Workshop on Single Arm Trials in Oncology Products
EMA, London, June 30, 2016

How to improve the reliability of Single Arm Trials

Paolo Bruzzi
Clinical Epidemiology Unit
IRCCS San Martino-IST - National Cancer Research Institute
Genoa - ITALY

Focus of the presentation

	Not discussed
Experimental studies	*Observational studies*
aimed at evaluating the efficacy (phase III)	*Phase I and II, dose finding, proof of principle*
of cancer drugs	*Other treatments, other diseases*
without a randomized control group	*Classical RCT's, (Cluster trials? Crossover trials?)*

Single Arm Trials

a) Without a control group

b) With a (or >1) control group

- Historical
- Concurrent

Single Arm Trials

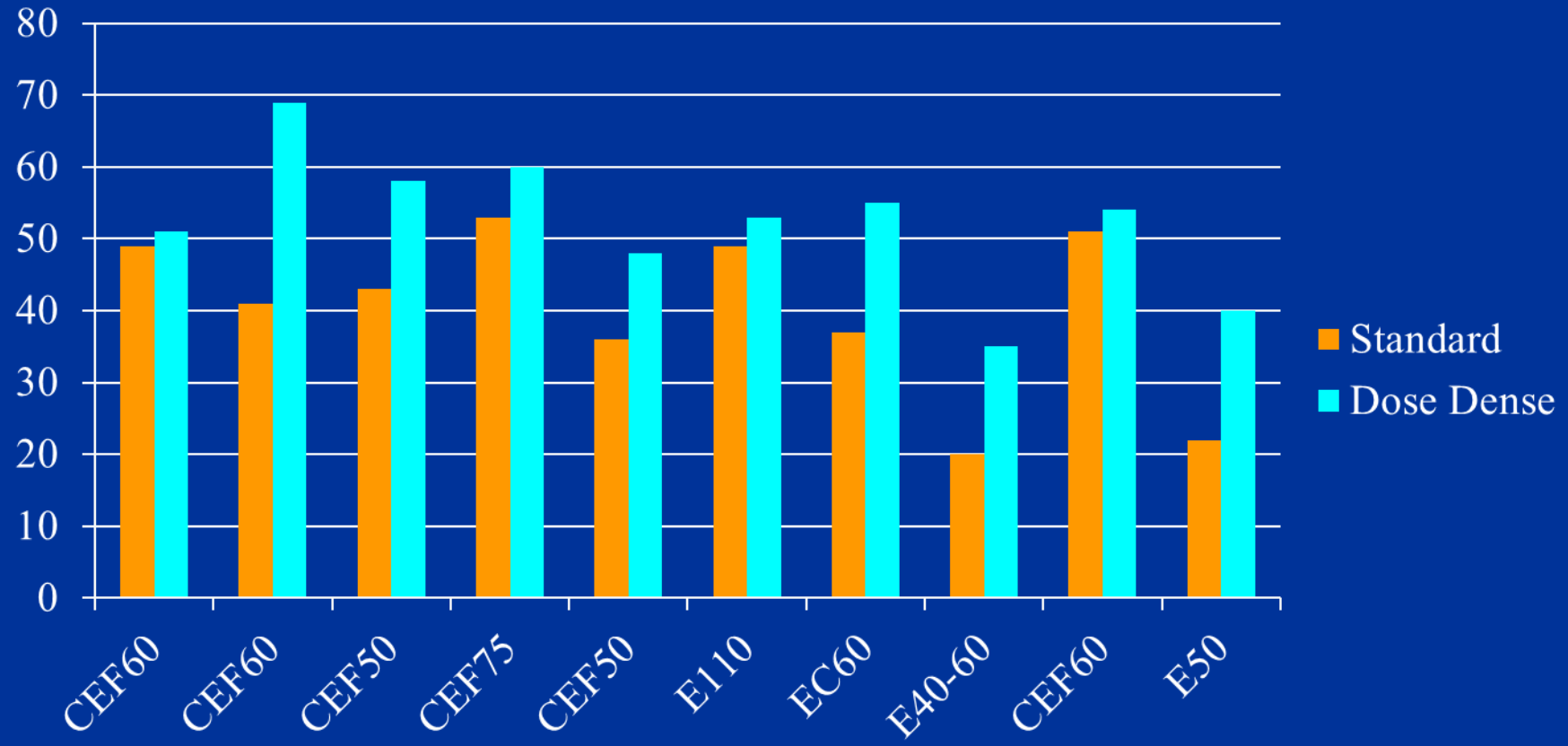
a) Without a control group

Absolute benefit = “success” rate

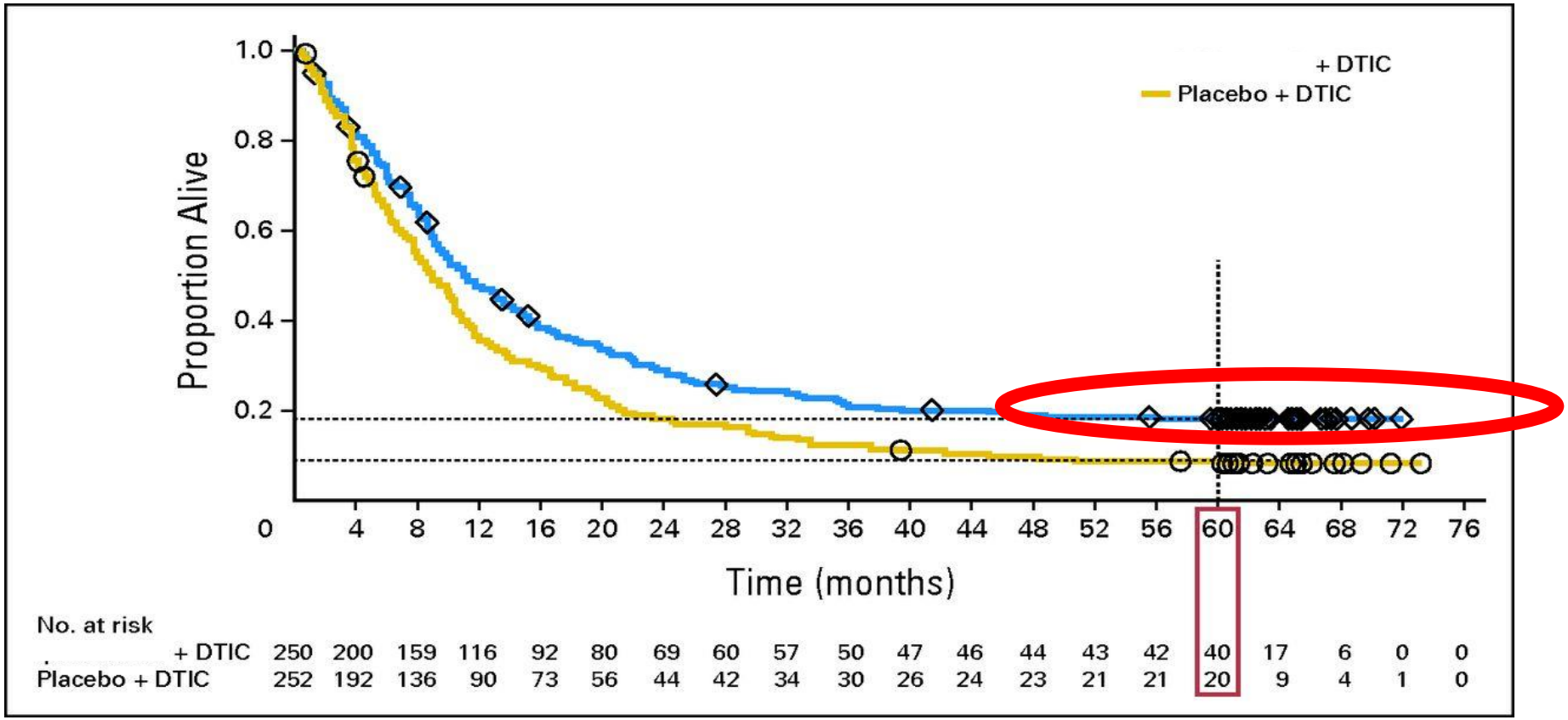
(e.g. % of Responses, of long-term survivors, etc.)

- Breakthrough drugs (e.g. Gleevec in GIST)
- Otherwise unreliable (large variations)

Response Rate in RCT's of standard & dose dense Anthracyclines in MBC



Kaplan-Meier estimates of overall survival in patients treated with ipilimumab plus dacarbazine (DTIC) or placebo plus DTIC in phase III CA184-024 study.



Michele Maio et al. JCO doi:10.1200/JCO.2014.56.6018

Metastatic Melanoma, DTIC/Plat./IF +/- IL2 - 2005 (Keilholz EU)

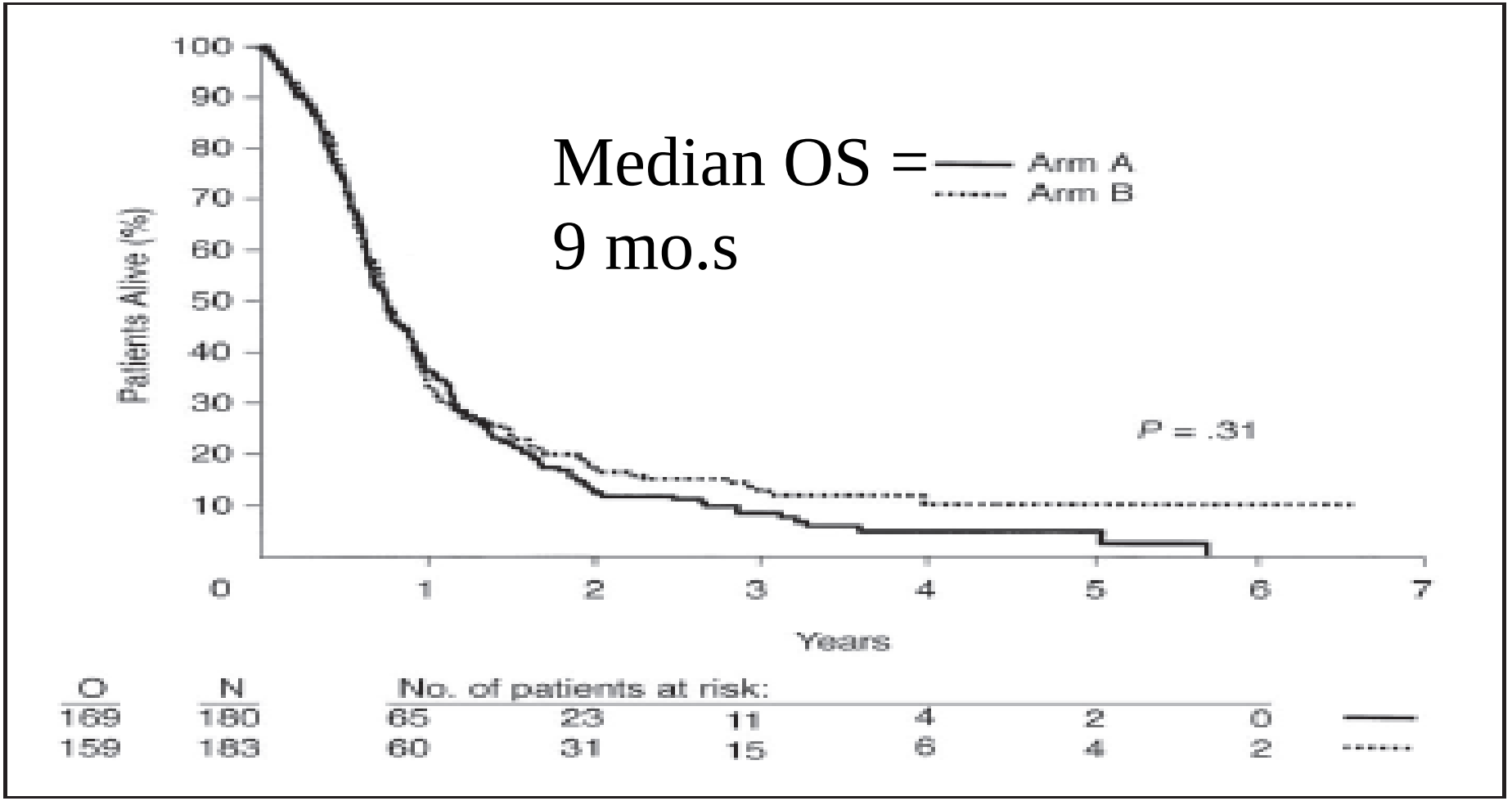


Fig 1. Overall survival according to treatment arm. O, number of observations; N, number of patients.

