con il Patrocinio dell'Associazione Italiana di Oncologia Medica

Aiom

Progetto **CANOA**

# CARCINOMA MAMMARIO:

## QUALI NOVITÀ PER IL 2013?

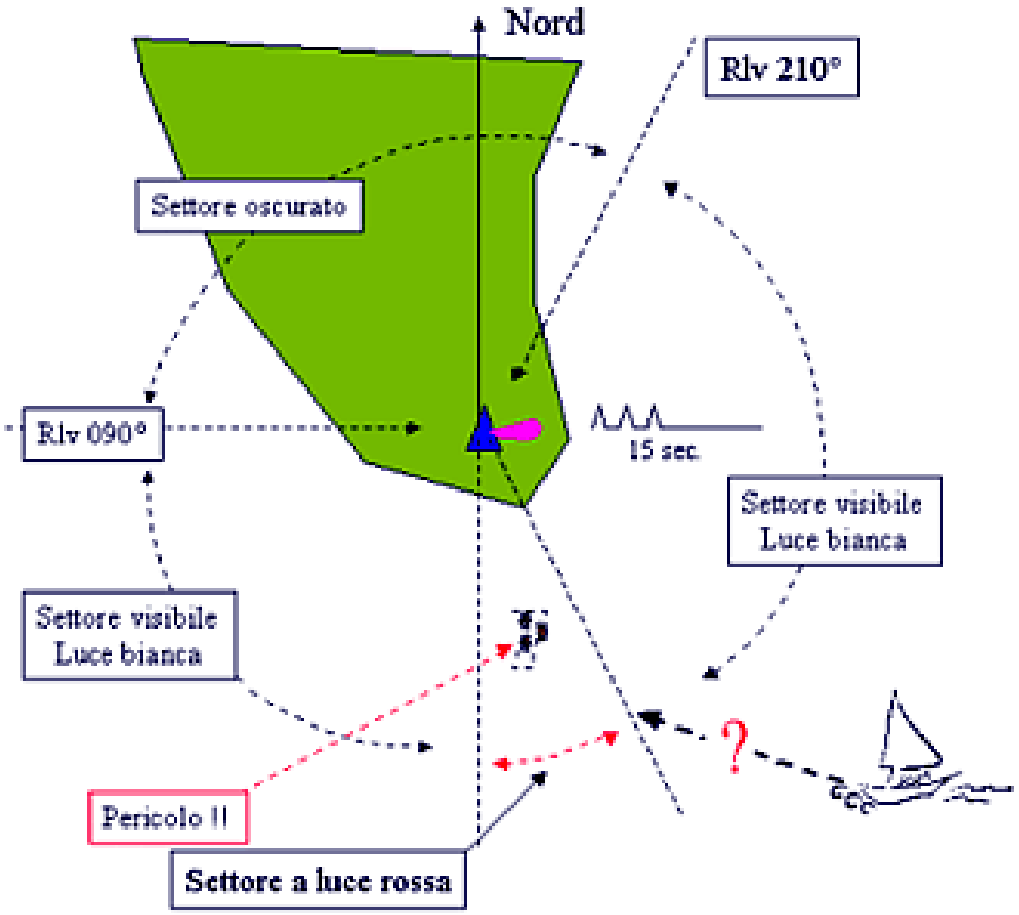
"Saper leggere" uno studio clinico per migliorare la pratica clinica


Coordinatori scientifici:
Stefania Gori
Giovanni L. Pappagallo

Comitato Scientifico:
Emilio Bria
Massimo Di Maio
Jennifer Foglietta
Alessia Levaggi


Negrar - Verona 22-23 marzo 2013
Ospedale Sacro Cuore - Don Calabria


ISTITUTO IDROGRAFICO DELLA MARINA
Elenco dei Fari
e Segnali da Nebbia
Genova 2011


Nord
Rlv 210°
Settore oscurato
Rlv 090°
15 sec
Settore visibile
Luce bianca
Settore visibile
Luce bianca
?
Pericolo !!
Settore a luce rossa

# Statistical Methods

- **Non inferiority randomized trial**
  - 2% variation in te[illegible]f absolute difference of recurrence
  - The 95% CI HR r[illegible]d not cross the 1.15 boundary
  - 1040 D[illegible]
    4 years
  - HR wer[illegible]
- **Accru[illegible]**

**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**

# Cardiac toxicity

| | 12 months (n=1690) | 6 months (n=1690) | P |
|---|---|---|---|
| Cardiac events* | 5.7% | 1.9% | <0.0001 |
| | | | |
| LVEF** < 50% | | | .04 |
| LVEF** < 50% and | | | 071 |
| LVEF** > 50% and | | | NS |

**RR 6 mesi / 12 mesi 0.3333**
**(LC95% 0.2247 – 0.4945)**

**↓185 eventi / 1000 pazienti**

* Investigator reported events (composite with clinical and LFEV finding)
** Based on more than > 25,000 assessments

