

**Adjusting the Crossover Effect in Overall
Survival Analysis Using a Rank
Preserving Structural Failure Time Model:
*The Case of Sunitinib GIST Trial***

Xin Huang¹ and Qiang (Casey) Xu²

¹Pfizer Oncology Statistics, ²Columbia University

Outline

- **The case of sunitinib GIST Trial**
- **The crossover issue in oncology clinical trials**
- **Rank preserving structural failure time (RPSFT) model**
- **Conclusions**

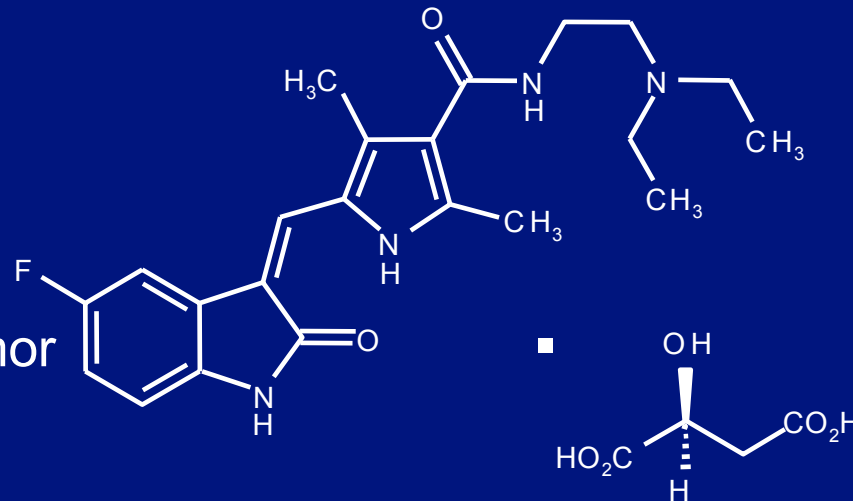
Sutent[®] (*sunitinib malate*)

- **Mechanism**

- Small molecule
- Inhibits multiple receptor tyrosine kinases (RTKs)

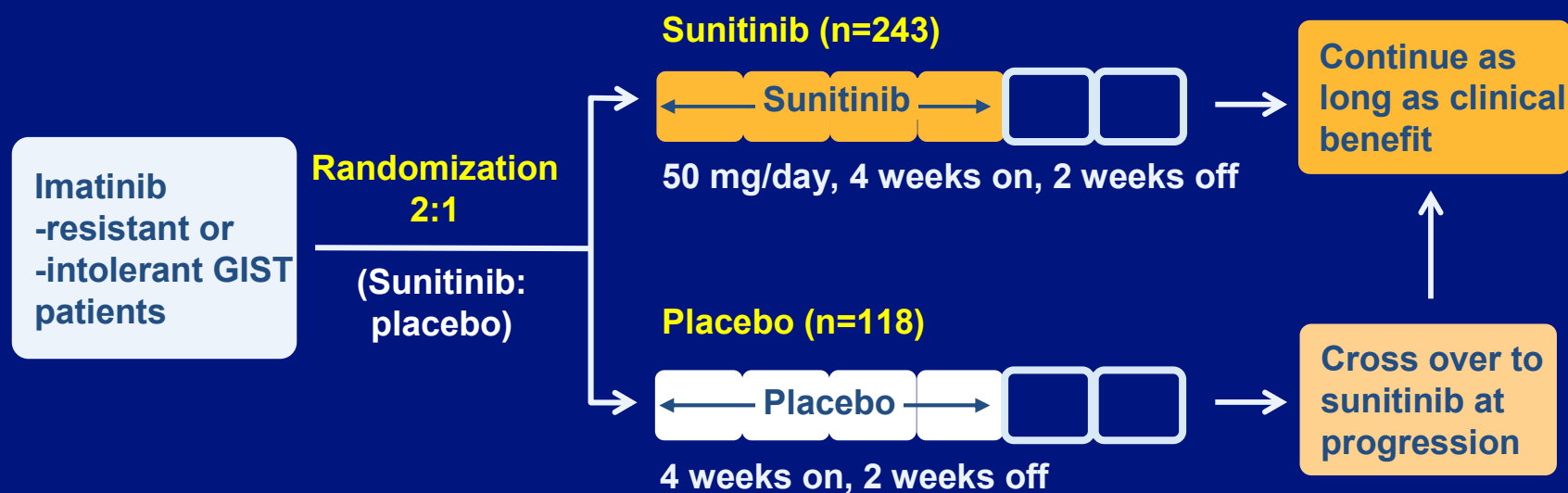
- **Indication**

- Second-line in gastrointestinal stromal tumor (GIST)
- First-line in renal cell carcinoma (RCC)
- First- or second-line in pancreatic neuroendocrine tumors (pNET)



Phase 3 Trial of Sunitinib in Imatinib-resistant/-intolerant GIST

GIST - Gastrointestinal Stromal Tumor



Analysis includes patients who enrolled during and subsequent to interim analysis

Study Objectives

- **Primary**

- **TTP (assessed by BICR according to RECIST)**
 - Sample size based on ability to detect a 50% increase in median TTP from 4 to 6 months

- **Secondary**

- **OS, PFS, ORR**
- **Patient-reported outcomes (pain control, health state)**
- **Safety monitoring**
- **Drug exposure and correlation with efficacy and safety**
- **Biomarkers and correlative kinase genotyping**

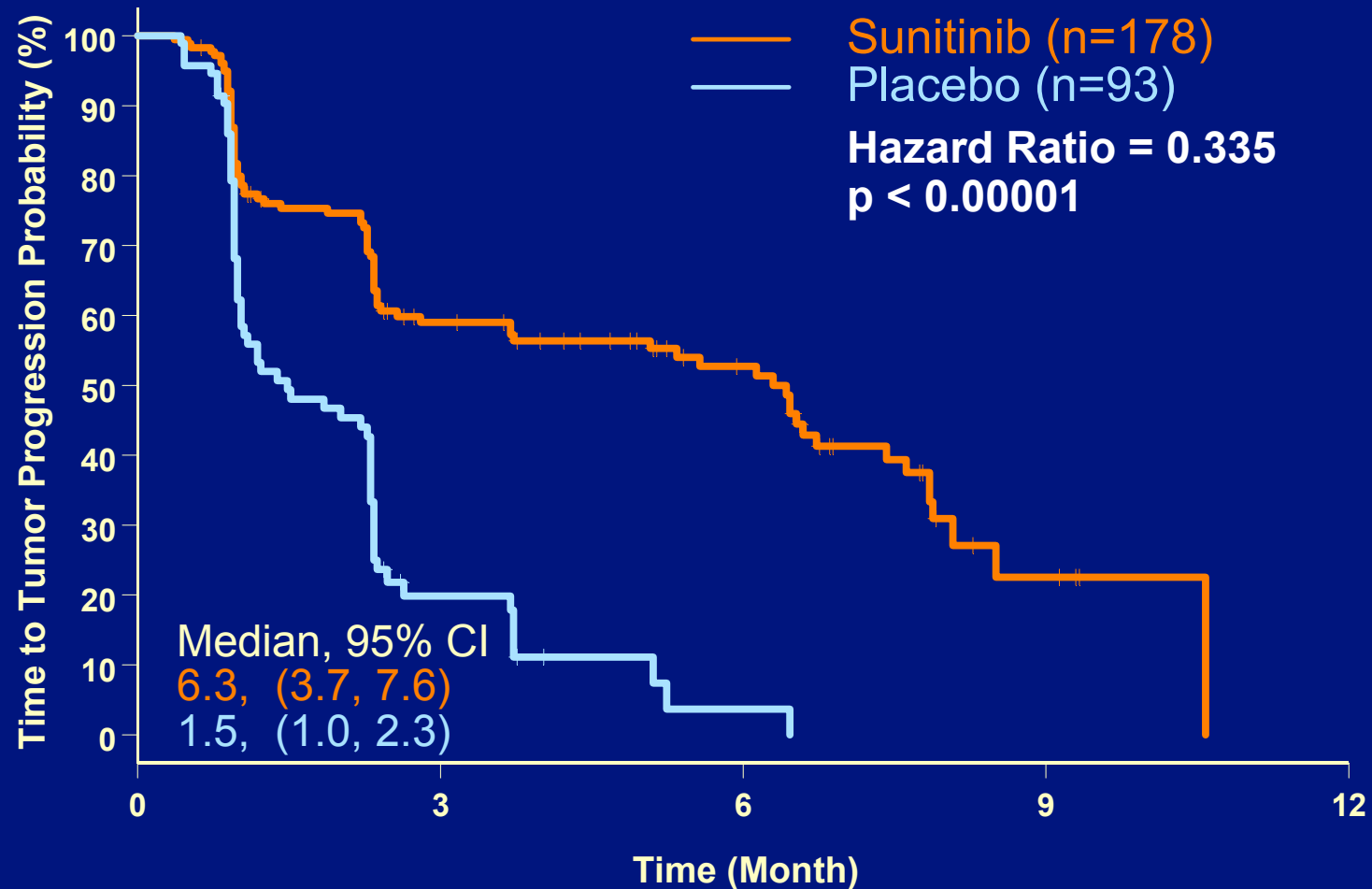
Key Oncology Endpoints

- **Progression Free Survival (PFS) – “surrogate”**
 - Time from randomization to tumor progression or death
 - Smaller studies, reflects tumor growth, not confounded by subsequent therapy
 - Not statistically validated as surrogate for OS, not precisely measured, requires balanced timing of assessments among treatment arms
- **Overall Survival (OS) – “gold standard”**
 - Time from randomization to death from any cause
 - Direct measure of benefit, easy and precise measurement
 - Larger studies, the treatment effect can be diluted or confounded by effective experimental or later line therapies