RCT's in Metastatic Melanoma (control group)

Year	Author - Treatment	Median OS (months)	HR
2001	DTIC, Young	4	1
2004	DTIC, Avril	5.6	0.71
2002	*PoliCT, Ridolfi*	*9.5*	*0.42*
2005	PoliCT, Keilholz	9	0.44
2011	DTIC, Robert	9.1	0.44

Stacchiotti JCO 2012
Phase II of Imatinib in Chordoma

- Uncontrolled, 14 centers
- 56 patients in 18 mos
- PD > 30 events (Median PFS = 9 months)
- OS ≈ 25 events (Median OS = 35 months)
- Response Rate
  - CR: 0
  - PR: 1/50 (2%),
  - SD: 35 (70%)

Stacchiotti JCO 2012
Phase II of Imatinib in Chordoma

*Conclusions: "…confirms that imatinib has some antitumor activity in chordoma. … The lack of RECIST responses and the potentially slow natural course of the disease … **do not allow us to affirm that this treatment is effective**"*

How to improve the reliability of Single Arm Trials

1. A control group should be available for comparison (and considered in study design)

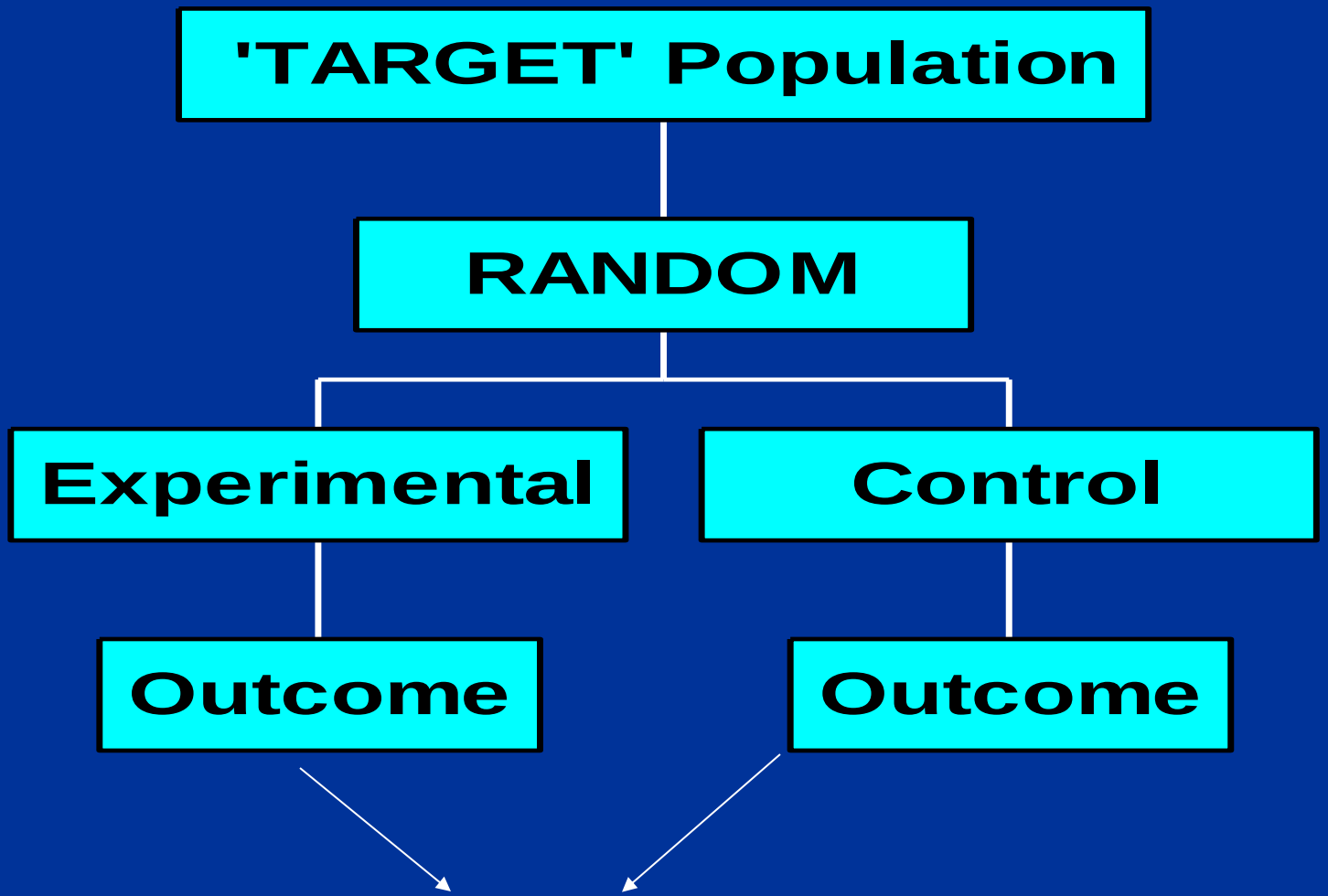
Single Arm trials

a) Without a control group

