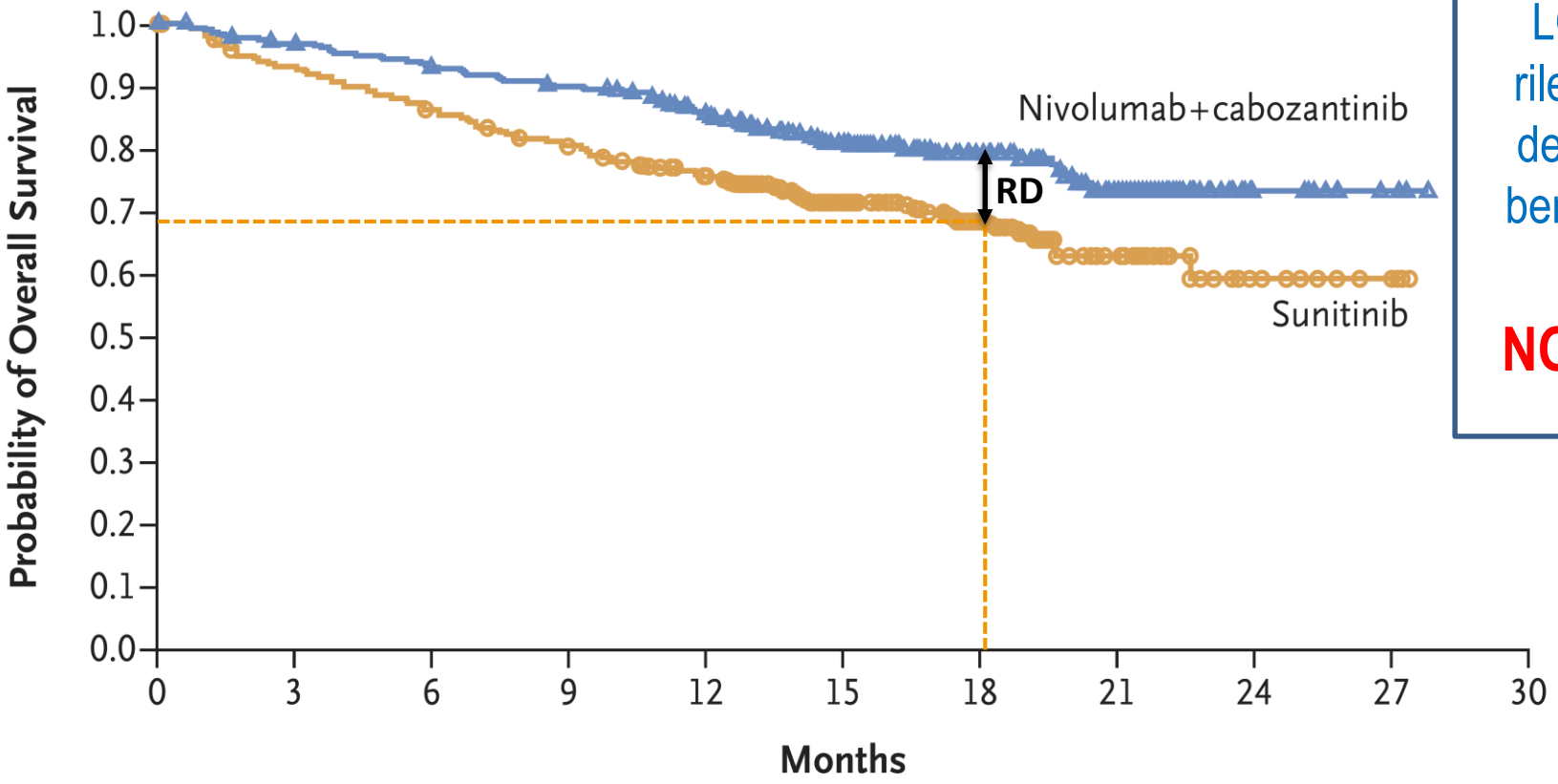
GRADE guidelines 6. Rating the quality of evidence—imprecision

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.





h, 0.60
0.89)

ns

n f.u.: 31%

LC95% consistenti con una rilevante efficacia (↓15 eventi decesso / 100 paz.) e con un beneficio marginale (↓5 eventi decesso / 100 paz.)