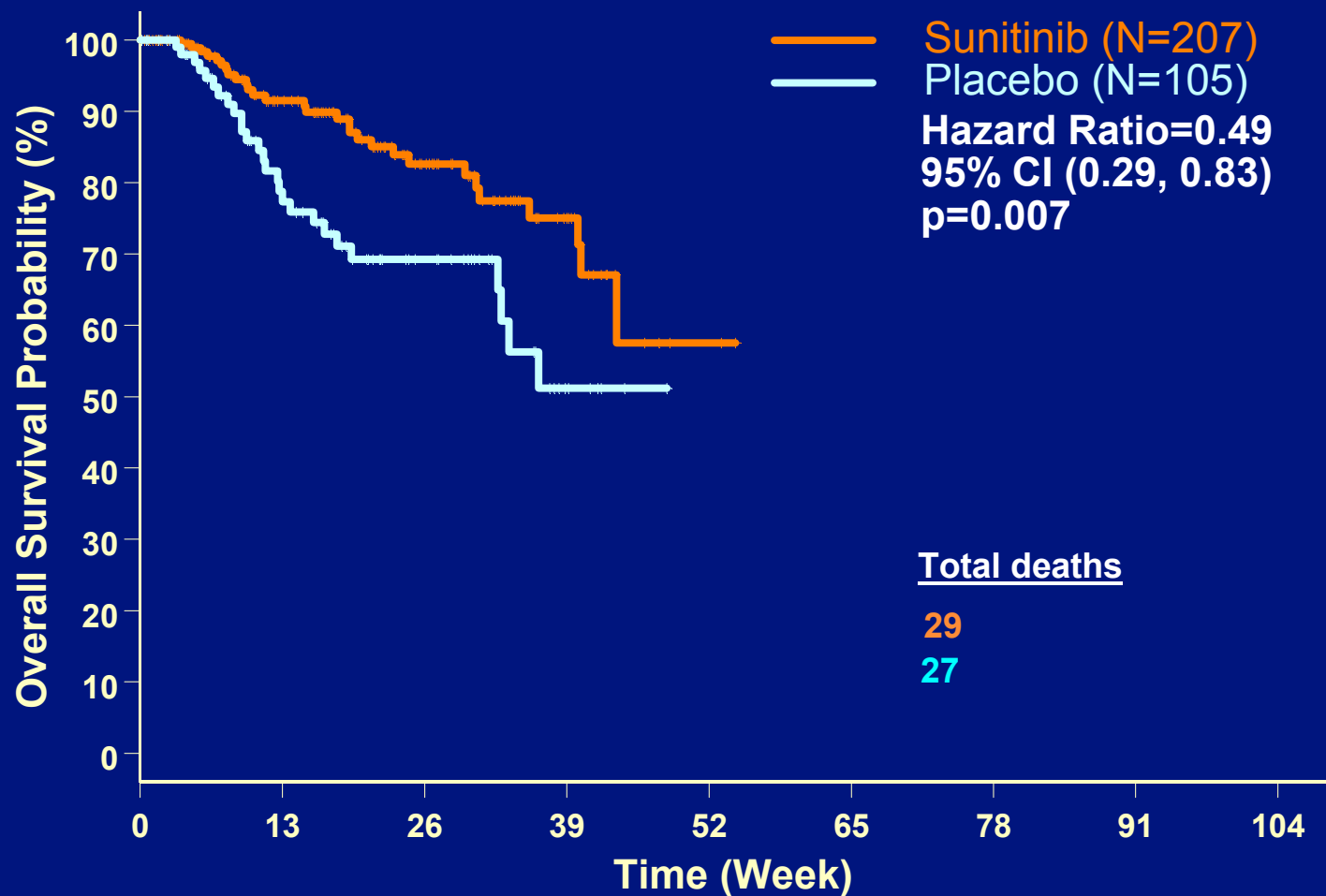
Results from the Primary Endpoint

- Patient accrual: December 2003 – May 2005
- First planned interim analysis for efficacy performed in January 2005
- Primary endpoint (TTP) statistically significant between sunitinib and placebo at interim analysis
- Treatment unblinded following recommendation by Independent Data Monitoring Committee (DMC) and all patients randomized to placebo were offered sunitinib.
- Based on these results, sunitinib received approval from US FDA (Jan 2006) and EU (July 2006) for treatment of GIST after disease progression on or intolerance to imatinib therapy

Time to Tumor Progression (Interim Analysis Based on BICR, 2005)

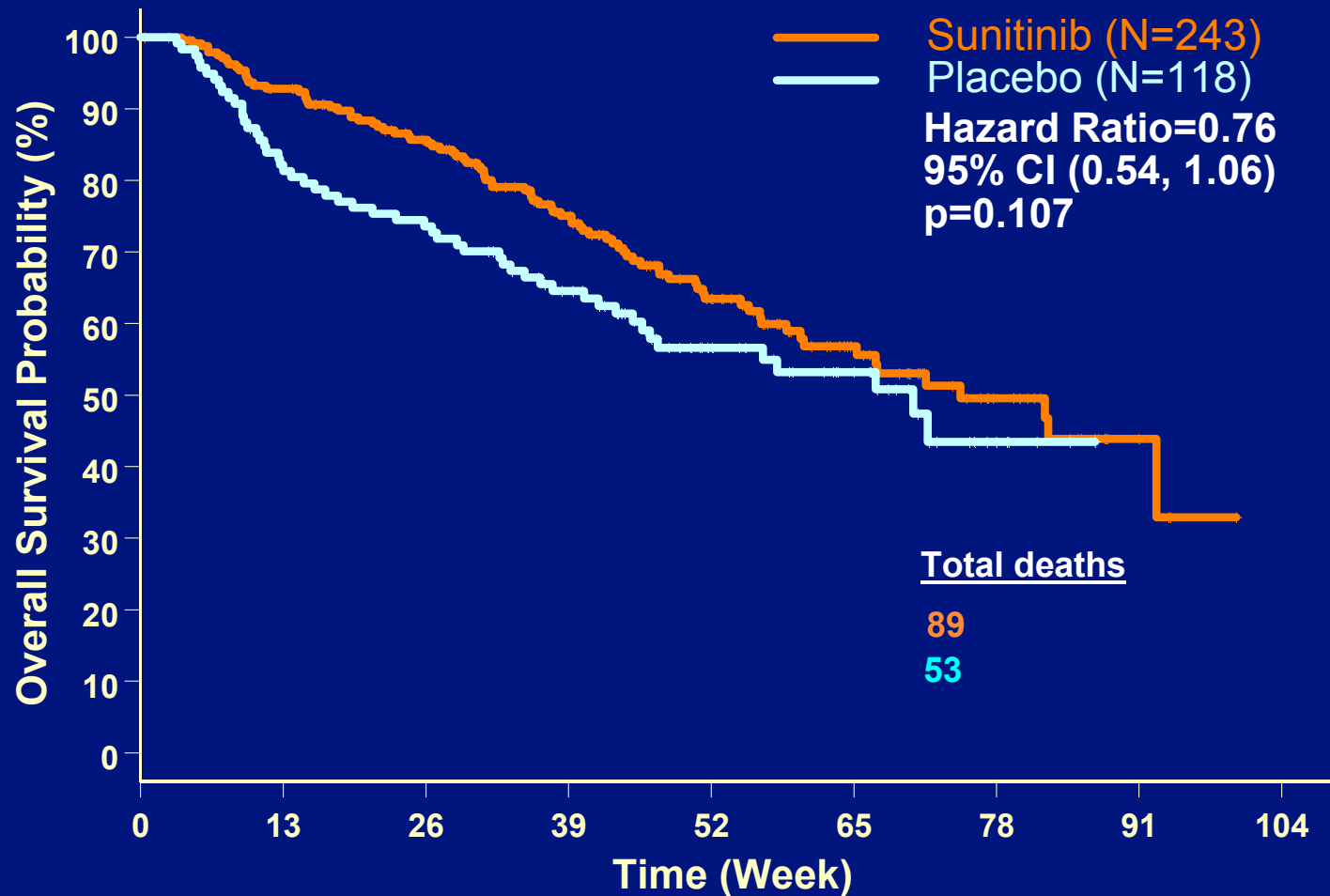


Overall Survival (NDA, 2005)