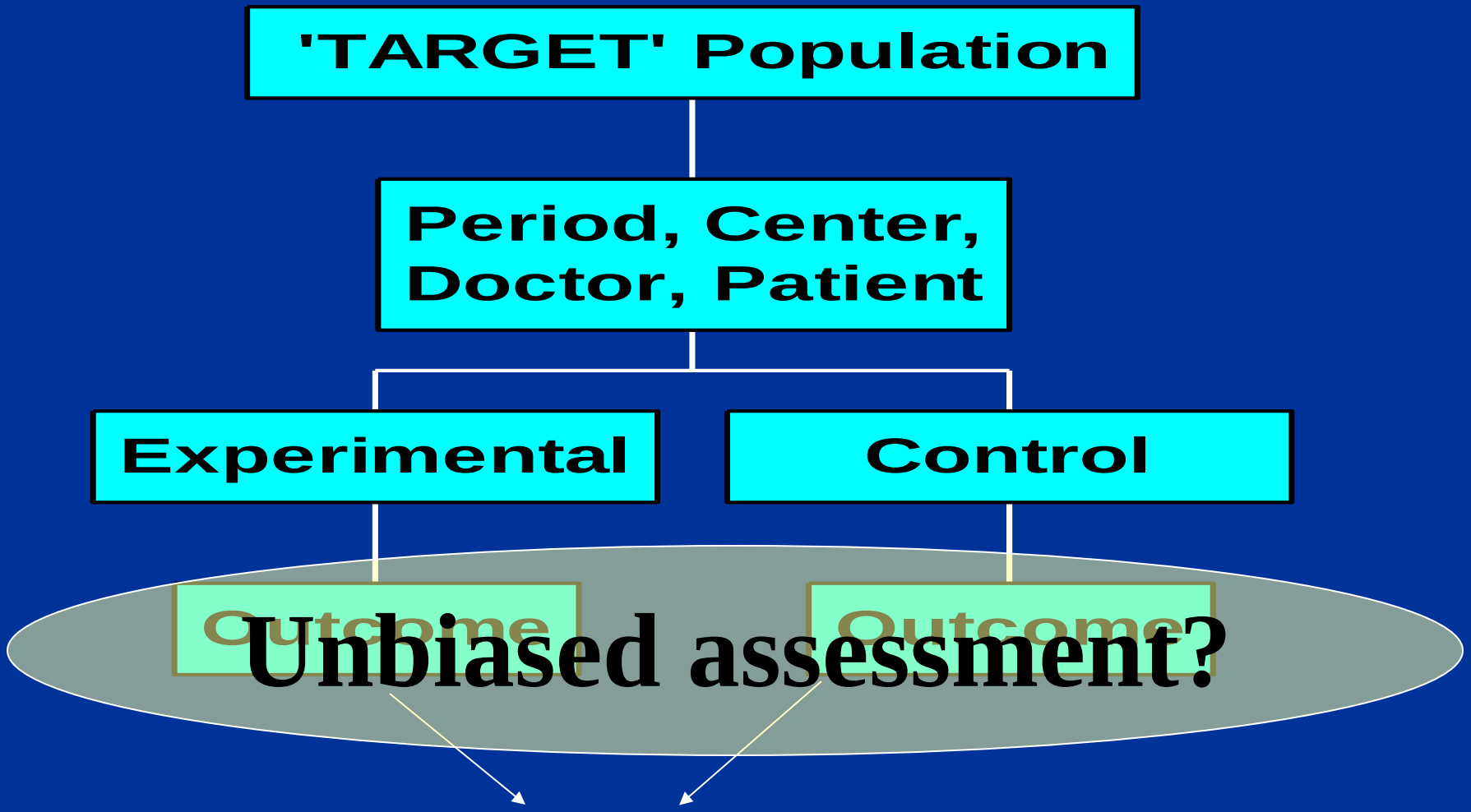
b) With a control group

- Historical
- Concurrent

Randomized Clinical Trial



Single Arm Trial



Unbiased assessment of outcome

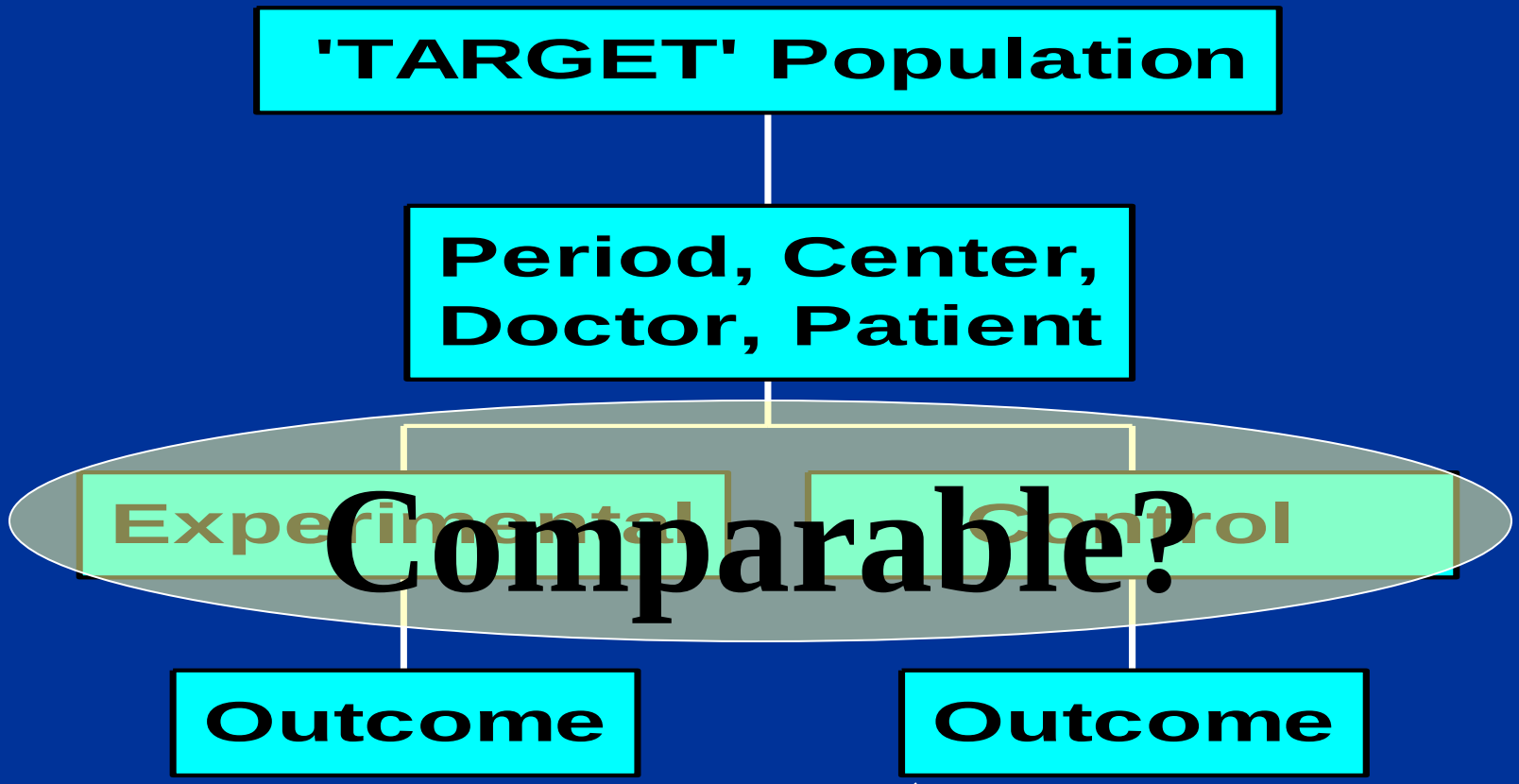
- **Hard Endpoint = Overall Survival**
- Soft endpoint (RR, PFS, RFS, etc.)
  - Valid Surrogate?
  - Unbiased (or comparable) assessment in different periods, centers, etc.?

Blinded review of endpoints (feasible and effective?)

How to improve the reliability of Single Arm Trials

1. A control group should be available for comparison (and considered in study design)
2. The study endpoint should be OS. Otherwise a bias must be always suspected

Single Arm Trial

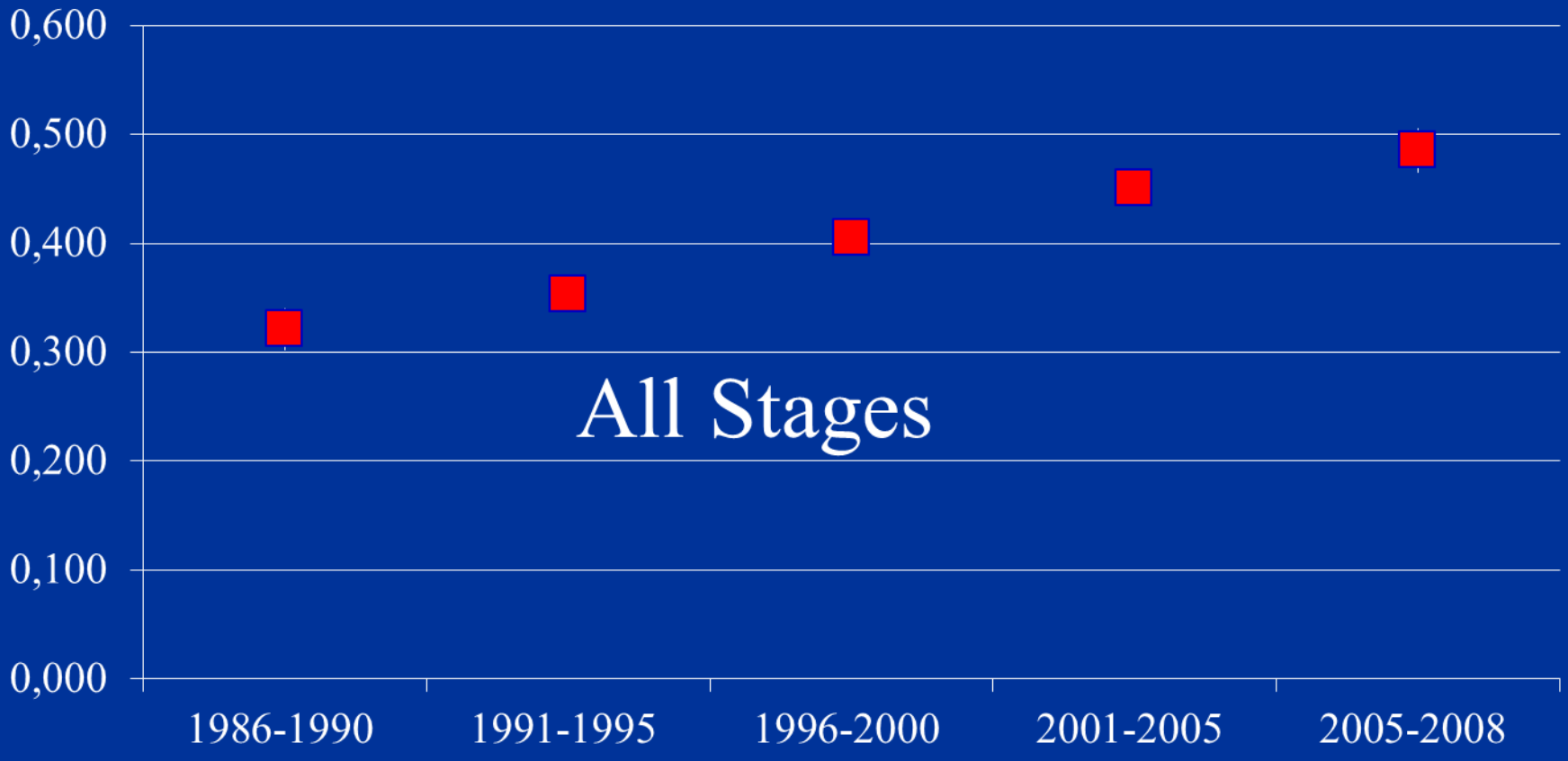


Problems with historical controls

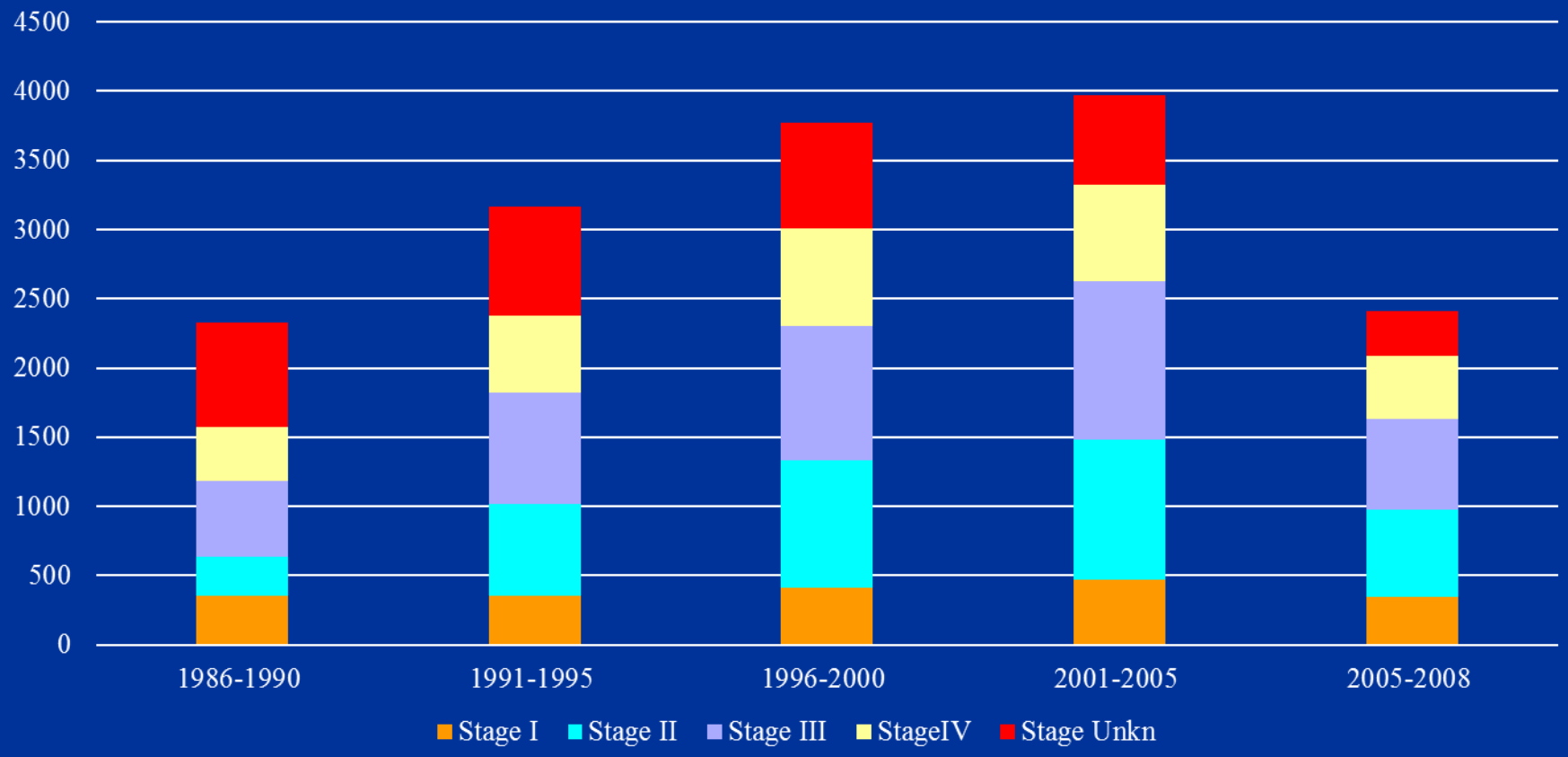
- Biological Variability (e.g. endomet. C., breast c.)
- Selection Criteria (Clinical Center, Period
- Disease Stage at diagnosis (earlier diagnosis, technologies)
- **Stage Migration (Will Rogers' phenomenon)**