NON seria imprecisione

- **Risk Difference** 11 events lower / 100 pts **(95%CI: 15 lower to 5 lower)**

N Engl J Med 2021;384:829-41.

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

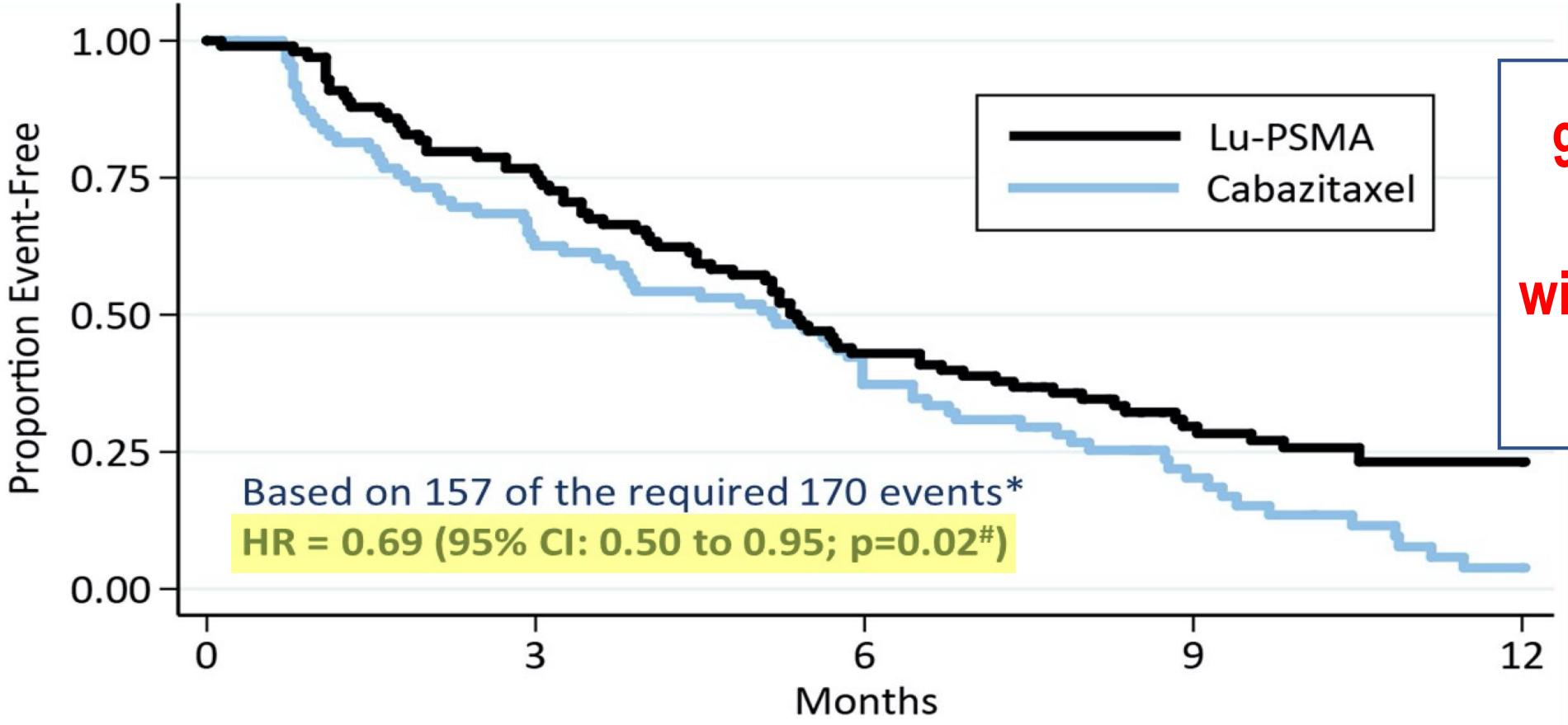
PRESENTED AT: 2020 ASCO ANNUAL MEETING #ASCO20 #TheraP

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

TheraP (ANZUP 1603)