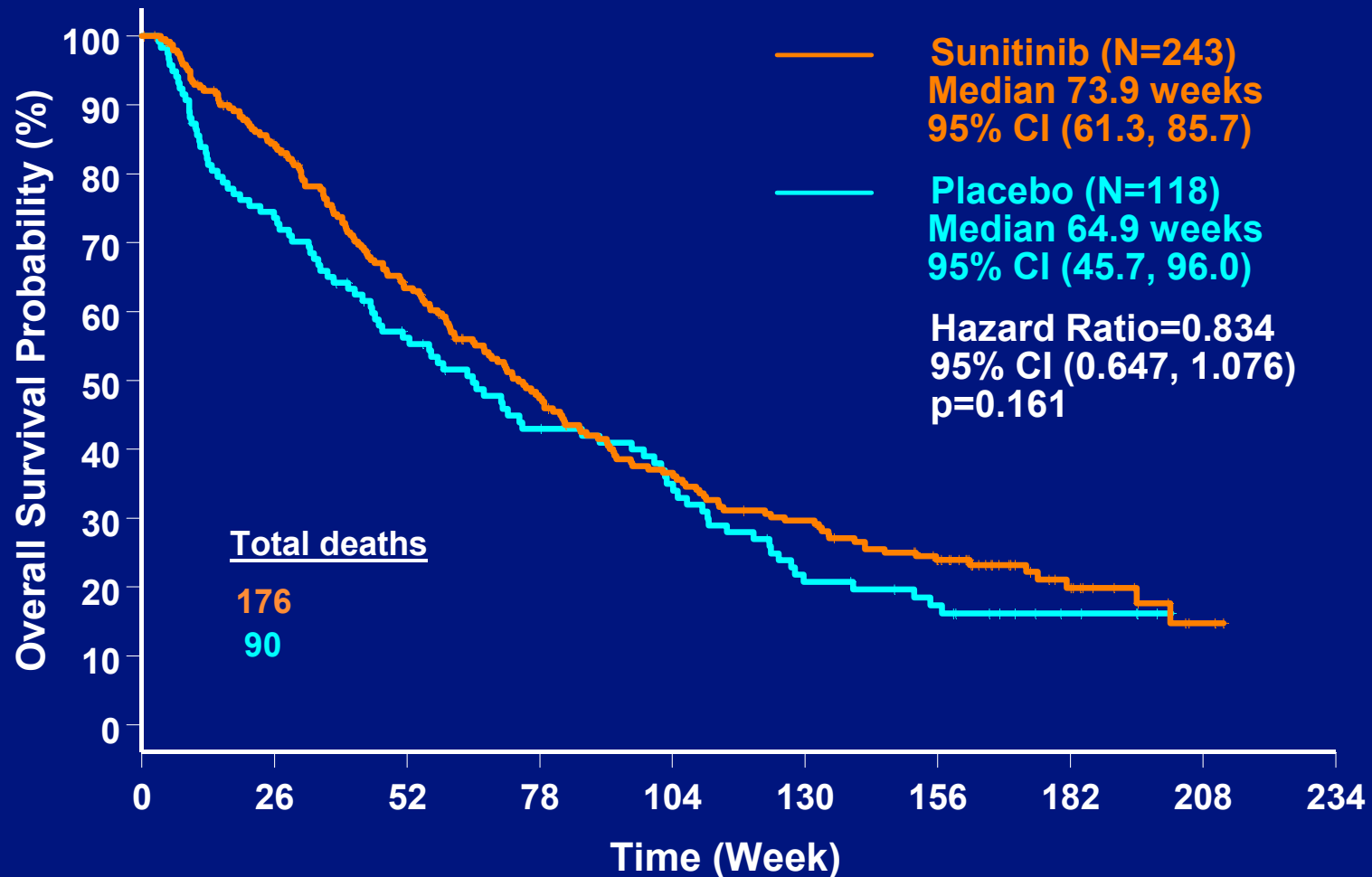
nRisk Sunitinib	207	13 / 114	9 / 61	4 / 25	3 / 2
nRisk Placebo	105	18 / 55	5 / 26	4 / 6	0 / NA

Overall Survival (ASCO, 2006)



nRisk Sunitinib	243	17 / 214	16 / 187	22 / 142	19 / 86	7 / 47	5 / 23	2 / 5
nRisk Placebo	118	22 / 96	9 / 84	10 / 66	7 / 37	2 / 25	3 / 6	0 / NA

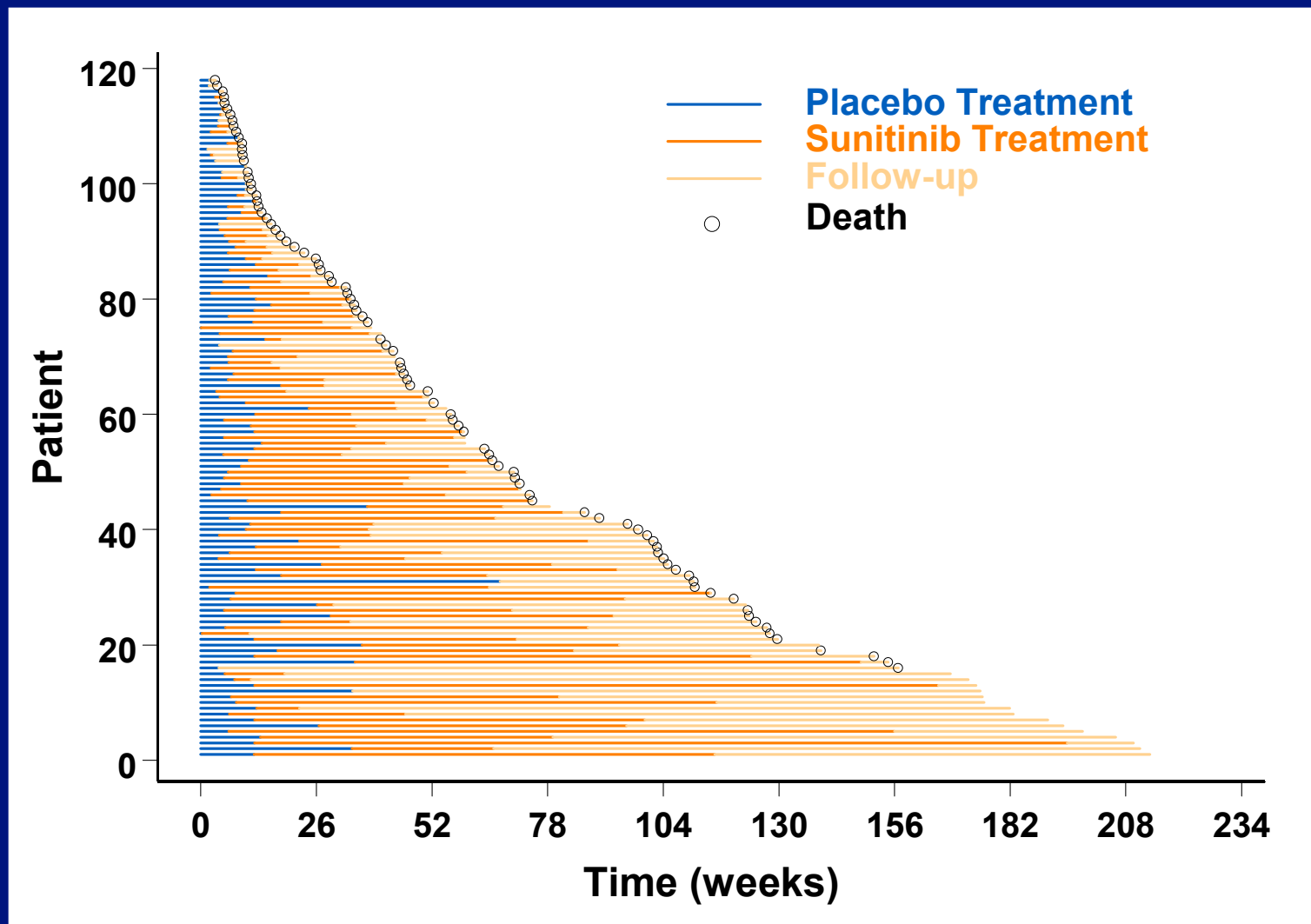
Overall Survival (ASCO, 2008)