# Cardiac toxicity

| | 12 months (n=1690) | 6 months (n=1690) | P |
|---|---|---|---|
| Cardiac events* | 5.7% | 1.9% | <0.0001 |
| | | | |
| LVEF** < 50% | 6.3% | 4.7% | 0.04 |
| LVEF** < 50% and ↓ > 10% | 4.8% | 3.6% | 0.071 |
| LVEF** > 50% and | | | NS |

**RR 6 mesi / 12 mesi 0.7453**
**(LC95% 0.5614 – 0.9895)**

**↓78 eventi / 1000 pazienti**

finding)

Based on more than > 25,000 assessments

# Cardiac toxicity

| | 12 months (n=1690) | 6 months (n=1690) | P |
|---|---|---|---|
| Cardiac events* | 5.7% | 1.9% | <0.0001 |
| | | | |
| LVEF** < 50% | 6.3% | 4.7% | 0.04 |
| LVEF** < 50% and ↓ > 10% | 4.8% | 3.6% | 0.071 |
| LVEF** > 50% and ↓ > 15% | 7.4% | 7.0% | NS |

**RR 6 mesi / 12 mesi 0.7531**
**(LC95% 0.5439 – 1.0427)**

**↓58 eventi / 1000 pazienti**

/ finding)

# Cardiac toxicity

| | 12 months (n=1690) | 6 months (n=1690) | P |
|---|---|---|---|
| Cardiac events* | 5.7% | 1.9% | <0.0001 |
| | | | |
| LVEF** < 50% | 6.3% | 4.7% | 0.04 |
| LVEF** < 50% and ↓ > 10% | 4.8% | 3.6% | 0.071 |
| LVEF** > 50% and ↓ > 15% | 7.4% | 7.0% | NS |

**RR 6 mesi / 12 mesi 0.9440**
**(LC95% 0.7408 – 1.2029)**

**↓20 eventi / 1000 pazienti**

V finding)

VIENNA 2012 ESMO

2012.org

# Statistical Methods

- **Non inferiority randomized trial**
  - 2% variation in te[illegible]f absolute difference of recurrence
  - The 95% CI HR r[illegible]d not cross the 1.15 boundary
  - 1040 D[illegible]
  - 4 years
  - HR wer[illegible]
- **Accru[illegible]**

**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**

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

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

**COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP)**

# GUIDELINE ON THE CHOICE OF THE NON-INFERIORITY MARGIN

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

# Statistical Methods

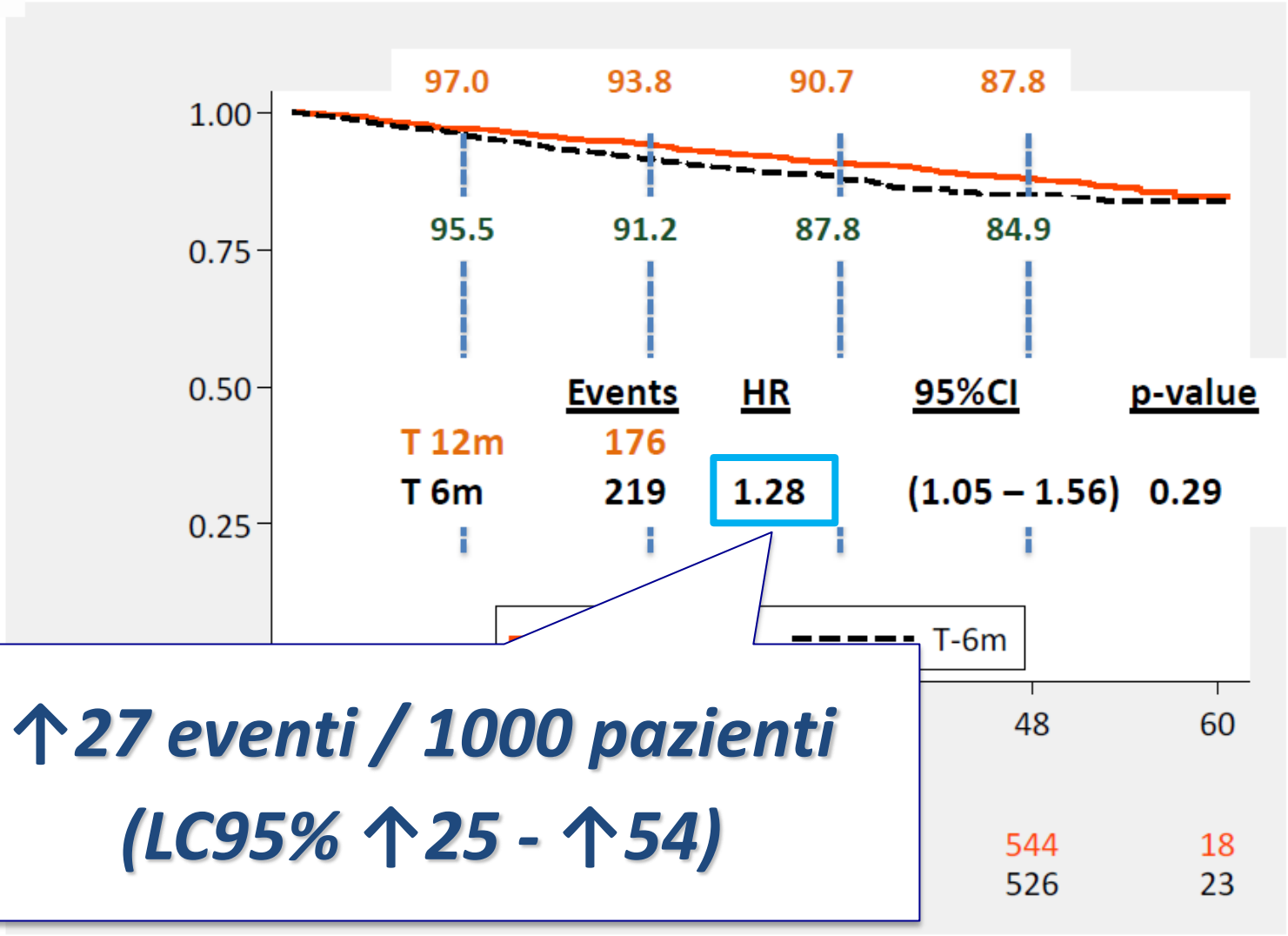
- **Non inferiority randomized trial**
  - 2% variation in terms of absolute difference of recurrence
  - The 95% CI HR margins should not cross the 1.15 boundary
  - 1040 DFS events required for 80% power at 5% level