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

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



95%CI of absolute effect consistent with different clinical interpretations

SACRO CUORE DON CALABRIA I R C C S

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

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

Scale	Cabazitaxel Pred. Mean (SE) {95% CI}	Lu-PSMA Pred. Mean (SE) {95% CI}	Diff. 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

Caso Vs Bias

CASO

↓

Errore Random

↓

Risultati Imprecisi

Errore in diminuzione con l'aumentare delle dimensioni del campione

BIAS

↓

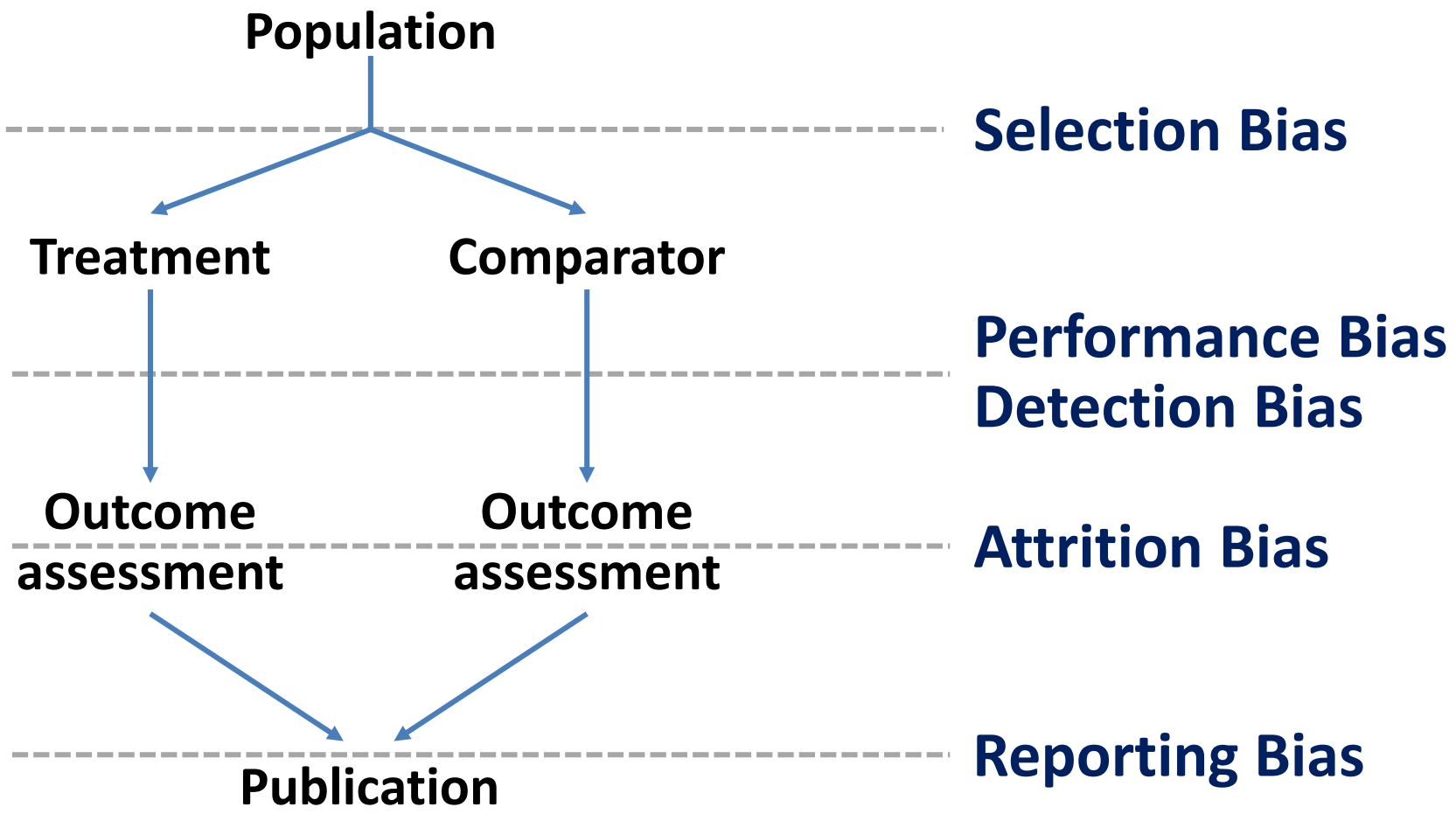
Errore Sistematico

↓

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

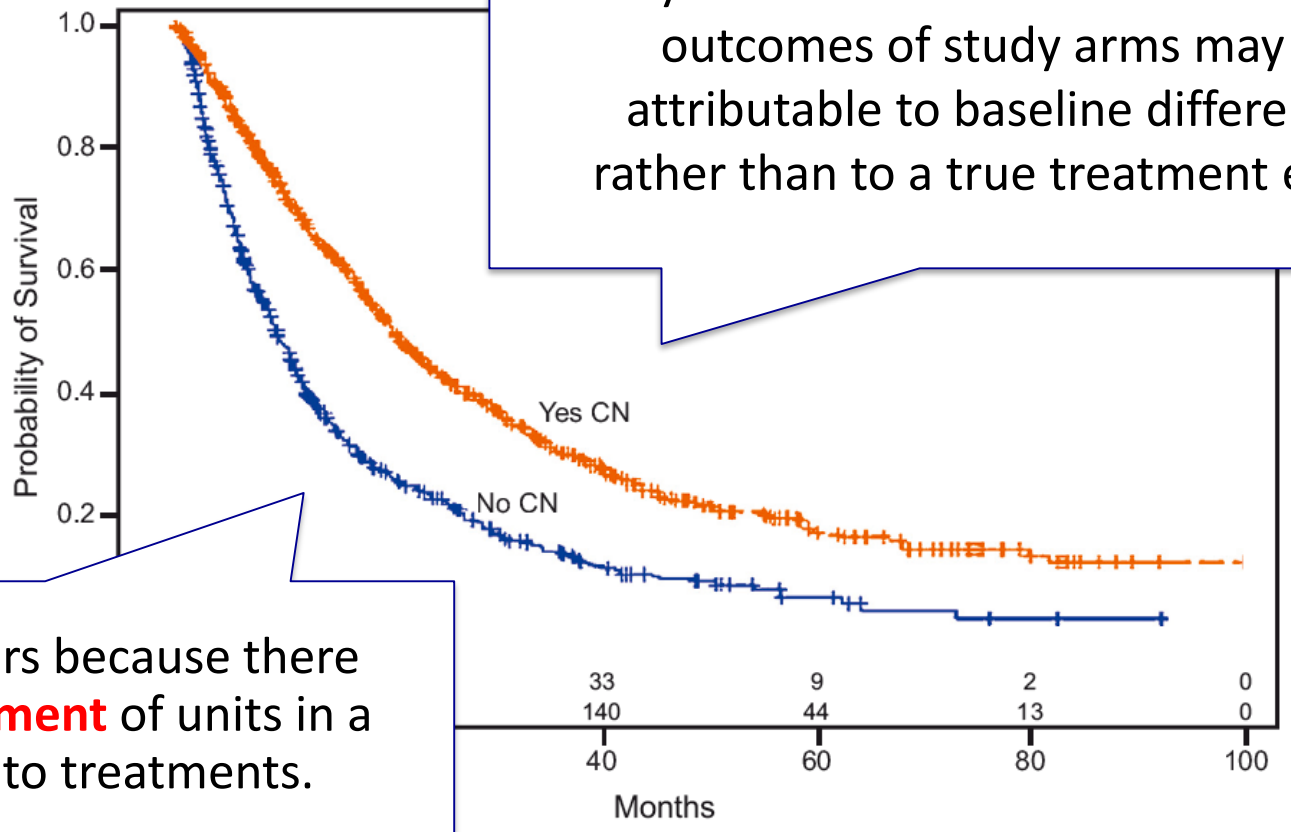
Version 5.1.0

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.