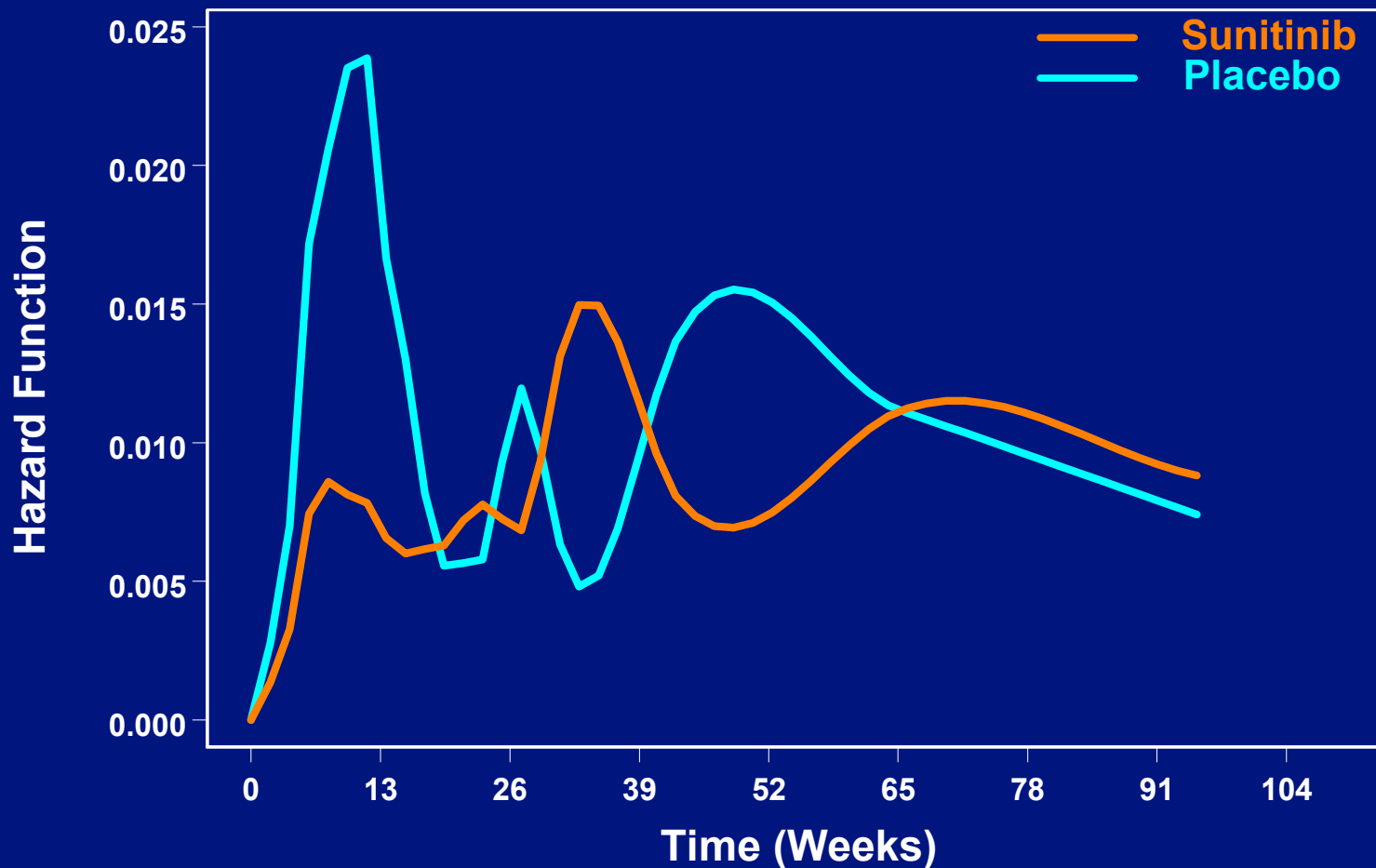
What We Know about Placebo Patients

- **118 patients were randomized to placebo arm**
 - 90 (76.3%) patients died
- **103 (87.3%) patients crossed over from placebo to sunitinib treatment**
 - 83 (70.3%) patients crossed over within 3 months
 - 19 (16.1%) patients crossed over before disease progression
 - 4 (3.4%) patients never treated with placebo
- **15 (12.7%) patients did not crossover**
 - 12 patients died

Treatment Duration at Patient Level for Placebo Arm



Estimated Hazard Functions by Treatment



Hazard function is the *instantaneous* failure rate at any point in time

Impact of Crossover to Overall Survival Analysis

- **Discontinuation or change in therapy (crossover effect) after disease progression or adverse event**
- **Whether the lack of efficacy results from a lack of benefit of the treatment, or the crossover has obscured the benefit of treatment?**
- **Conventional methods cannot fully adjust for the bias caused by crossover and clinical effect may be underestimated**

Consequences of not Having an OS Improvement

- **Higher hurdle to obtain the regulatory approval**
 - Bevacizumab breast cancer indication
- **Great challenge to be accepted by the payers**
 - UK National Institute for Health and Clinical Excellence (NICE) cost effectiveness analyses

Standard Approaches to Analyze OS

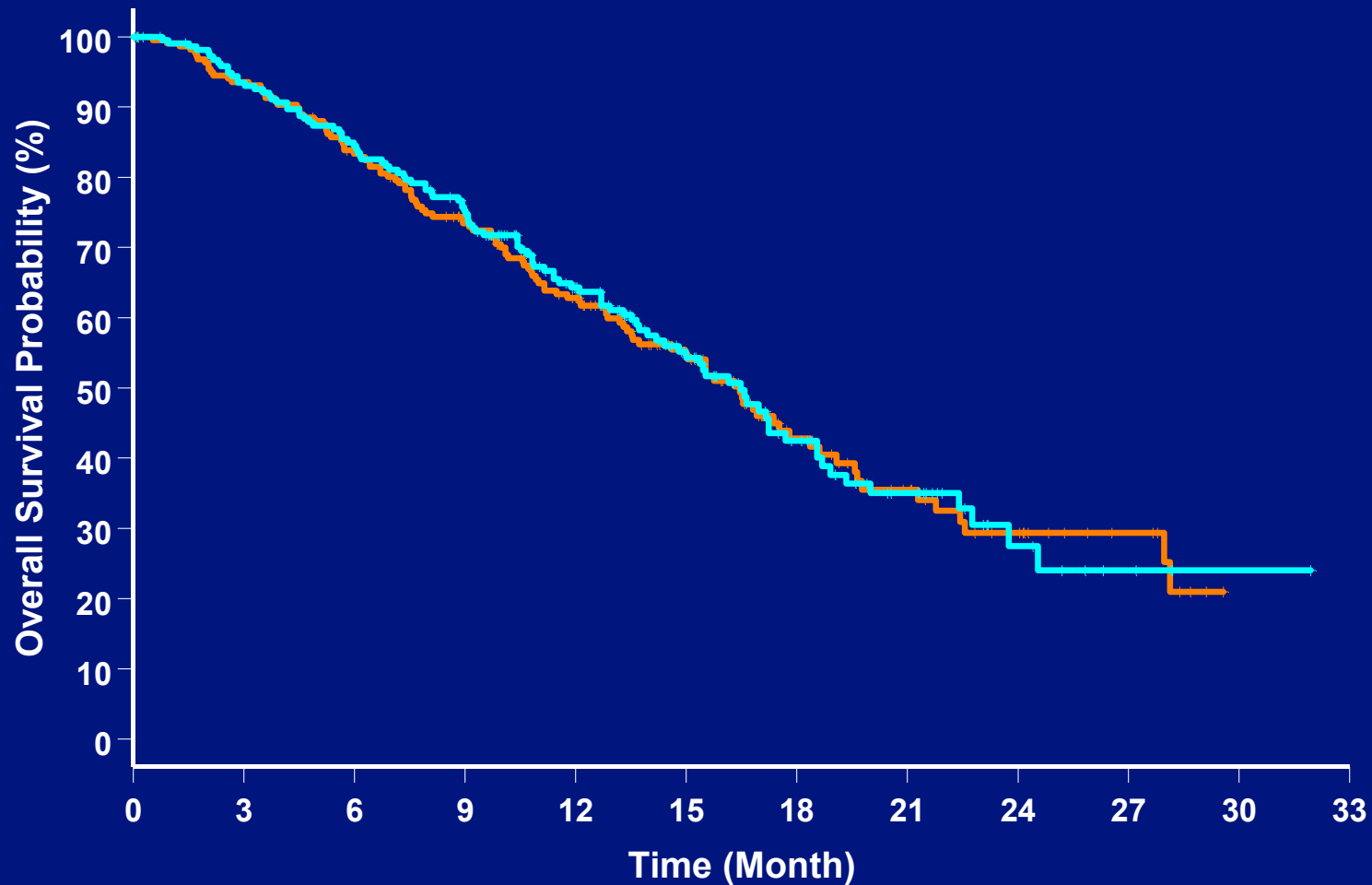
- **ITT Analysis**
 - Analyze patients as randomized and makes no adjustment for crossover
- **Per-protocol Analysis**
 - Drop patients who crossed over
- **On-Treatment Analysis**
 - Censor patients when they crossover
- **Time Dependent Treatment Analysis**

The later 3 approaches break the “exchangeability” created by randomization. Outcomes are not representative of what would be expected from the originally randomized group

Why/When Does Crossover Cause Bias?

- **Outcomes in the control arm reflect the benefit of the experimental drug**
- **Crossover is a selective process**
 - Not all of the patients who progress with the control drug cross over
- **The timing of crossover**
- **For crossover to cause bias, the experimental treatment must have some benefit for the endpoint**
- **If two treatments have the same effect, switching from one to the other shouldn't affect the endpoint**
- **The indication that crossover may cause bias can be identified by early signs of benefit**
 - Earlier separation of survival curves, with gradual approaching due to increasing crossover

Unlikely Crossover Can Cause Bias in This Case



Estimation of Treatment Effect with Crossover

- Prefer intent-to-treat analysis and desire to compare treatment groups as randomized
- Estimate the treatment effect, *as if no patients in placebo arm had ever crossed over to experimental arm.*
- Rank preserving structural failure time (RPSFT) model proposed by Robins and Tsiatis (1991) can be used to dealing with the crossover problem

RPSFT Model

- RPSFT model maintains the original randomization group and preserves the validity of between group comparison
- RPSFT model produces a “randomization-based estimate” of treatment effect corrected for the bias introduced by crossover
- Assuming treatment has a multiplicative effect ($e^{-\eta}$) on a patient’s lifetime – **Accelerated Failure Time (AFT)** model
 - Values of $e^{-\eta} > 1$ or $\eta < 0$ reflect a beneficial treatment effect
 - Values of $e^{-\eta} < 1$ or $\eta > 0$ indicate a detrimental effect of treatment
 - Equality of expected time to event for experimental and placebo arms holds when $e^{-\eta} = 1$ or $\eta = 0$
- AFT model describes a relationship between the survival functions of any two individuals
 - *A good example is the conventional wisdom that a year for a dog is equivalent to 7 years for a human*
- In RPSFT, $e^{-\eta}$ is estimated under a “test-based” procedure

RPSFT Assumptions

- **Rank preserving**

- Assume treatment doesn't change the order of events
- It just extends or shortens the time when the events happen
- The proportion of the extension/shortening is the same for all treated patients