  **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**

# Disease Free Survival

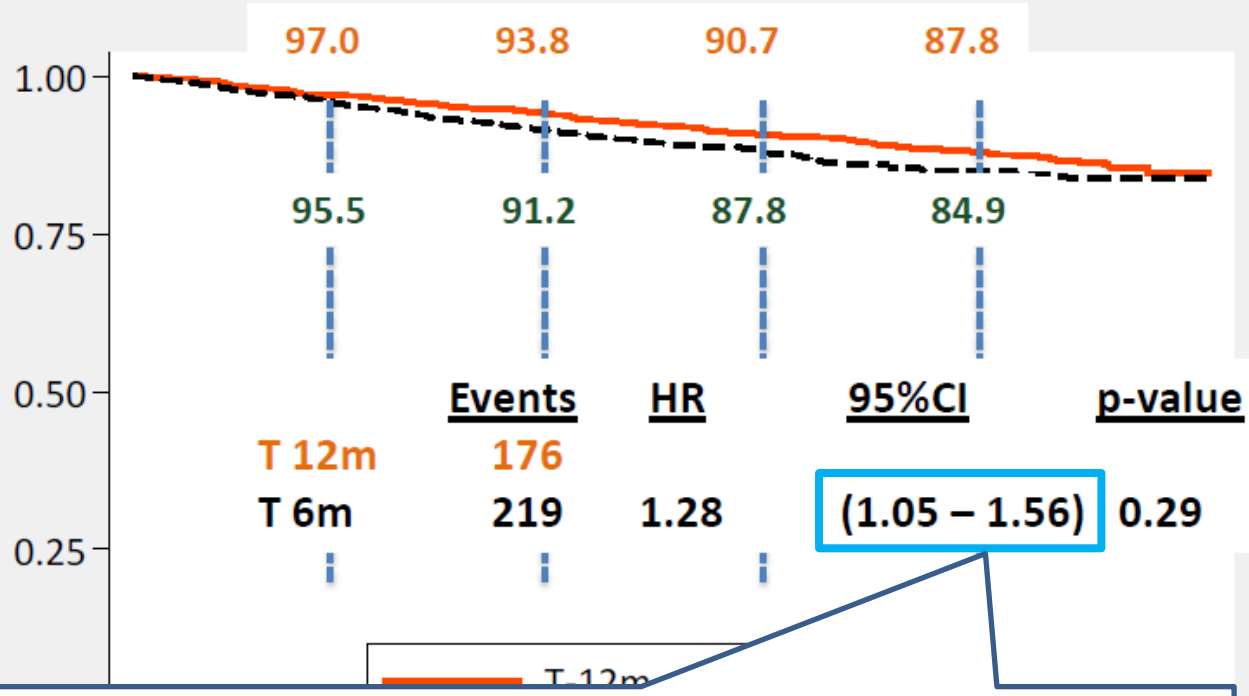



***↑27 eventi / 1000 pazienti***

***(LC95% ↑25 - ↑54)***

* Cox model stratified by ER status and concomitant chemotherapy

# Disease Free Survival




***1.56 > 1.15 = lo studio ha fallito la dimostrazione di non-inferiorità***

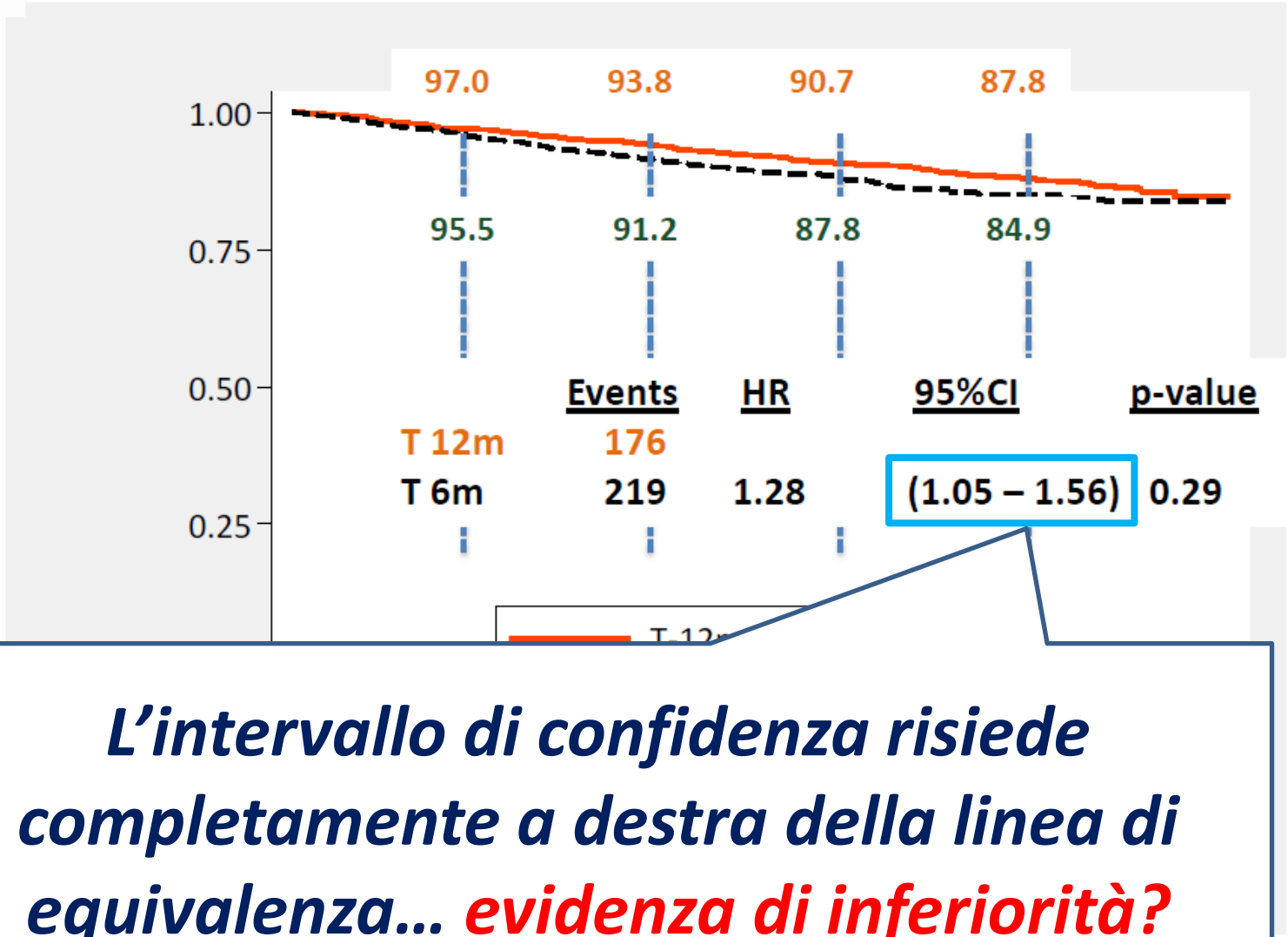
* Cox model stratified by ER status and concomitant chemotherapy

# PHARE CONCLUSIONS

- **Observed HR 1.28 (CI: 1.05-1.56) so inconclusive in terms of noninferiority**

Sandra Swain, ESMO 2012

# Disease Free Survival




***L'intervallo di confidenza risiede completamente a destra della linea di equivalenza… evidenza di inferiorità?***

# PHARE CONCLUSIONS

- **Observed HR 1.28 (CI: 1.05-1.56) so inconclusive in terms of noninferiority**
- **Since Lower CI > 1.0, conclude that 12 mo is better than 6 mo and increase in HR for 6 mo is at least 5%, not sure if less than 15%**

Sandra Swain, ESMO 2012

# 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'.

# Reporting of Noninferiority and Equivalence Randomized Trials

## An Extension of the CONSORT Statement


Gilda Piaggio, PhD
Diana R. Elbourne, PhD
Douglas G. Altman, DSc
Stuart J. Pocock, PhD
Stephen J. W. Evans, MSc
for the CONSORT Group


*JAMA. 2006;295:1152-1160*

# Reporting of Noninferiority and Equivalence Randomized Trials

## An Extension of the CONSORT Statement


| | |
|---|---|
| Gilda Piaggio, PhD | Stuart J. Pocock, PhD |
| Diana R. Elbourne, PhD | Stephen J. W. Evans, MSc |
| Douglas G. Altman, DSc | for the CONSORT Group |


*JAMA. 2006;295:1152-1160*


†This CI is inconclusive in that it is still plausible that the true treatment difference is less than Δ, but the new treatment is significantly worse than the standard.