Statistical adjustements ineffective

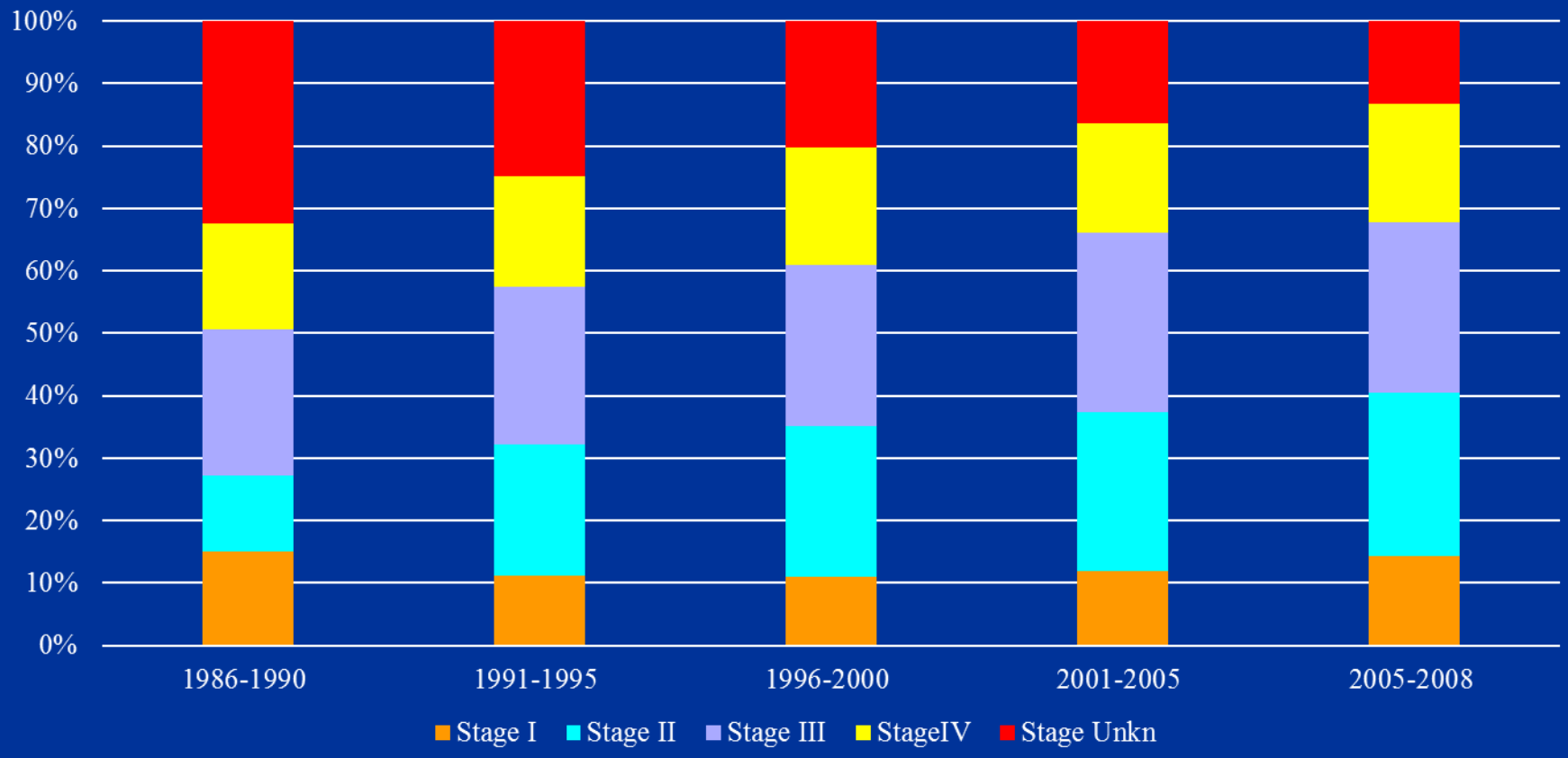
Genoa Cancer Registry
Colorectal cancer - 5 yrs Survival



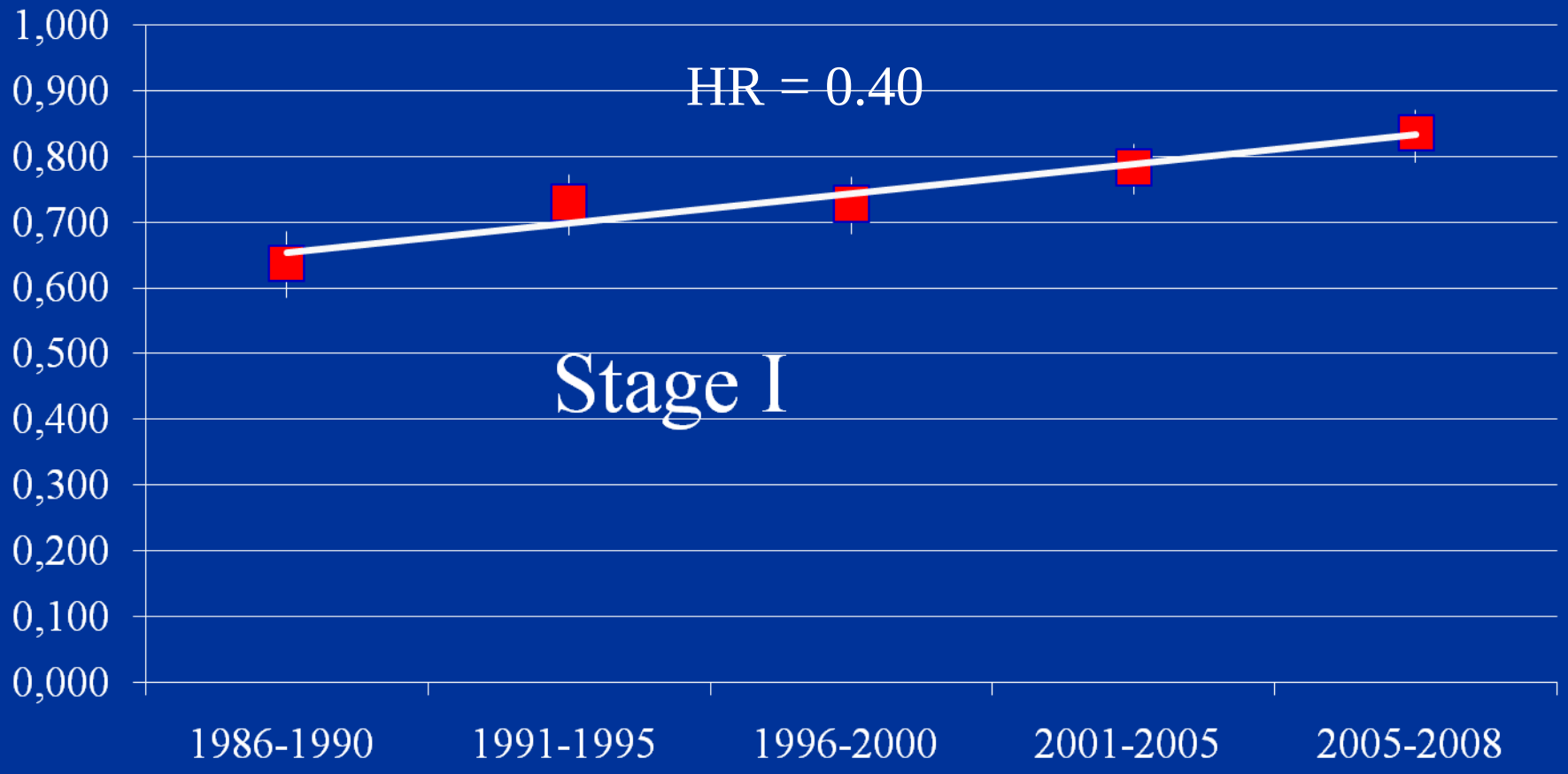
Genoa Cancer Registry – Cases of Incident Colon Cancer



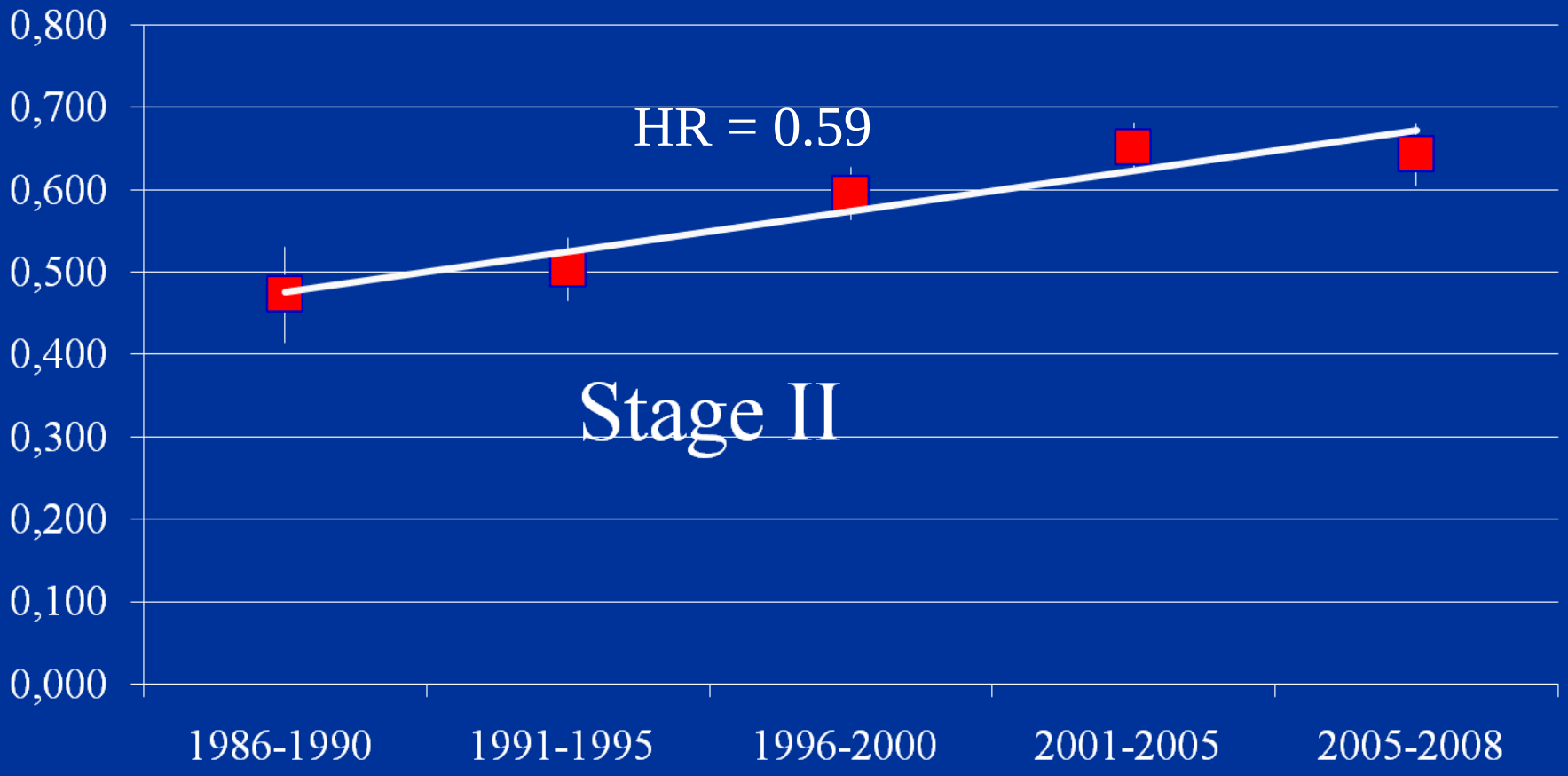
Genoa Cancer Registry – Cases of Incident Colon Cancer