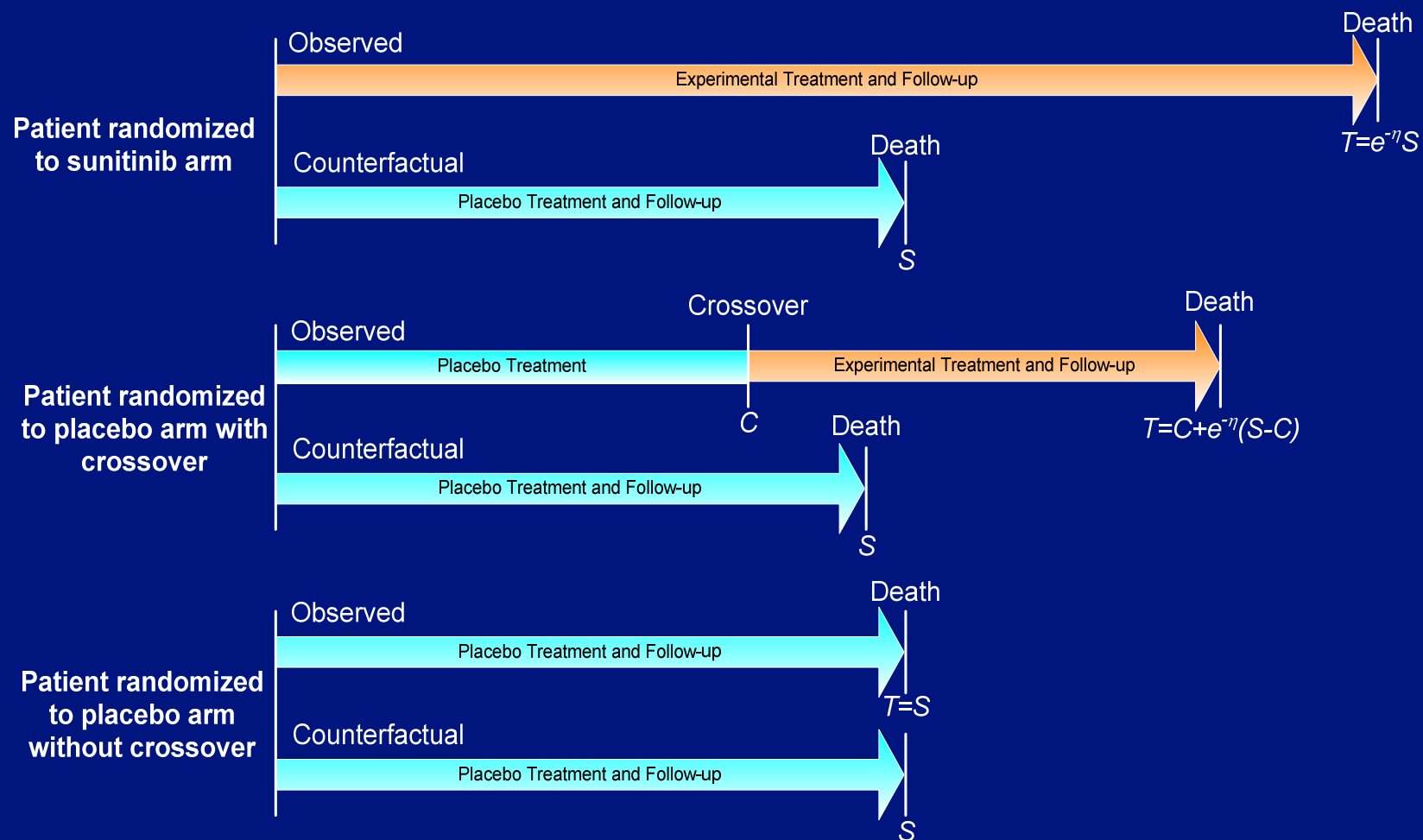
- **Structural = Casual**

- A casual assumption of time proportionality instead of a proportional hazards assumption of the Cox model
- The model describes the relationship between the event times of a given individual if she or he receives the experimental drug or control drug
- Only one of these times is observed
- Depend on the structural model being correct, or at least reasonably captures the effect of treatment

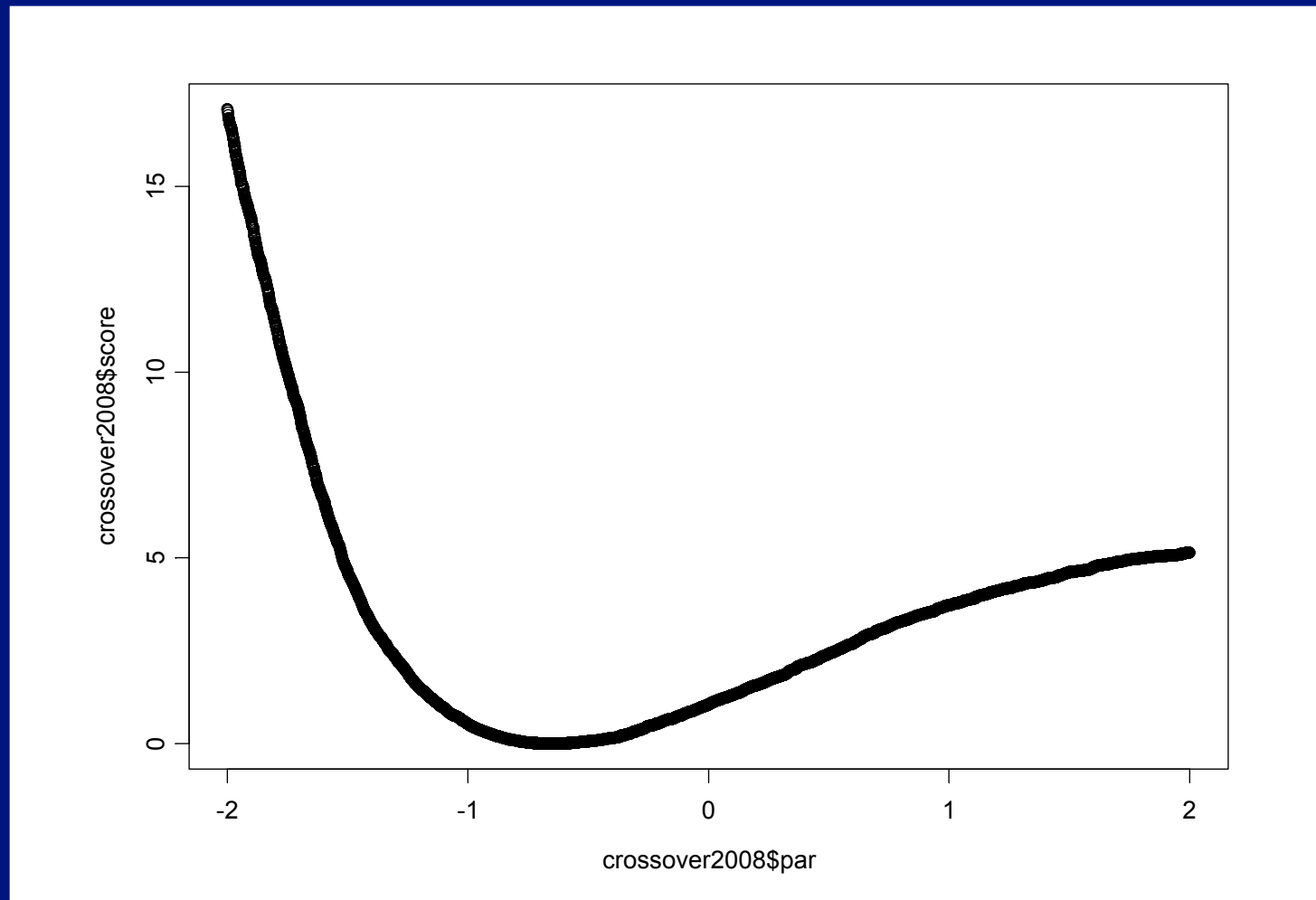
Estimation of η

- An individual's *counterfactual* survival time, S , does not depend on which treatment arm the individual is randomized
- Assumption of *Accelerated Failure Time (AFT)*
 - $T = e^{-\eta} S$, T is observed survival time
 η is an unknown parameter
- For patients randomized to experimental arm
 - $S = e^{\eta} T$
- For patients randomized to placebo arm (crossover)
 - $S = C + e^{\eta} (T - C)$, C is the crossover time
- For patients randomized to placebo arm (no crossover)
 - $S = T$
- A grid search over possible values is used to determine the point estimate for η
- The point estimate for η is given by the log-rank statistic $Z(\eta)=0$