# PHARE CONCLUSIONS

- Observed HR 1.28 (CI: 1.05-1.56) so inconclusive in terms of noninferiority
- Since Lower CI > 1.0, conclude that 12 mo is better than 6 mo and increase in HR for 6 mo is at least 5%, not sure if less than 15%

**Significantly worse**
**≠ (clinically) inferior**

Sandra Swain, ESMO 2012

# 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'.

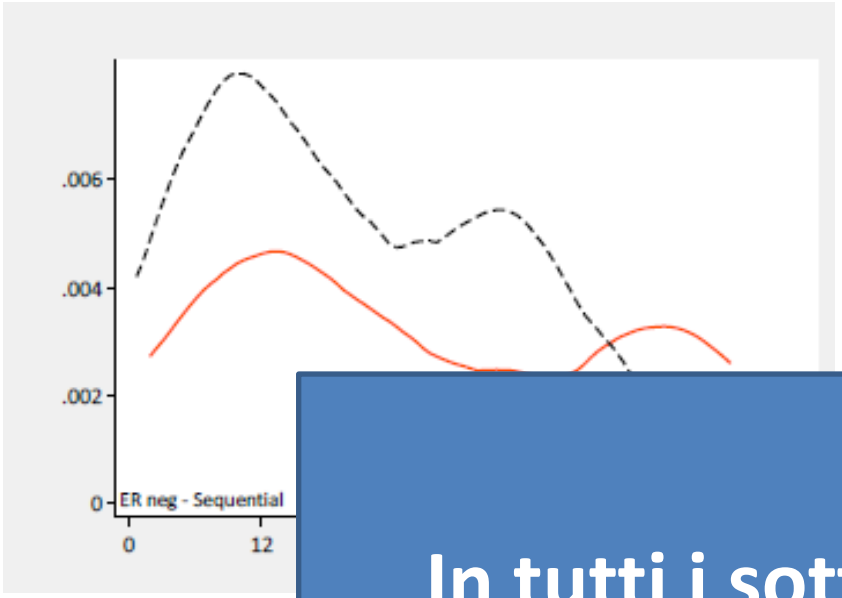
# PHARE CONCLUSIONS

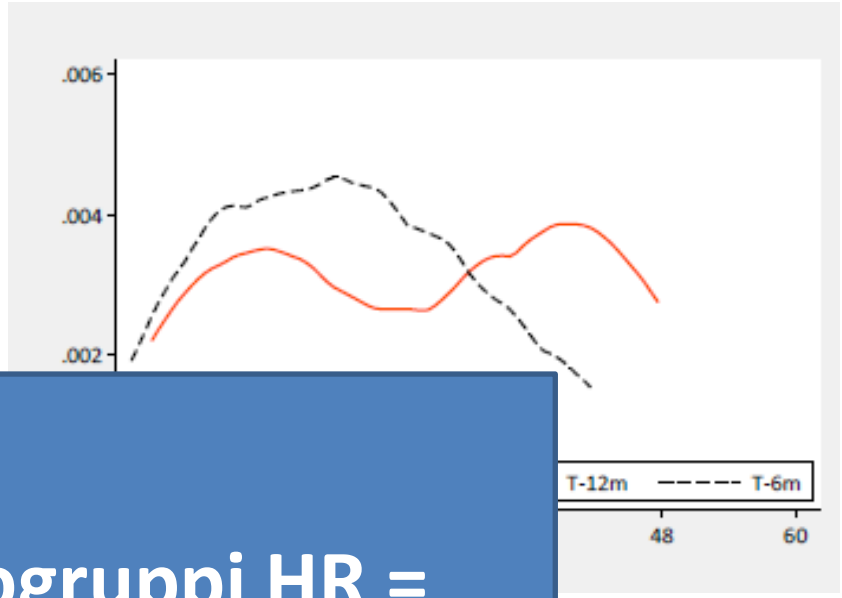
- Observed HR 1.28 (CI: 1.05-1.56) so inconclusive in terms of noninferiority
- Since Lower CI > 1.0, conclude that 12 mo is better than 6 mo and increase in HR for 6 mo is at least 5%, not sure if less than 15%
- 395 events much less than planned 1040 events, so if trial had continued may have been able to statistically show HR with 6 mo greater than 15% with tighter point estimates