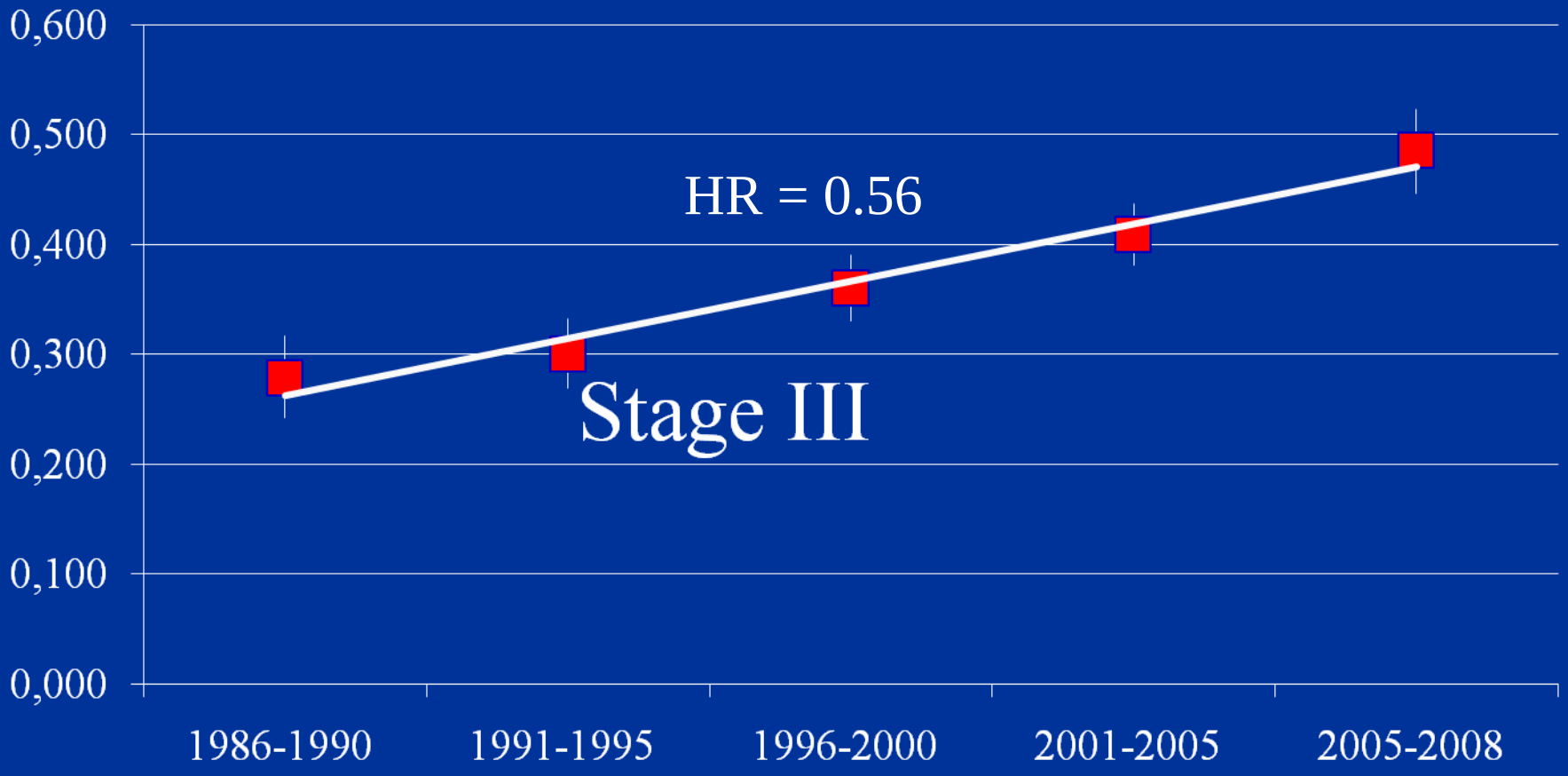
Genoa Cancer Registry
Colorectal cancer - 5 yrs Survival



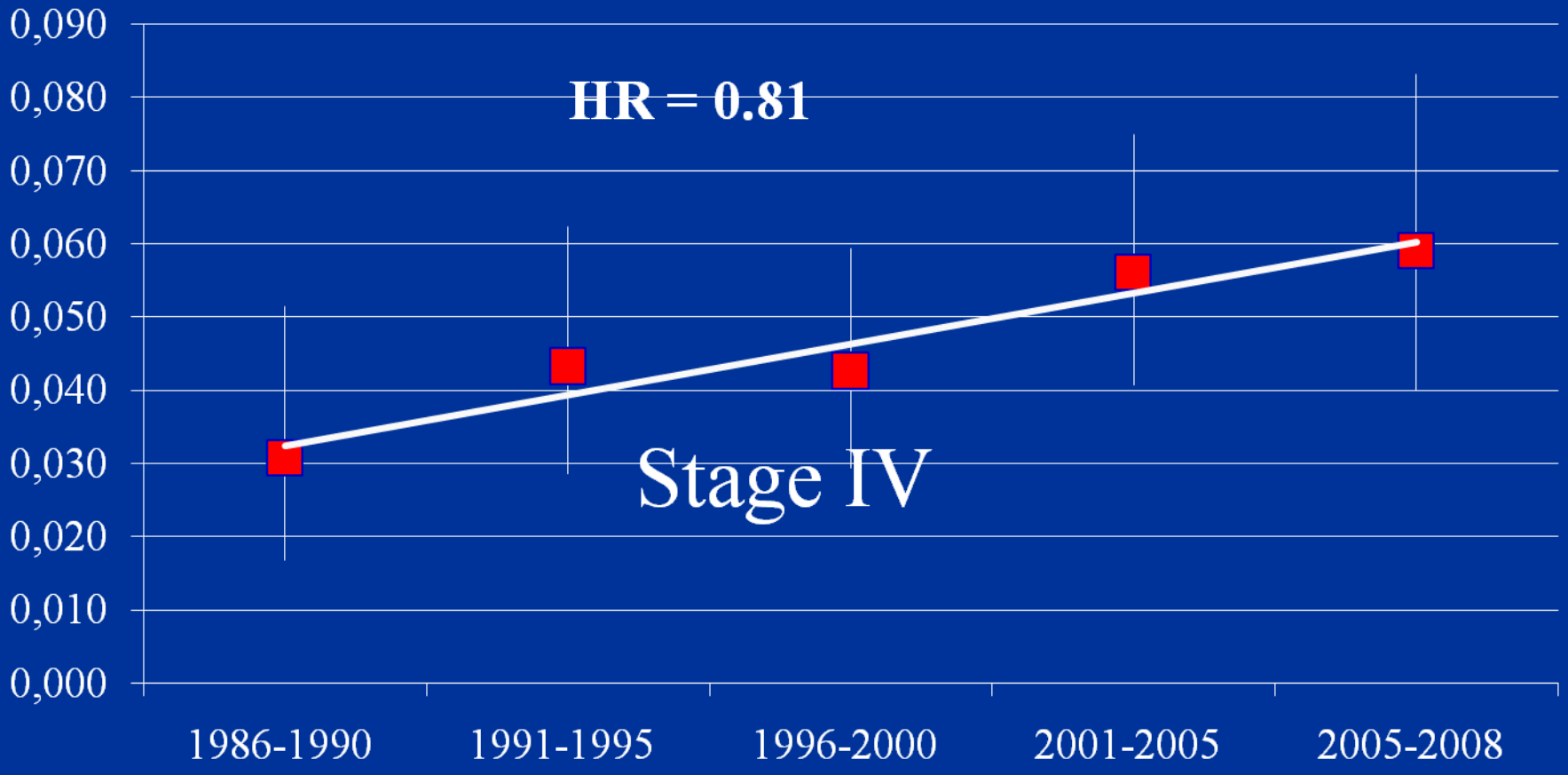
Genoa Cancer Registry
Colorectal cancer - 5 yrs Survival



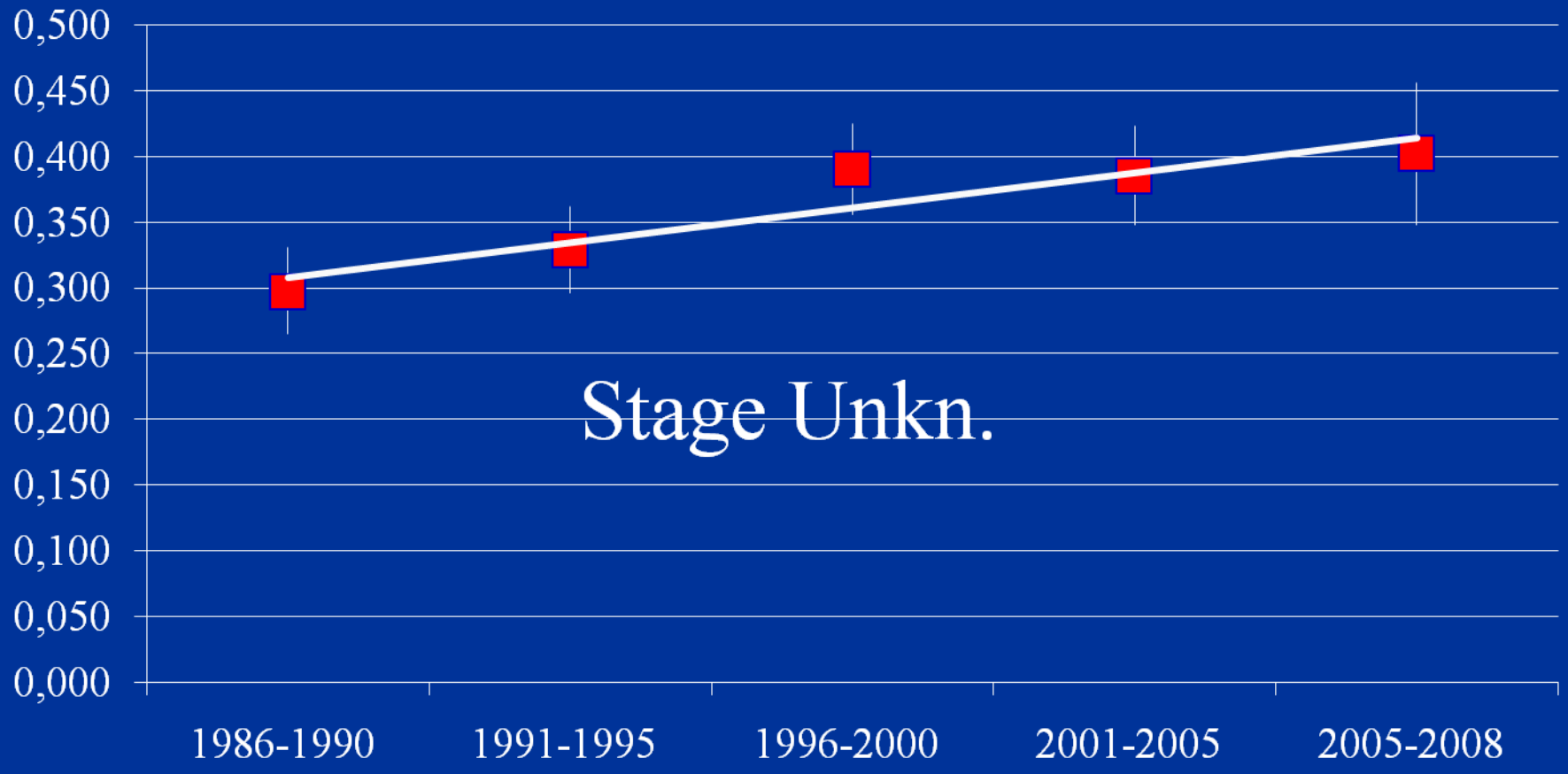
Genoa Cancer Registry
Colorectal cancer - 5 yrs Survival