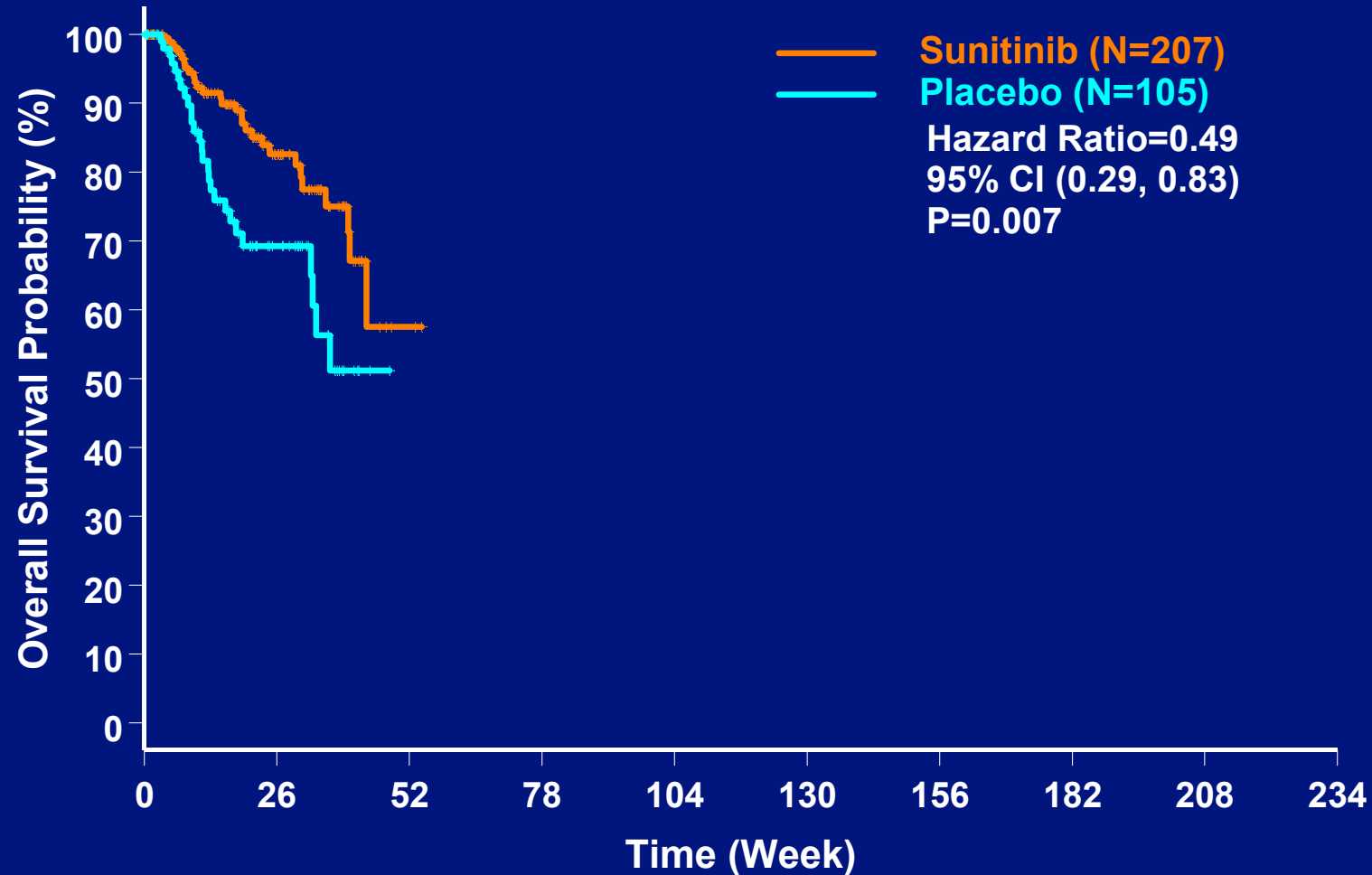
Illustration of Observed and Counterfactual Survival Times for Individual Patients ($e^{-\eta} > 1$)



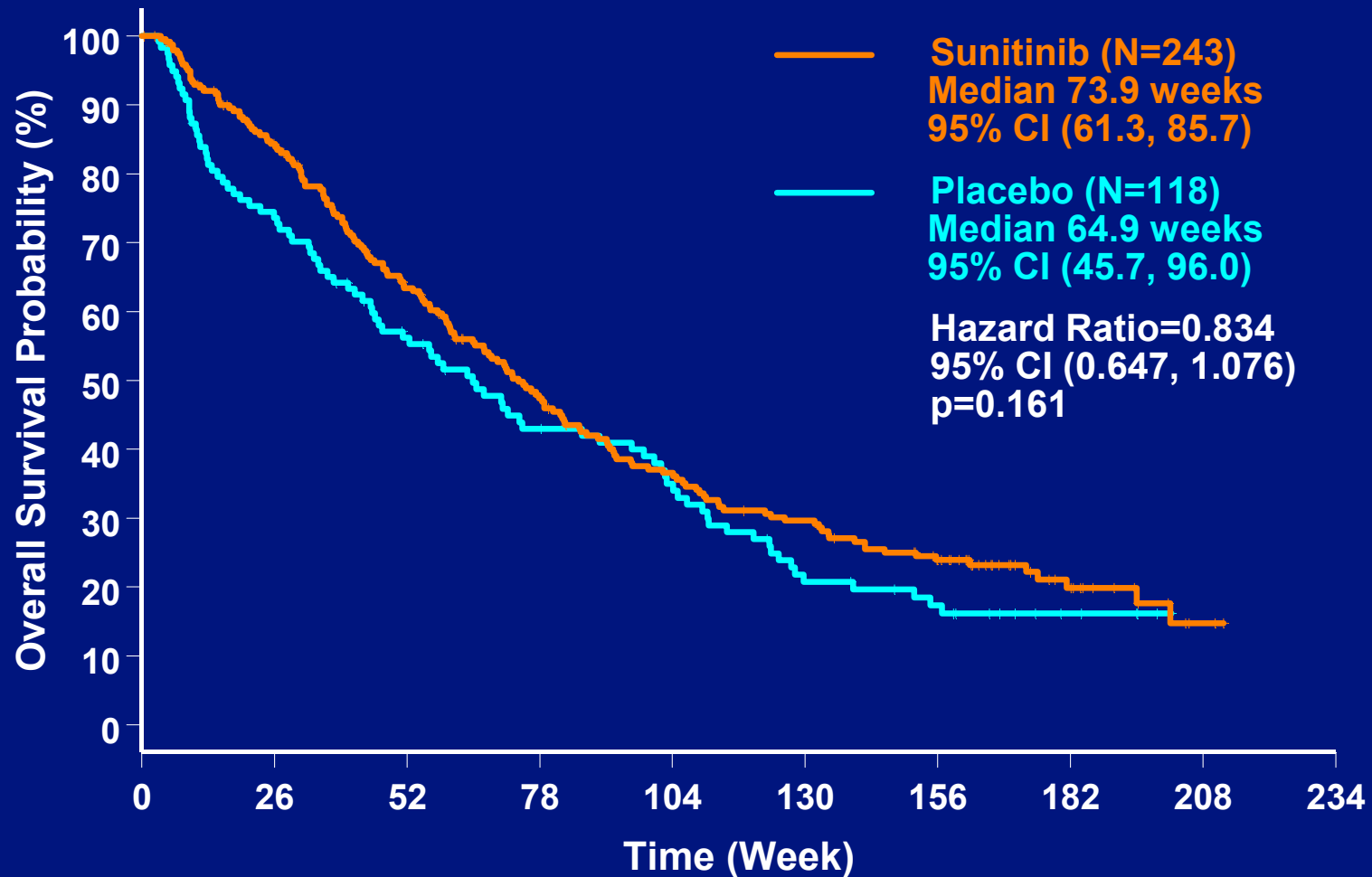
A Grid Search over Possible Values for η



Overall Survival (NDA, 2005)

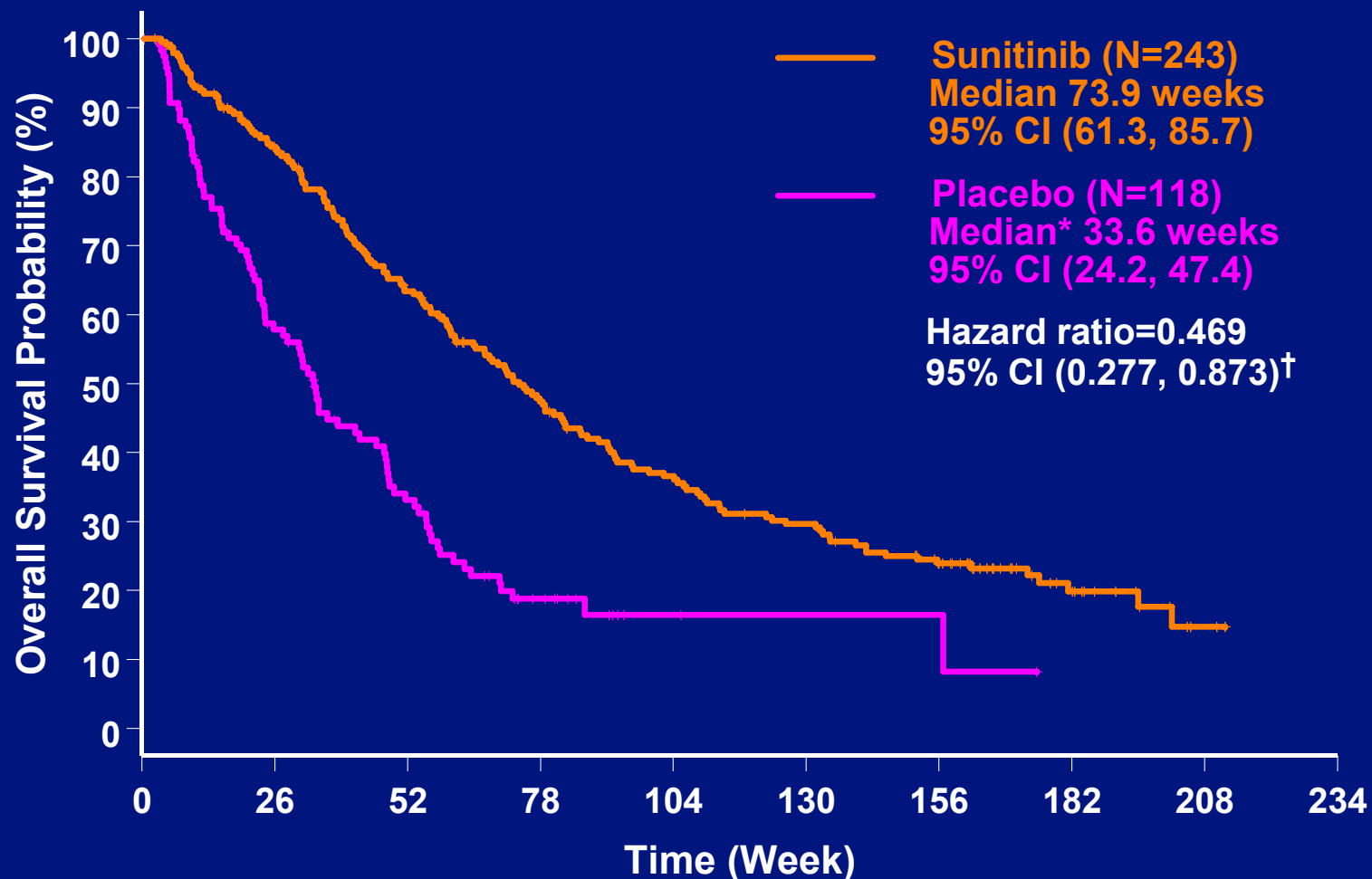


Overall Survival (ASCO, 2008)