Sandra Swain, ESMO 2012

# Objective and strategy of subgroup analysis

- The aim of this analysis was to assess the trastuzumab duration effects for specific subgroups
  - Identified by pre-defined stratification factors
    - Estrogen receptor status
    - Sequential or concomitant administration of trastuzumab with chemotherapy
- Assess heterogeneity by interaction tests
  - 0.10 significance level

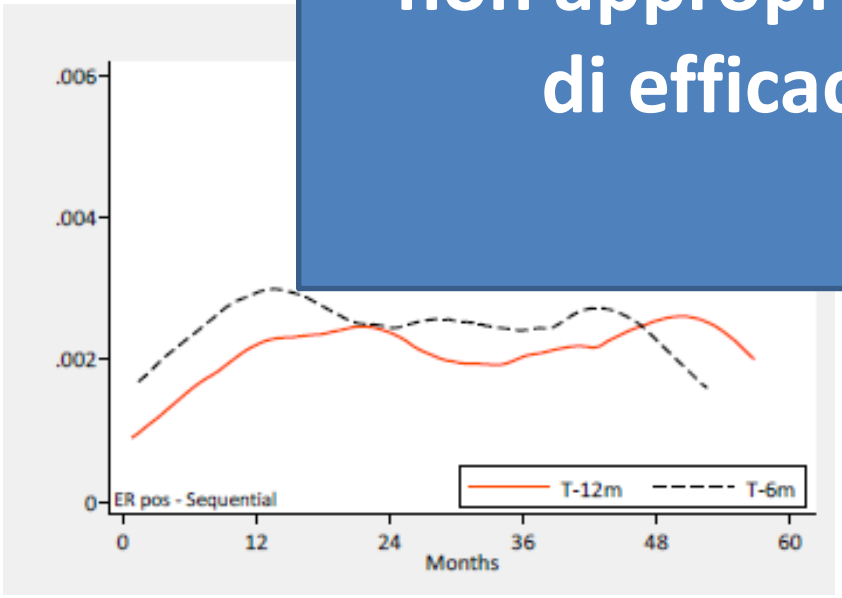


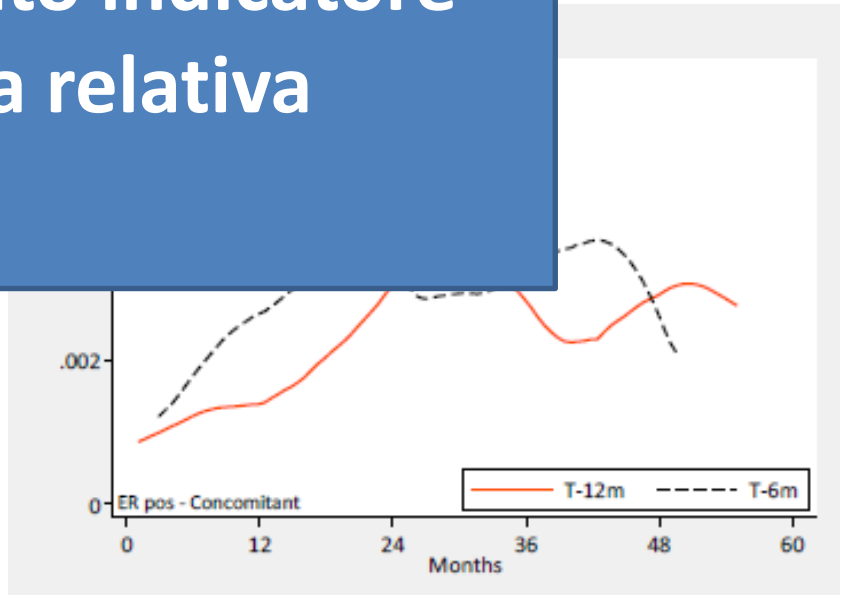

ER

ant

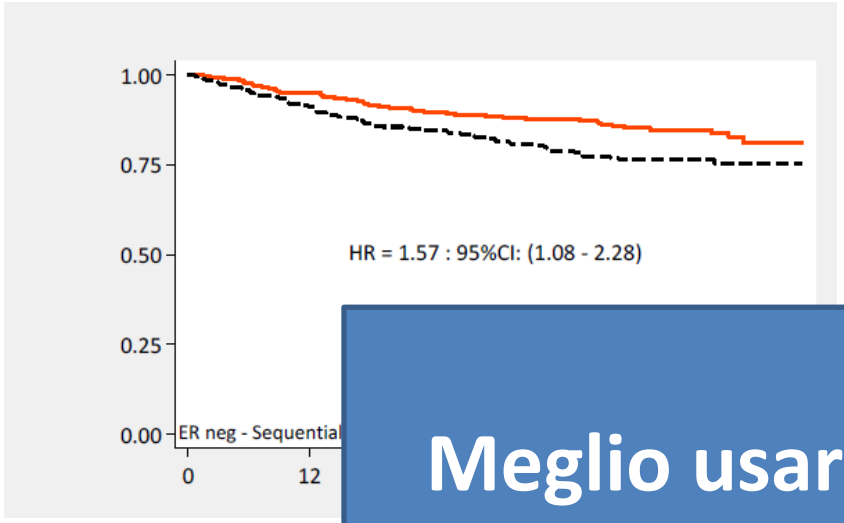
**In tutti i sottogruppi HR = non appropriato indicatore di efficacia relativa**