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

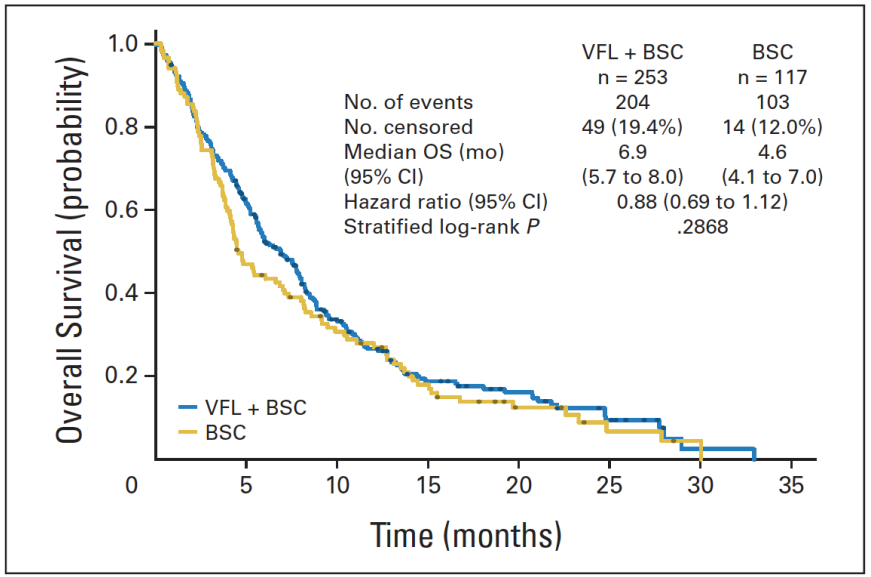


Fig 2. Overall survival (OS) in the intent-to-treat population (n = 370). VFL, vinflunine; BSC, best supportive care.

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

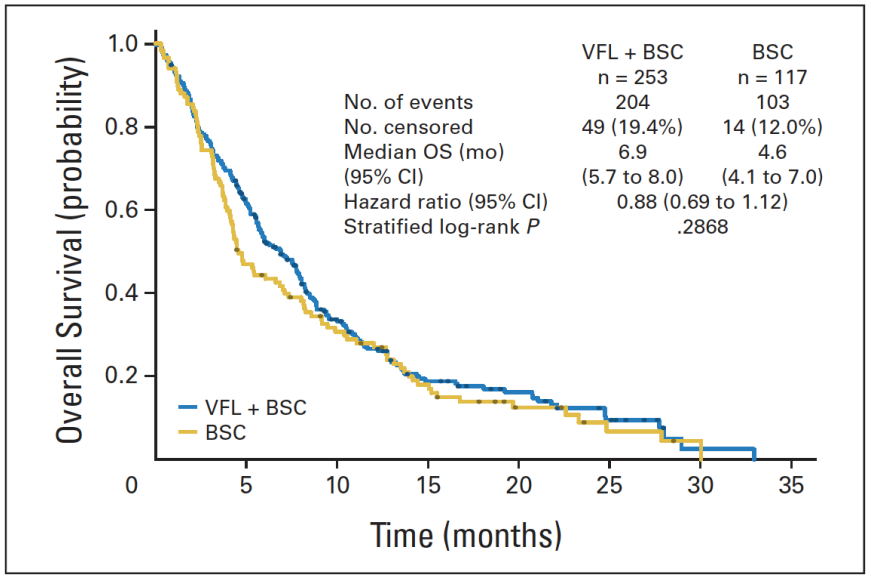


Fig 2. Overall survival (OS) in the intent-to-treat population (n = 370). VFL, vinflunine; BSC, best supportive care.