Genoa Cancer Registry
Colorectal cancer - 5 yrs Survival



Genoa Cancer Registry
Colorectal cancer - 5 yrs Survival

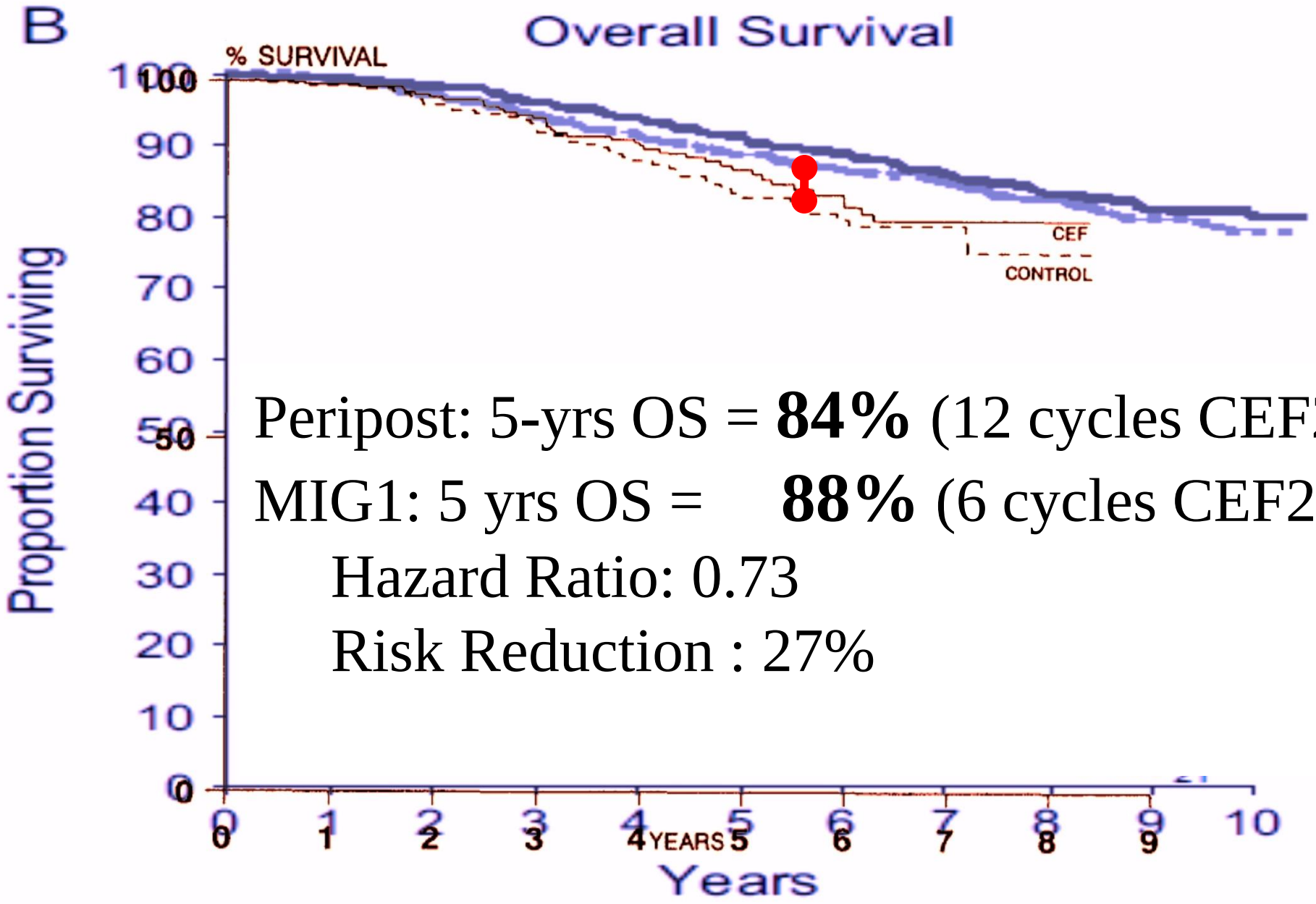


5-years-survival in control groups from two consecutive RCT's in early BC

- Peripost (85-92): Control group: 12 cycles - CEF21
- MIG1 (92-95): Control group: 6 Cycles CEF21

Same therapy, same selection criteria, largely same centers

Sertoli, JCO 1995 - Venturini, JNCI 2005



Peripost: 5-yrs OS = **84%** (12 cycles CEF21)

MIG1: 5 yrs OS = **88%** (6 cycles CEF21)

Hazard Ratio: 0.73

Risk Reduction : 27%

Same therapy, same selection criteria, largely same centers

How to improve the reliability of Single Arm Trials

1. A control group should be available for comparison (and considered in study design)
2. The study endpoint should be OS. Otherwise a bias must be always suspected
3. Historical trends in survival must be accounted for (concurrent controls?)

Problems with Concurrent Controls

- Where? (Same centers/Other centers?)
- How? (Diagnostic (molecular)/staging procedures? Informed consent?)
- *(Sample Size?)*

a) Natural (cluster) experiments

b) Registries?

a) Natural (cluster) experiments

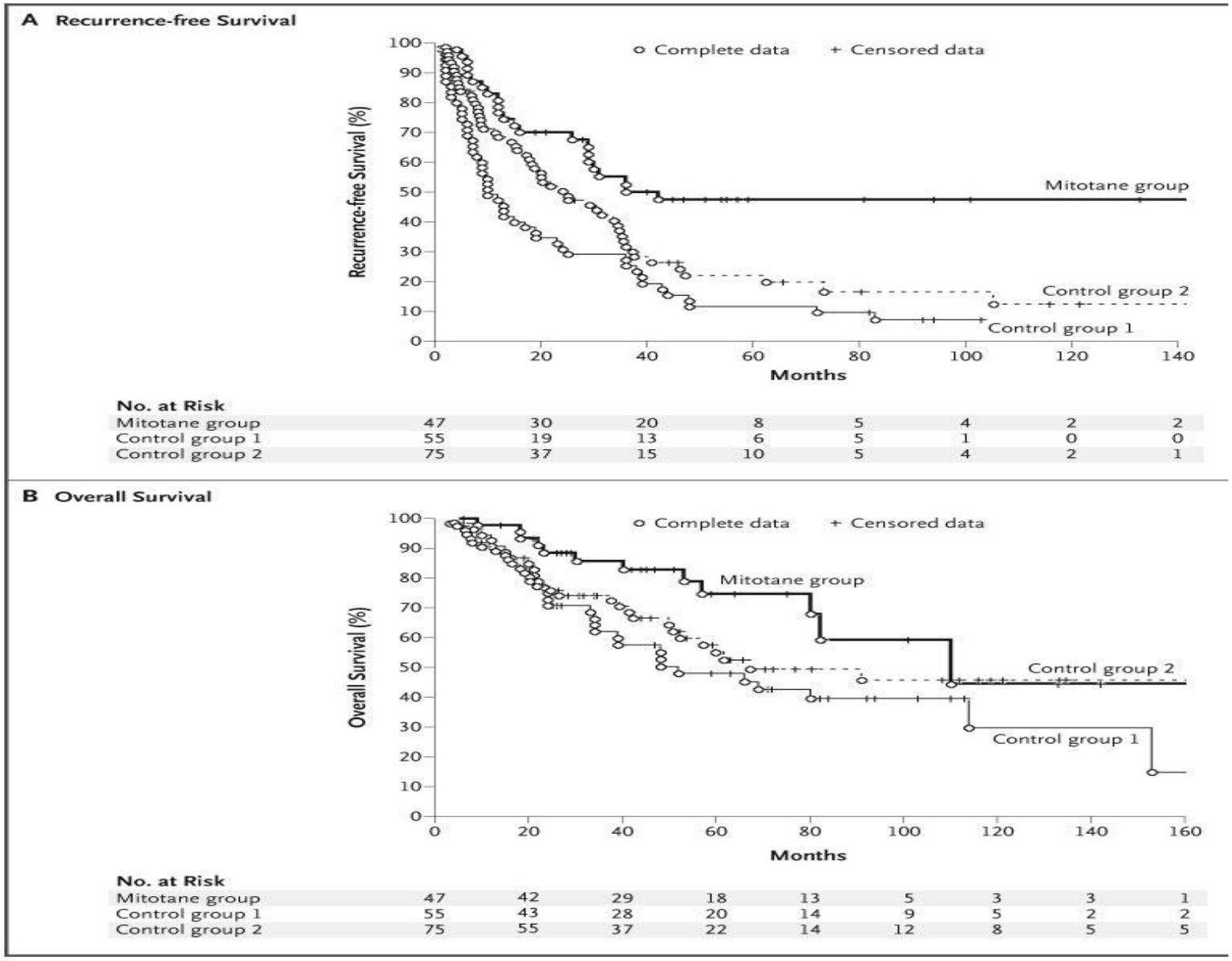
Adrenocortical carcinoma, rare tumor, grim prognosis

Adjuvant therapy after resection? Efficacy Unknown

Eight tertiary referral centers for ACC in Italy, 4 routinely used adjuvant Mitotane, 4 never used it (Control group 1)

Control group 2 : All cases included in the German ACC Registry (adjuvant therapy never used)

Adjuvant Mitotane in adrenocortical Carc.