Overall Survival (ASCO, 2008)

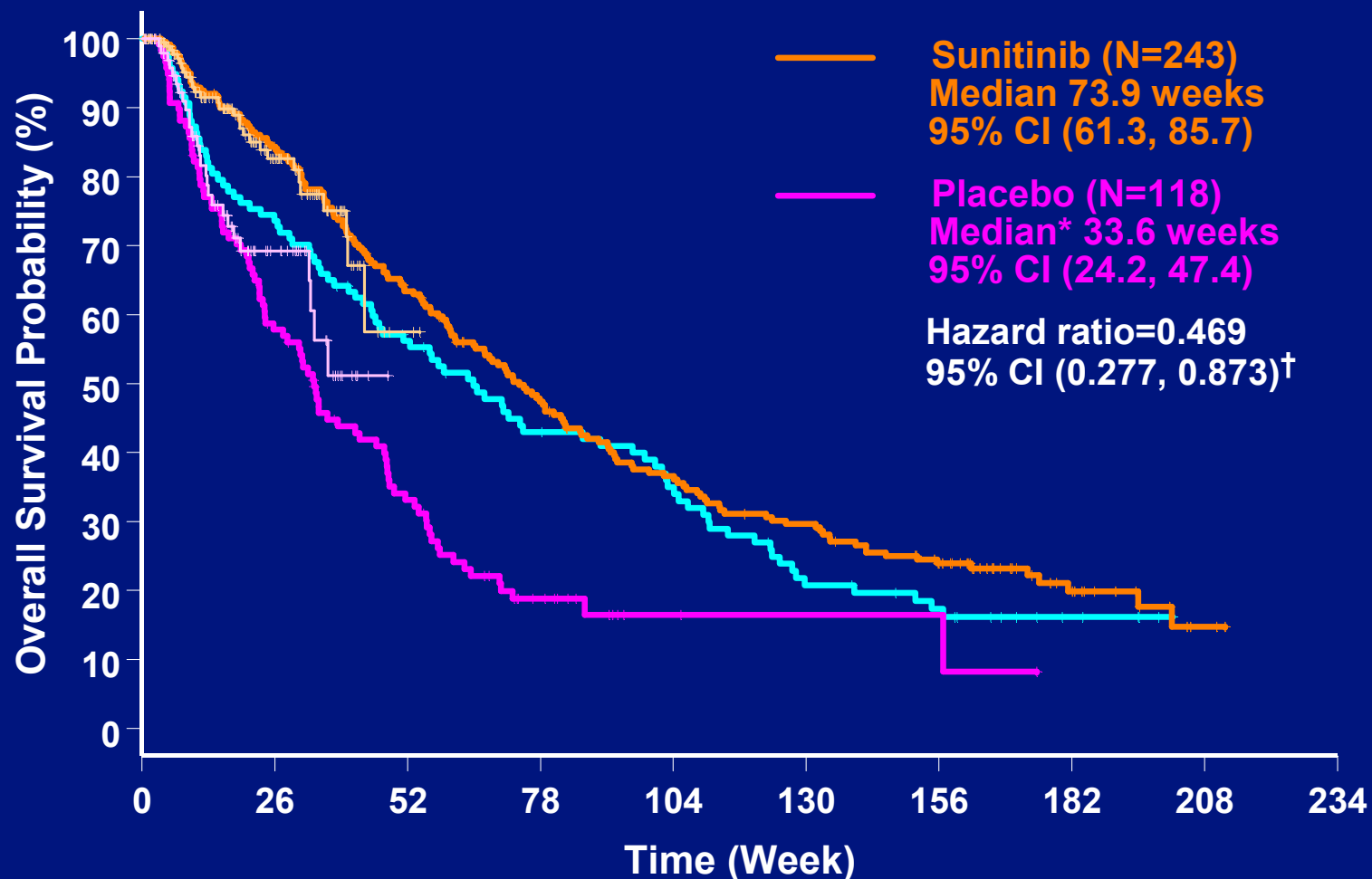
Crossover corrected by RPSFT



*Estimated by RPSFT model †Empirical 95% CI obtained using bootstrap samples.

Overall Survival (ASCO, 2008)

Crossover corrected by RPSFT



*Estimated by RPSFT model †Empirical 95% CI obtained using bootstrap samples.

The End of the Story

- **The final GIST OS result is included in USPI (2009)**
 - “..... Ninety-nine of the patients initially randomized to placebo crossed over to receive SUTENT in the open-label treatment phase. At the protocol specified final analysis of OS, the median OS was 72.7 weeks for the SUTENT arm and 64.9 weeks for the placebo arm [HR= 0.876, 95% CI (0.679, 1.129)].”
- **NICE final appraisal determination sunitinib for the treatment of GIST (May, 2009)**
 - “Sunitinib is recommended as a treatment option for people with unresectable and/or metastatic malignant gastrointestinal stromal tumours

Applications of RPSFT in Cost Effectiveness Analyses

Therapy and Tumor Type	Method
Sunitinib for the second-line treatment of GIST	RPSFT
Pazopanib for the first-line treatment of RCC	RPSFT IPCW*
Everolimus for the second-line treatment of RCC	RPSFT IPCW
Trastuzumab + anastrozole for postmenopausal women with HER2+ and HR+ BC	RPSFT
Letrozole and anastrozole vs. tamoxifen as adjuvant therapy in postmenopausal women with early BC	RPSFT IPCW

*IPCW – Inverse probability of censoring weighting methods

Conclusions

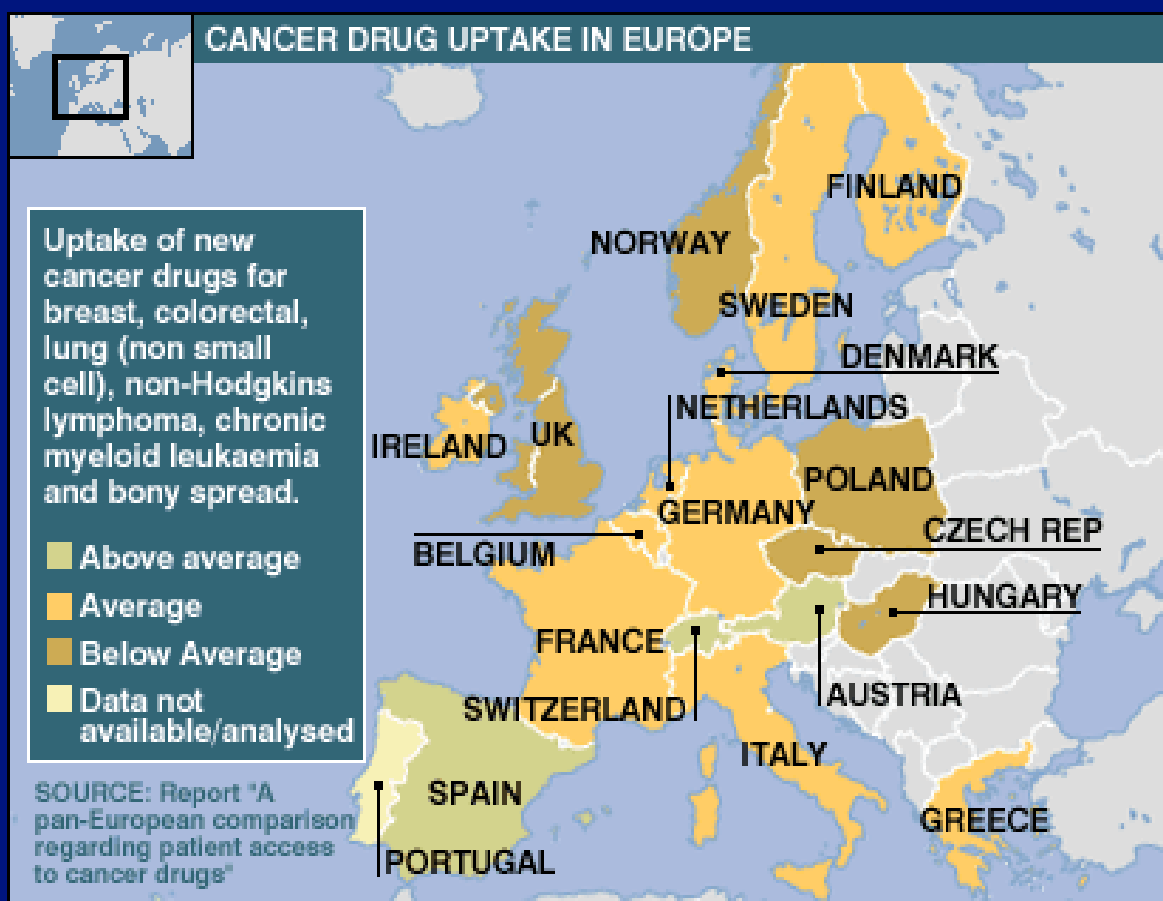
- Crossover is a common and unavoidable issue in oncology clinical trials and its impact should be properly addressed
- RPSFT model conceptually can address the problem if assumptions are considered to be reasonable
- RPSFT model provides a randomization-based estimate of treatment effect corrected for the bias caused by crossover
- RPSFT model can be used to perform a sensitivity analysis to support the ITT analysis
- The benefit demonstrated in TTP in sunitinib GIST trial is likely to translate into an OS benefit in Imatinib-resistant or intolerant GIST patients

References

- Robins, J. M., and Tsiatis, A. A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics. - Theory and Methods*, 20, 2609-2631.
- Robins, J. M., Mark, S. D., and Newey, W. K. (1992). Estimating exposure effect by modeling the expectation of exposure conditional on confounders. *Biometrics*, 48, 279-495.
- Mark, S. D., and Robins, J. M. (1993). Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. *Statistics in Medicine*, 12, 1605-1628.
- Branson, M., and Whitehead, J. (2002). Estimating a treatment effect in survival studies in which patients switch treatment. *Statistics in Medicine*, 21, 2449-2463.