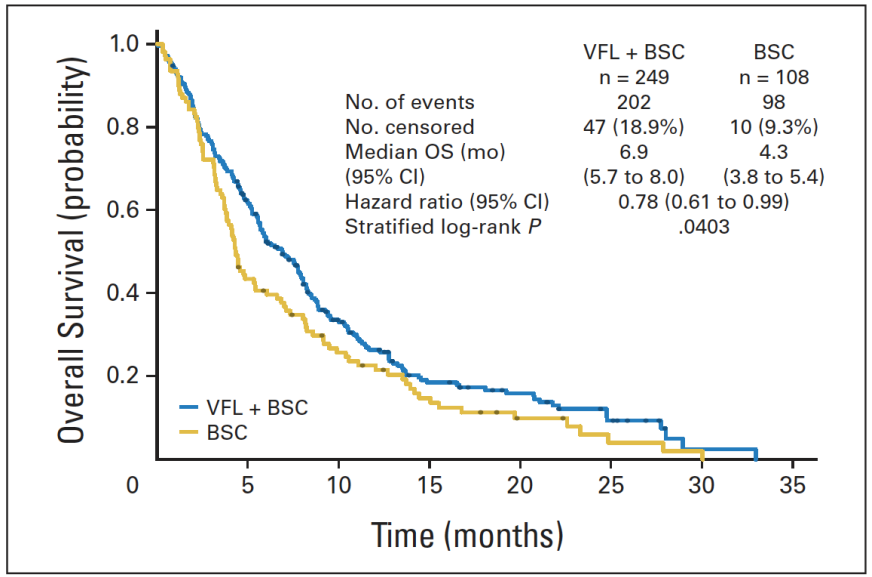


Fig 3. Overall survival (OS) in the eligible population (n = 357; 96.5% of intent-to-treat population). VFL, vinflunine; BSC, best supportive care.

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).

What is Selective Crossover?

- Special case of *non-adherence to a randomized treatment* following the report of *positive trial results*: control group patients selectively cross over to the experimental treatment
- *Disturbs the randomized comparison* in updated analyses performed subsequent to the first results

BIG 1-98 Monotherapy Update

Median Follow-up 76 months

	Hazard Ratio	Events Letrozole N=2463	Events Tamoxifen N=2459	Hazard Ratio (95% CI)	ITT P-value
Disease-free Survival					
Intent-to-treat*		509	565	0.88 (0.78–0.99)	0.03
Censored				0.84 (0.74–0.95)	
Overall Survival					
Intent-to-treat		303	343	0.87 (0.75–1.02)	0.08
Censored				0.81 (0.69–0.94)	
Time to distant recurrence					
Intent-to-treat		257	298	0.85 (0.72–1.00)	0.05
Censored				0.81 (0.68–0.96)	

0.5 0.75 1.0 1.25 1.5

Favors Letrozole Favors Tamoxifen

International Breast Cancer Study Group
IBCSG SABCS 2008

*Let:Tam: breast cancer events, 321:363
second (non breast) malignancy, 101:115
deaths without prior cancer event, 87:87

BIG
BREAST INTERNATIONAL GROUP

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?

Rischio di bias legato all'assenza di mascheramento

END POINTS AND ASSESSMENTS

The primary end point was progression-free survival, which was defined as the time from randomization to documented disease progression (as evaluated by independent central review by radiologists who were unaware of the treatment assignments) or death from any cause. Secondary end points included the objective response rate, overall survival (defined as the time from randomization to death from any cause), safety, and the side-effect profile.

Basso rischio di Detection Bias
(valutazione indipendente in cieco)

Alto rischio di Detection Bias

Alto rischio di Performance & Detection Bias

Basso rischio di Detection Bias
(per caratteristica intrinseca dell'outcome)