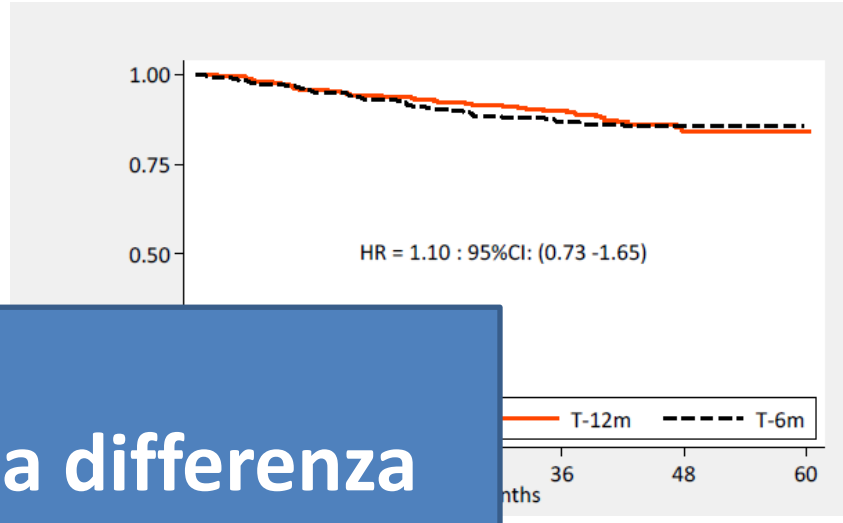

ER(+), sequential




ER(+), concomitant

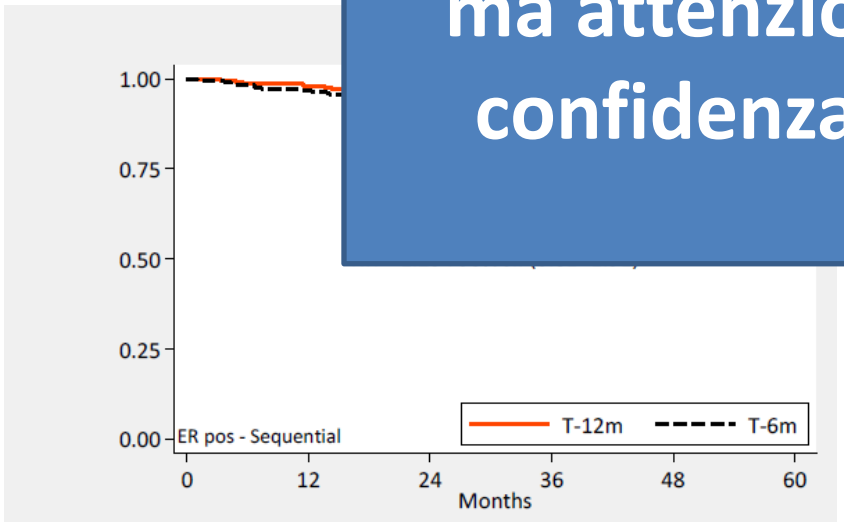



ER(-

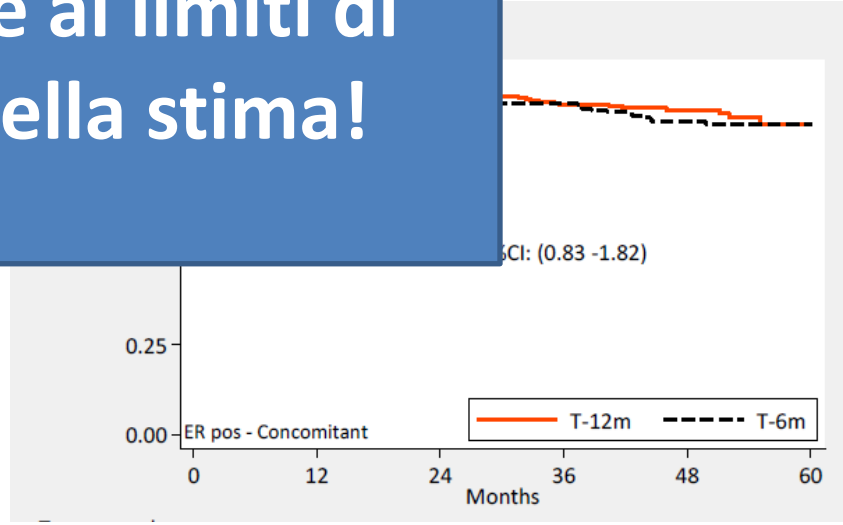



itant

**Meglio usare la differenza assoluta tra le curve… ma attenzione ai limiti di confidenza della stima!**




ER(+), sequential




ER(+), concomitant

# Come valutare la Qualità delle Evidenze disponibili in Letteratura

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

## GRADE: an emerging consensus on rating quality of evidence and strength of recommendations

The members of the Grade Working Group

BMJ | 26 APRIL 2008 | VOLUME 336 924

Graduazione della qualità delle prove.