Mitotane: All resected pts receiving. Adjuv. Mitotane

Control 1: Pts from Italian Centers not using adjuv. Mitotane

Control 2: Pts from German centers (Mitot. never used)
(Terzolo M, NEJM 2007)

b) Registries

1. **Population based** (Cancer Registries)

Pro's

- Unselected, representative series of patients
- Historical trends assessable

Con's

- Poorly diagnosed and staged
- Molecular Classification
- Heterogeneous therapies

Registries

2. **Institution based** (Hospital Registries)

Pro's

- Thoroughly diagnosed and staged
- Molecular Classification
- All outcomes assessed

Con's

- Selected series of patients (referral patterns?)
- Historical trends not assessable

Registries

3. **Disease based** (one disease, many institutions)

Pro's

- Thoroughly diagnosed and staged
- Molecular Classification
- All outcomes assessed
- Historical trends assessable

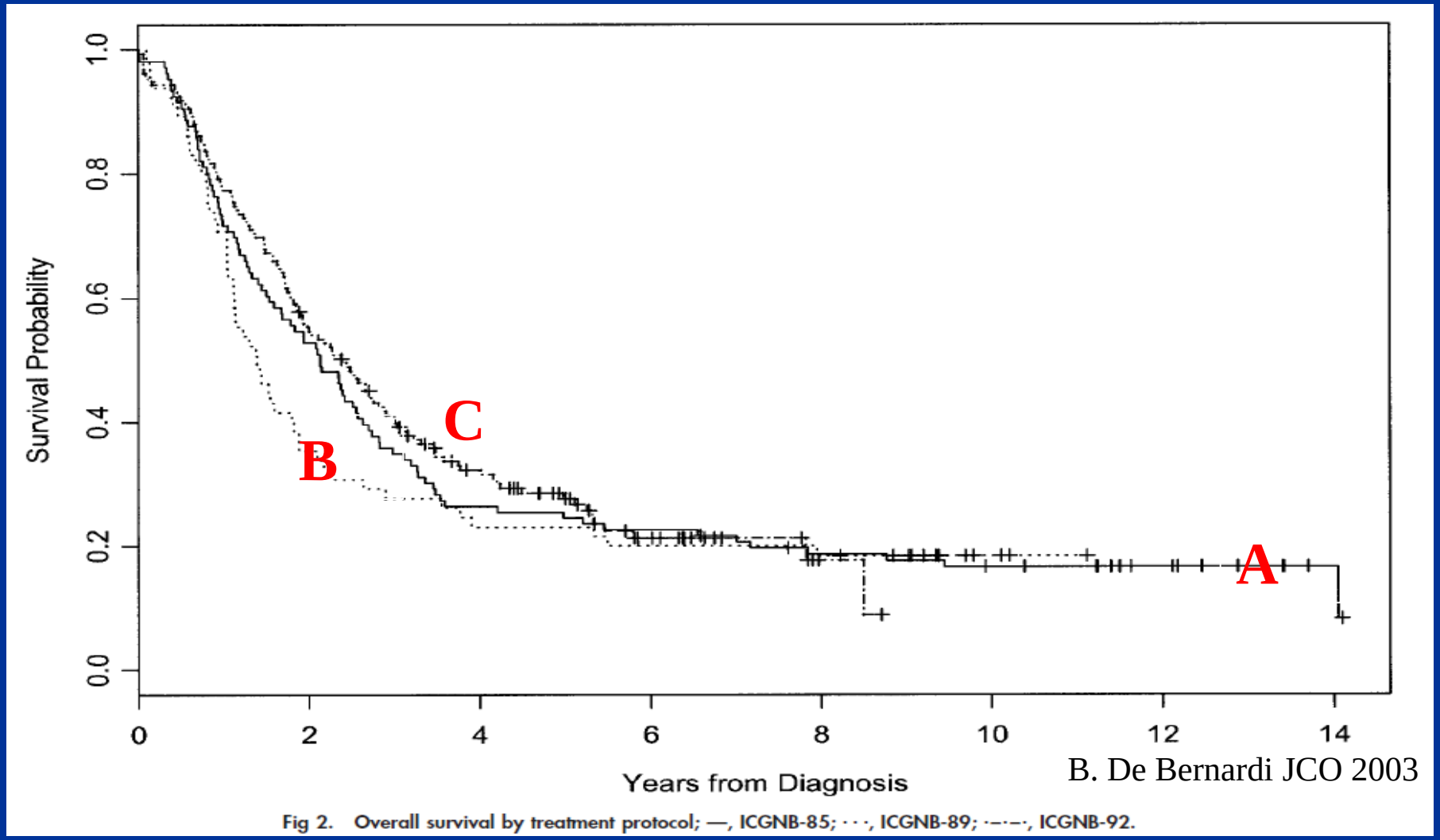
Con's

- Selection bias ? (Population coverage)

Italian Neuroblastoma Registry

- Population based (Italy)
- Acceptable Coverage (>80%)
- Centralised pathology & molecular biology
- Long term follow-up data
- Homogeneous treatment protocols
- Historical trends evaluable

1985-97: Three Consecutive protocols of the Italian Neuroblastoma Cooperative Group
A -> B: Intensification; B -> C: more drugs



B. De Bernardi JCO 2003

Fig 2. Overall survival by treatment protocol; —, ICGNB-85; · · ·, ICGNB-89; ·–·–·, ICGNB-92.

Conclusion: The therapeutic modifications adopted in the ICGNB-89 and ICGNB-92 protocols **were not associated with a significant improvement in response rate or in the 5-year OS** and EFS as compared with the ICGNB-85 protocol.

Attempts at intensifying chemotherapy were associated with greater toxicity

How to improve the reliability of Single Arm Trials

1. A control group should be available for comparison (and considered in study design)
2. The study endpoint should be OS. Otherwise a bias must be always suspected
3. Historical trends in survival must be accounted for (concurrent controls?)
4. Prospectively planned comparisons with appropriate registry data may be very useful

Study Designs of single arm trials

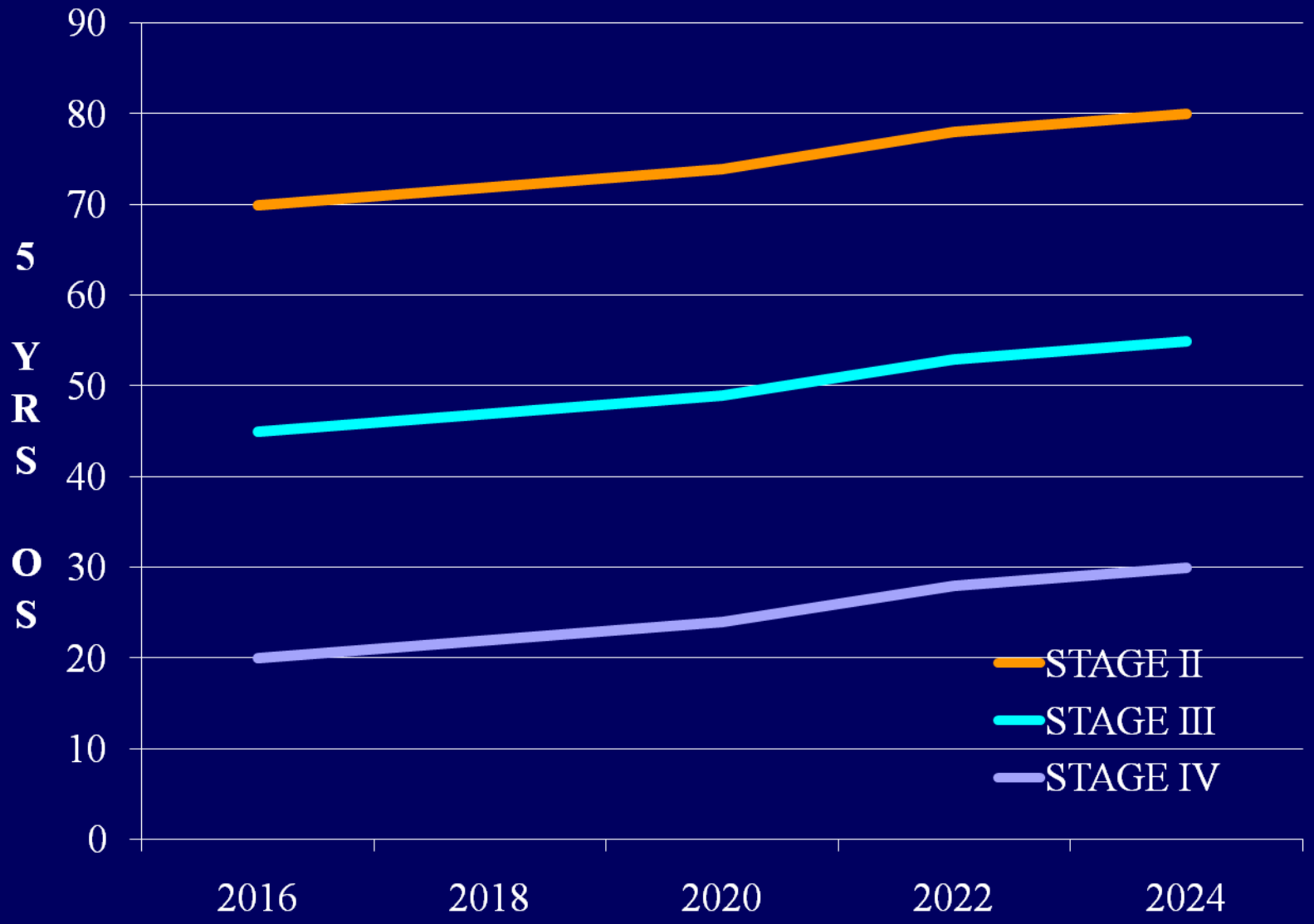
- True prospective cohort studies, where the comparison is between centers participating/not participating to the trial of the new drug