Lancet 2021; 397: 797–804

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

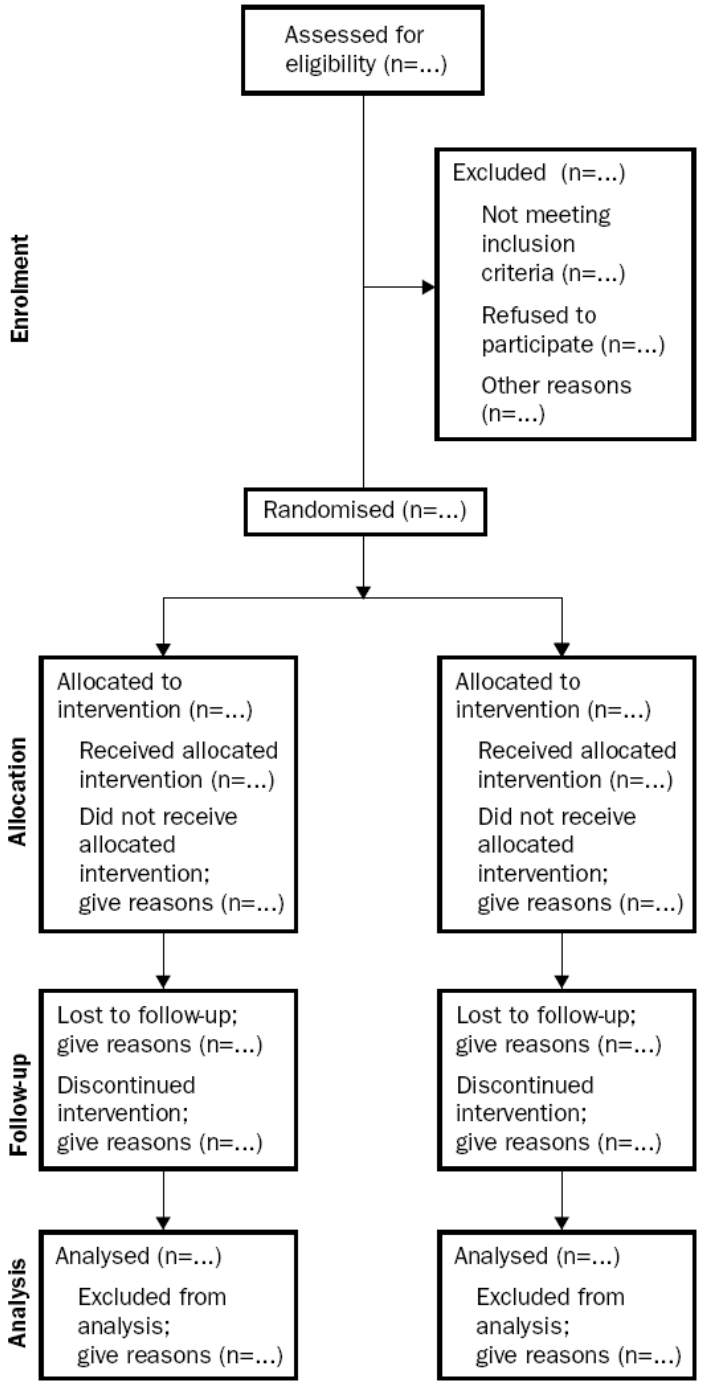
TRANSPARENT REPORTING of TRIALS

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*

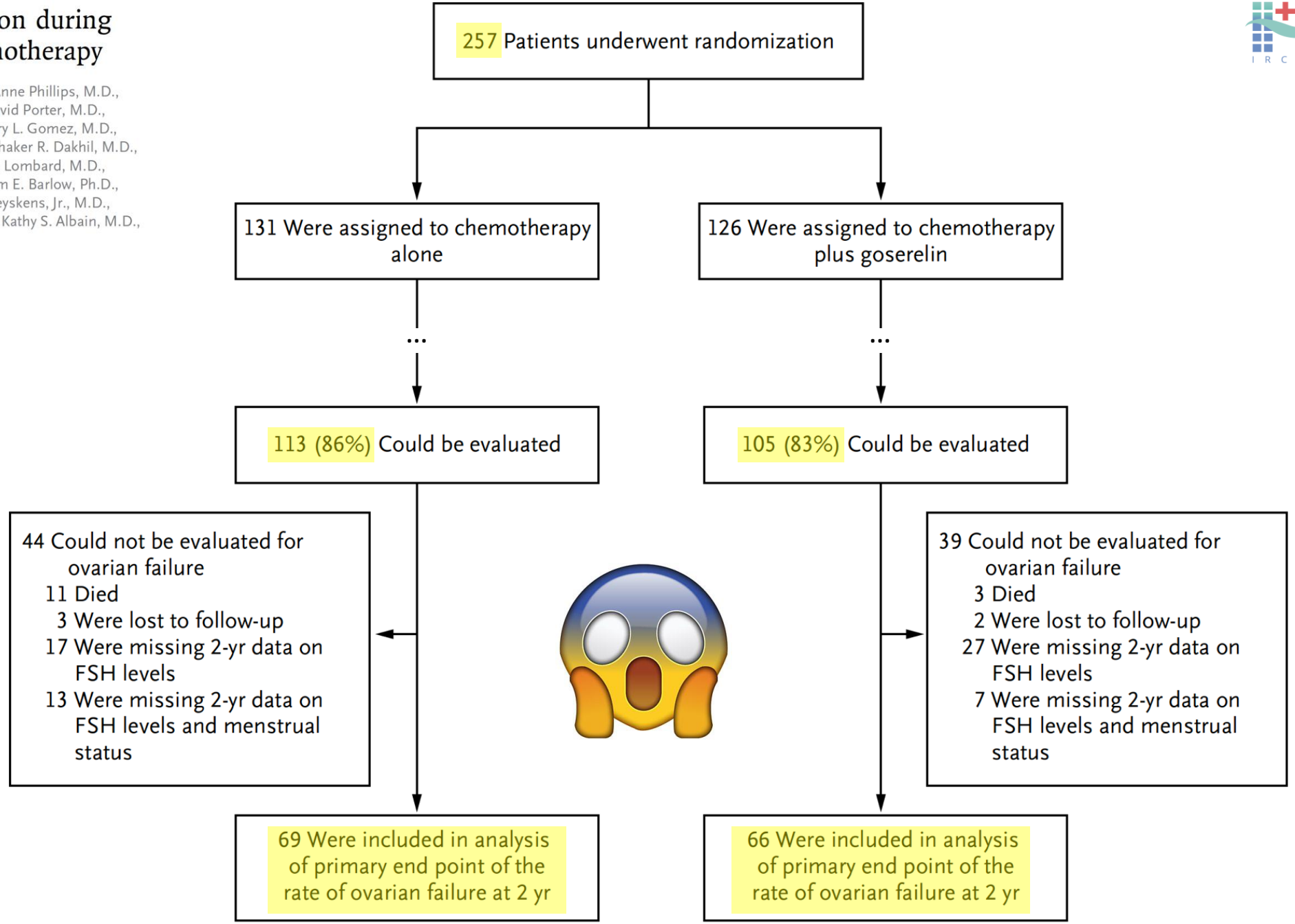
BMJ 2013;346:f1798 doi: 10.1136/bmj.f1798 (Published 22 April 2013)