| Livello qualità | Significato | Conseguenza |
|---|---|---|
| Alta | Alto grado di confidenza nei risultati | È molto improbabile che ulteriori studi possano cambiare la fiducia nella stima di effetto |
| Moderata | Discreto grado di confidenza nei risultati | È probabile che ulteriori studi possano confermare o cambiare la fiducia nella stima di effetto |
| Bassa | I risultati sono poco credibili | È necessaria ulteriore ricerca per ottenere stime affidabili sugli effetti positivi e negativi dell'intervento |
| Molto bassa | I dati esaminati sono totalmente inaffidabili | Non è possibile fare affidamento sulle stime di effetto disponibili |

# Come valutare la Qualità delle Evidenze disponibili in Letteratura

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

## GRADE: an emerging consensus on rating quality of evidence and strength of recommendations

The members of the Grade Working Group

BMJ | 26 APRIL 2008 | VOLUME 336 924

**Study design**

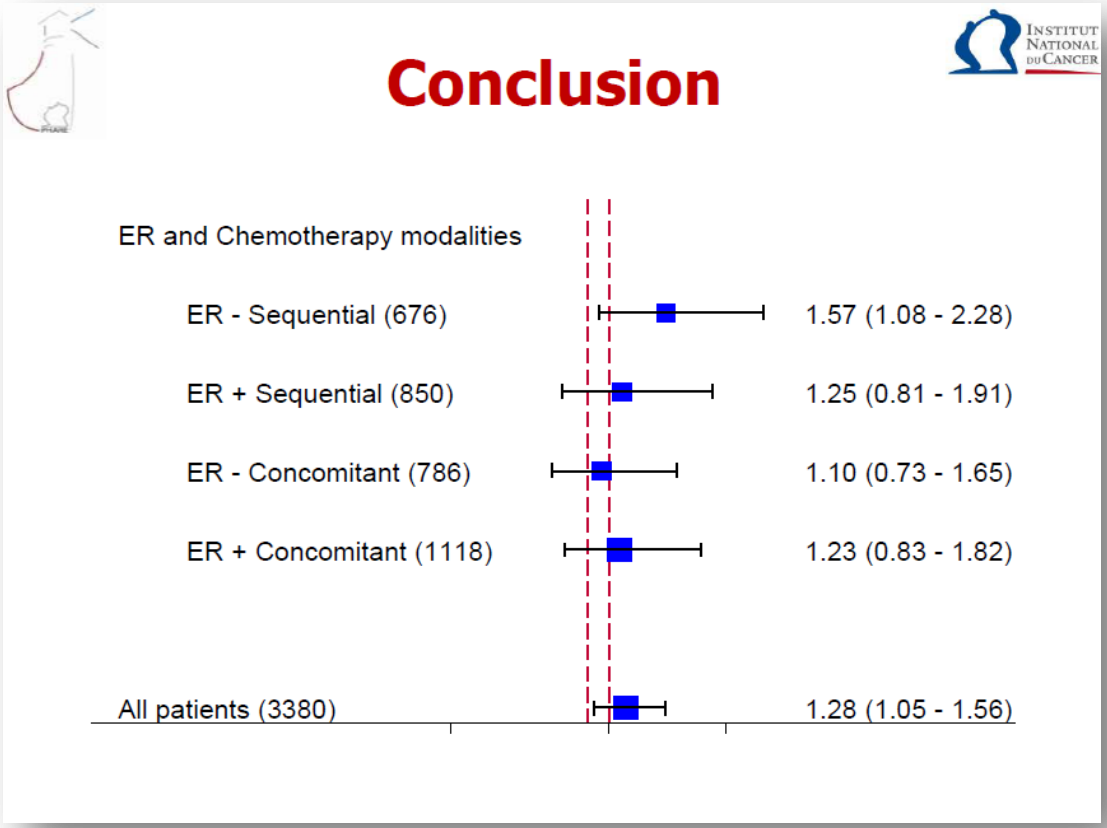
RCTs ⊕⊕⊕⊕

observational studies ⊕⊕◯◯

## Determinants of quality

5 factors that can lower quality

1. limitations of detailed design and execution *(risk of bias criteria)*
2. Inconsistency *(or heterogeneity)*
3. Indirectness *(PICO and applicability)*
4. Imprecision *(number of events and confidence intervals)*
5. Publication bias




**DFS endpoint: downgrade (-3) HIGH → VERY LOW**

- 395 events observed / 1040 planned
- non proportional hazards (subgroup analysis)
- wide confidence limits (subgroup analysis)