- **Single arm trials nested into population or disease registries**

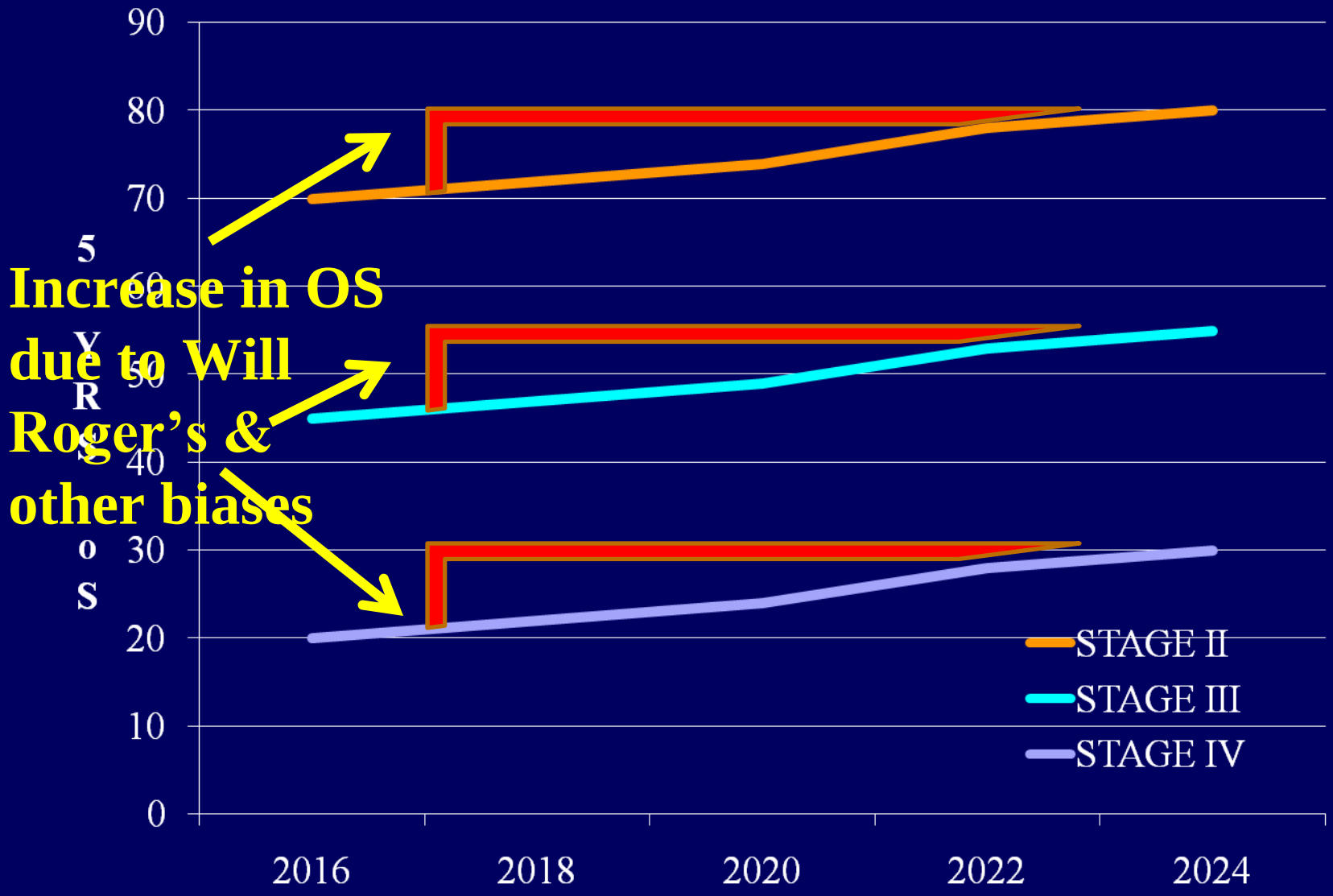
Study design to assess the efficacy of a new drug with an external control group

- The drug is introduced into a group of patients selected from a registry of incident cases
- (A control group of patients with similar characteristics is identified in the same registry)
- Historical trends in prognosis are estimated
- A sudden change in prognosis following the introduction of the new drug is expected **only among eligible patients**

OS by stage (mutation, etc.) over time

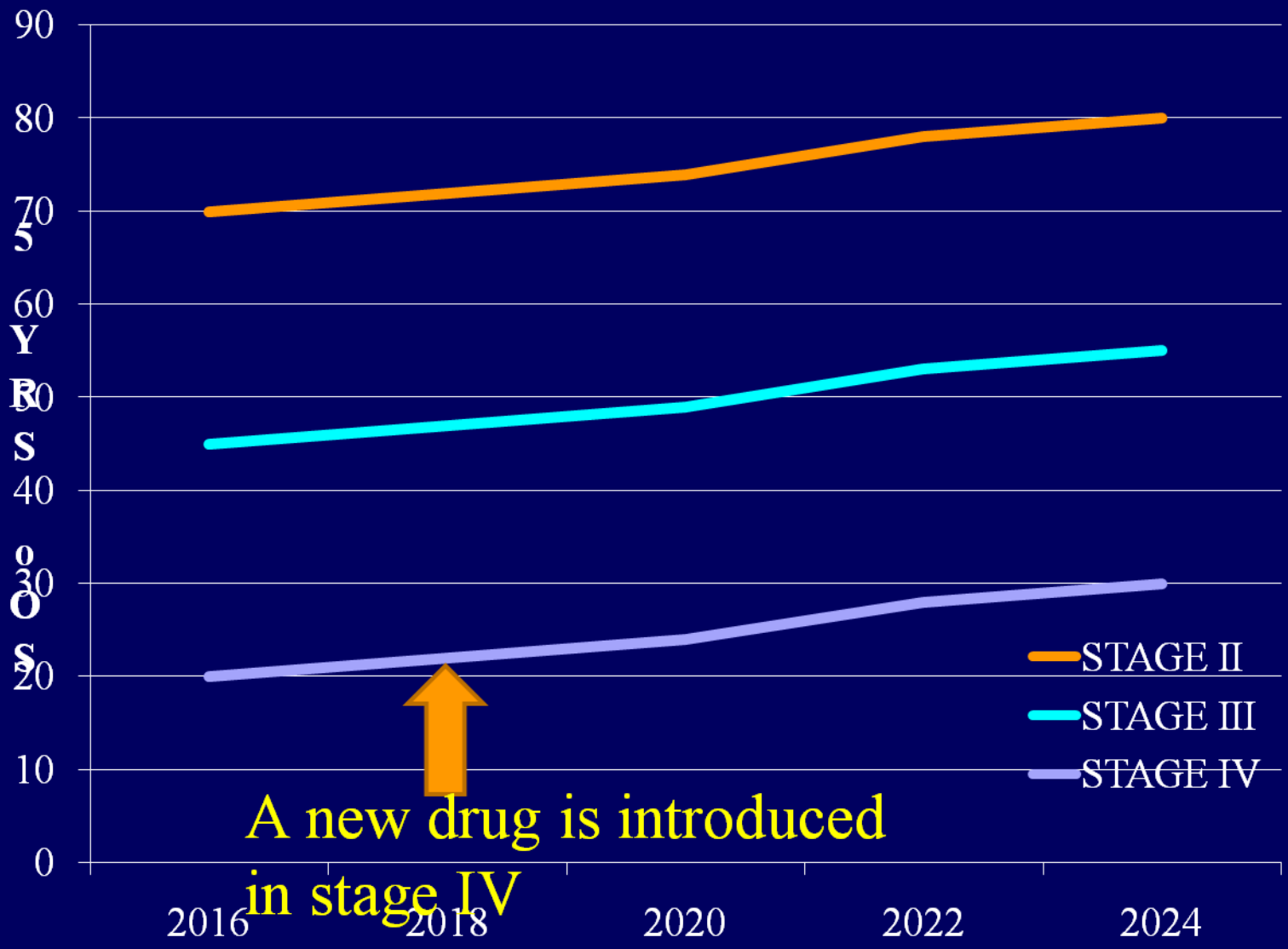


OS by stage over time

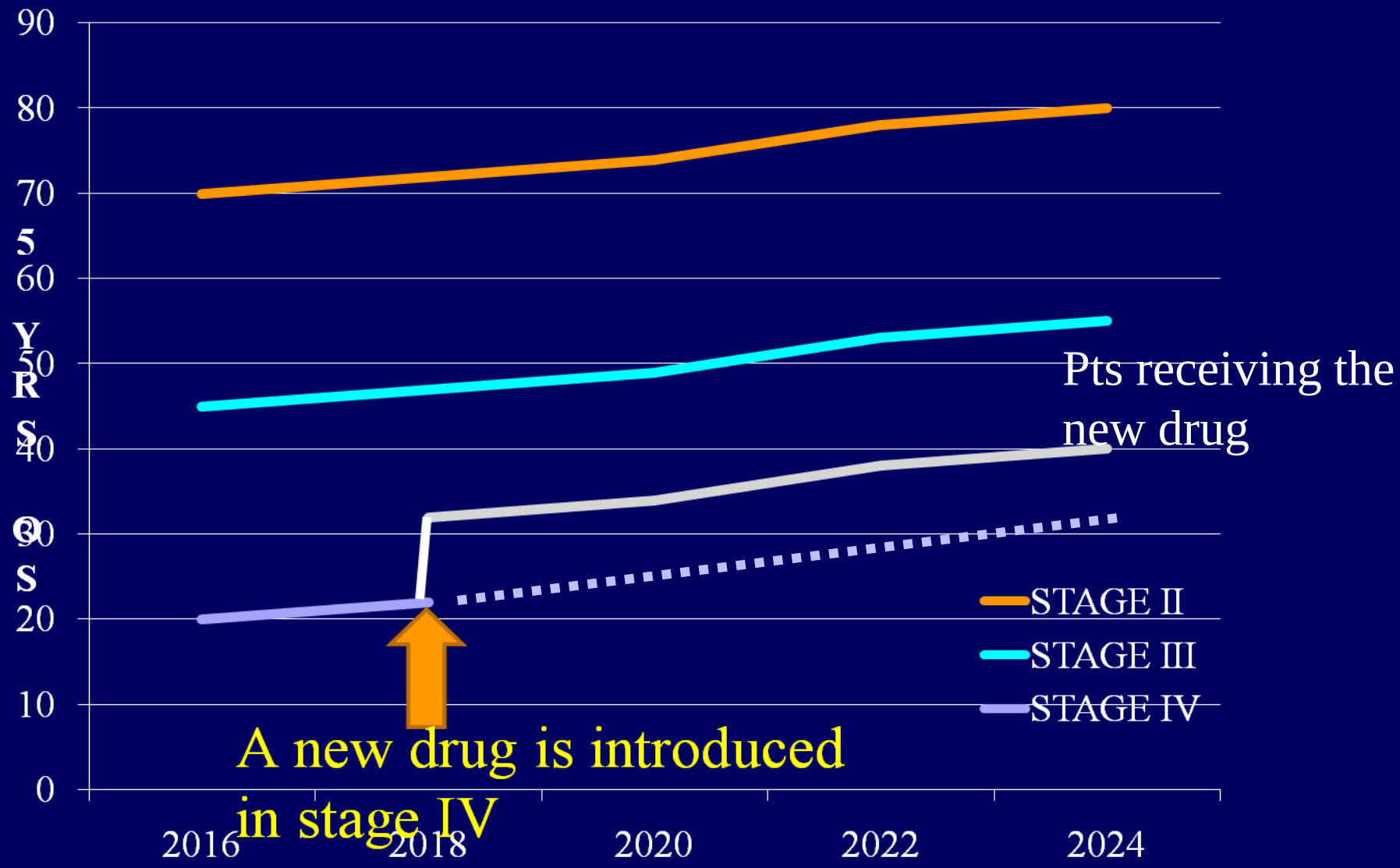


Increase in OS due to Will Roger's & other biases

OS by stage over time



OS by stage over time



Requirements

- Blinded review of selection criteria and prognostic factors to ensure comparability
- Analysis resembles that of a randomised trial (Intention to treat, Consort Chart, etc.)
- Comparison is with “expected” OS based on historical trends or/**and** OS observed in centers not participating to the trial
- Limited Reliance on statist. adjustments (multivariate analyses, propensity score)

Question

Positive evidence of a study of this kind would be deemed sufficient for registration of a new drug?

If yes, the establishment of prospective disease registries with adequate clinical and molecular data/material is a priority, particularly in cancer conditions where RCT's are problematic (grim prognosis, rarity, etc.)

Single Arm Trial: not a shortcut to registration

- Rationale
- Primary Aim
- Design
- Endpoint-Masking
- Selection Criteria
- ~~Randomization~~
- Treatment Protocol
- Statistical Plan
- Interpretation of Results

Treated and untreated patients are comparable?

Conclusion

In rare cancers, we should not forget what we've learnt from frequent tumors