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.

Goserelin for Ovarian Protection during Breast-Cancer Adjuvant Chemotherapy

Halle C.F. Moore, M.D., Joseph M. Unger, Ph.D., Kelly-Anne Phillips, M.D., Frances Boyle, M.B., B.S., Ph.D., Erika Hitre, M.D., David Porter, M.D., Prudence A. Francis, M.D., Lori J. Goldstein, M.D., Henry L. Gomez, M.D., Carlos S. Vallejos, M.D., Ann H. Partridge, M.D., M.P.H., Shaker R. Dakhil, M.D., Agustin A. Garcia, M.D., Julie Gralow, M.D., Janine M. Lombard, M.D., John F. Forbes, M.B., B.S., Silvana Martino, D.O., William E. Barlow, Ph.D., Carol J. Fabian, M.D., Lori Minasian, M.D., Frank L. Meyskens, Jr., M.D., Richard D. Gelber, Ph.D., Gabriel N. Hortobagyi, M.D., and Kathy S. Albain, M.D., for the POEMS/S0230 Investigators

N Engl J Med 2015;372:923-32.



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

SACRO CUORE DON CALABRIA I R C C S

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.

✗

Riassumendo:

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

  ***rimedio:* randomizzazione (e stratificazione)**

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

  ***rimedio:* mascheramento**

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

  ***rimedio:* disegno di studio e procedure di follow-up**

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

  ***rimedio:* database degli studi clinici in corso**

SACRO CUORE DON CALABRIA

ECI CANCER CENTRE

I R C C S

SCUOLA DI METODOLOGIA DELLA RICERCA CLINICA

2024 - 10ª EDIZIONE

FORMAZIONE DI BASE (Core School) - B1 - 1° Modulo

VENERDÌ 26 - SABATO 27 GENNAIO 2024

NEGRAR DI VALPOLICELLA (VR)

Centro Formazione IRCCS "Sacro Cuore - Don Calabria"

- Etica della ricerca
- Quesito clinico di riferimento
- Fasi della sperimentazione clinica
- Disegni di studio osservazionali e interventistici
- Misure di effetto relativo e assoluto
- Principi di statistica medica (errori statistici, verifica di ipotesi, dimensionamento campionario)
- Rilevanza clinica Vs significatività statistica
- Trasferibilità delle evidenze al quesito clinico
- Affidabilità delle prove (imprecisione degli effetti, rischio di bias)
- **Dialogo tra clinico e metodologo (dal quesito alle evidenze…)**