UK National Institute for Health and Clinical Excellence (NICE) cost effectiveness analyses

Some of the Slowest and Lowest Uptake Rates for New Cancer Drugs in Europe



Only Czech Rep, Poland, Slovenia and Denmark are worse than the UK

Without NICE Approval, Access is Typically Limited vs. Other EU Markets

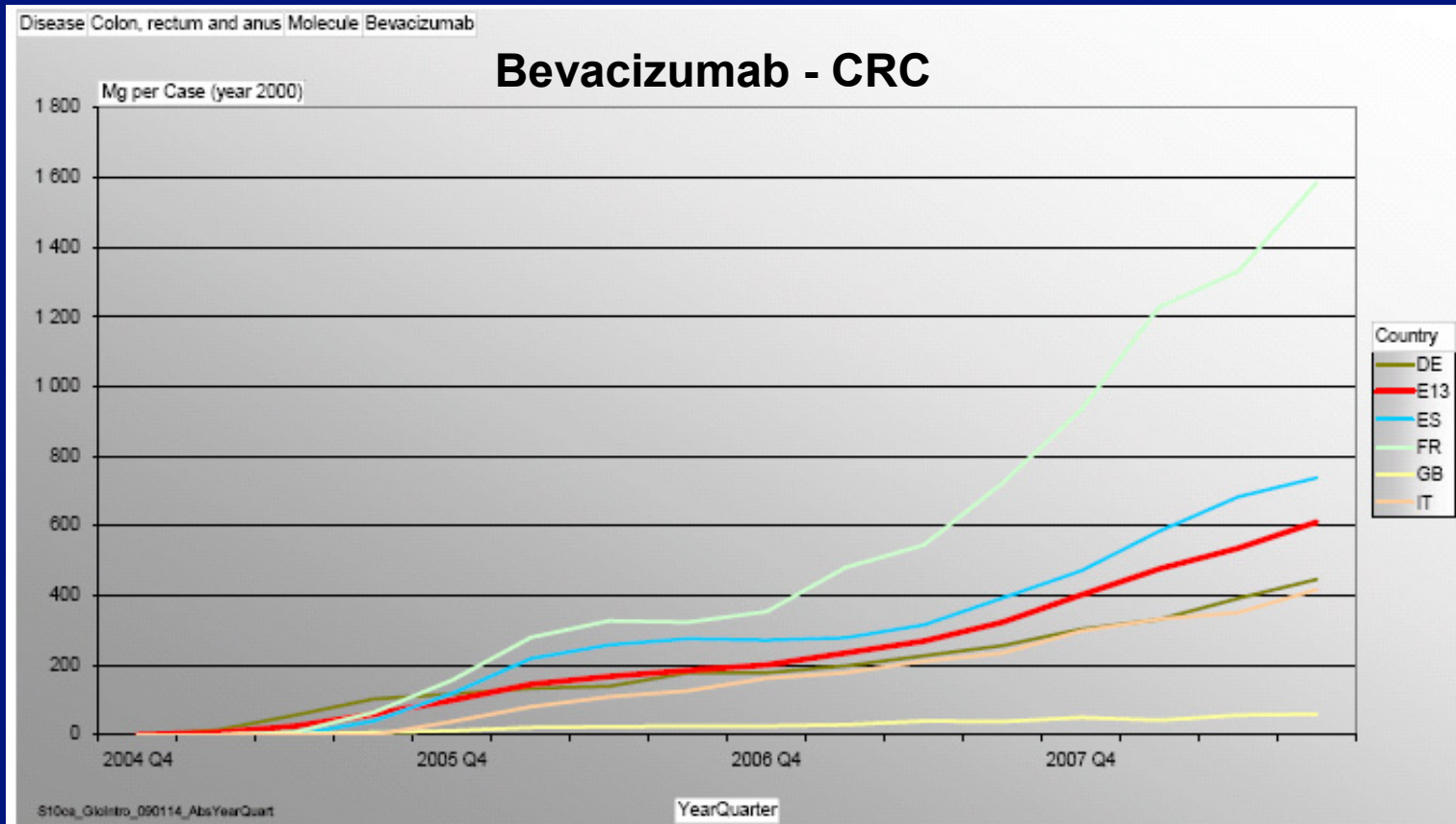


Figure 3-26. Usage of bevacizumab expressed as mg/case (related to mortality in colorectal cancer in 2000) in E13, France, Germany Italy, Spain and the UK. Please note that bevacizumab is also indicated for breast-, lung- and renal cell cancer.

Positive NICE Guidance Leads to Open Access

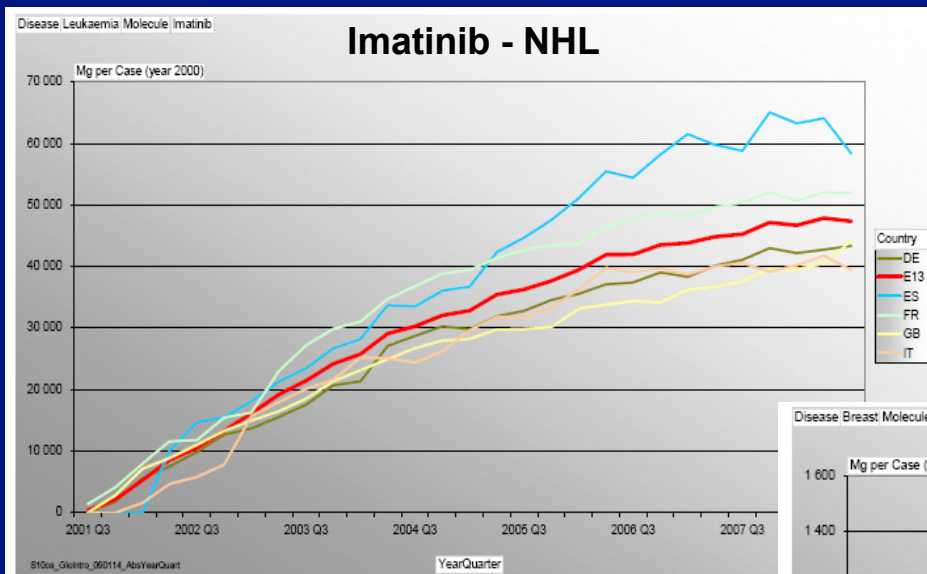


Figure 3-30. Usage of imatinib expressed as mg/case (related to mortality in 2000) in E13, France, Germany Italy, Spain and the UK.

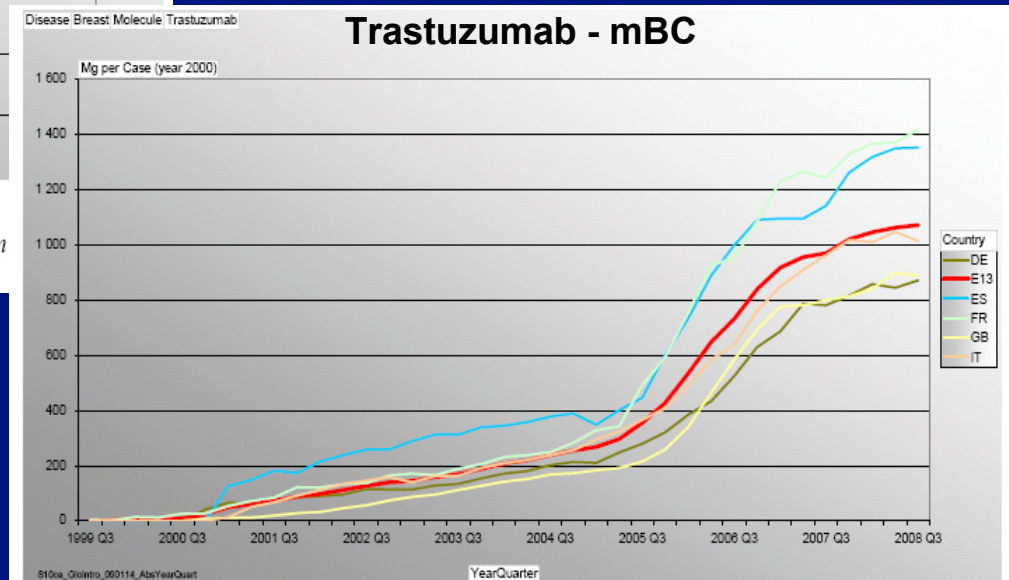
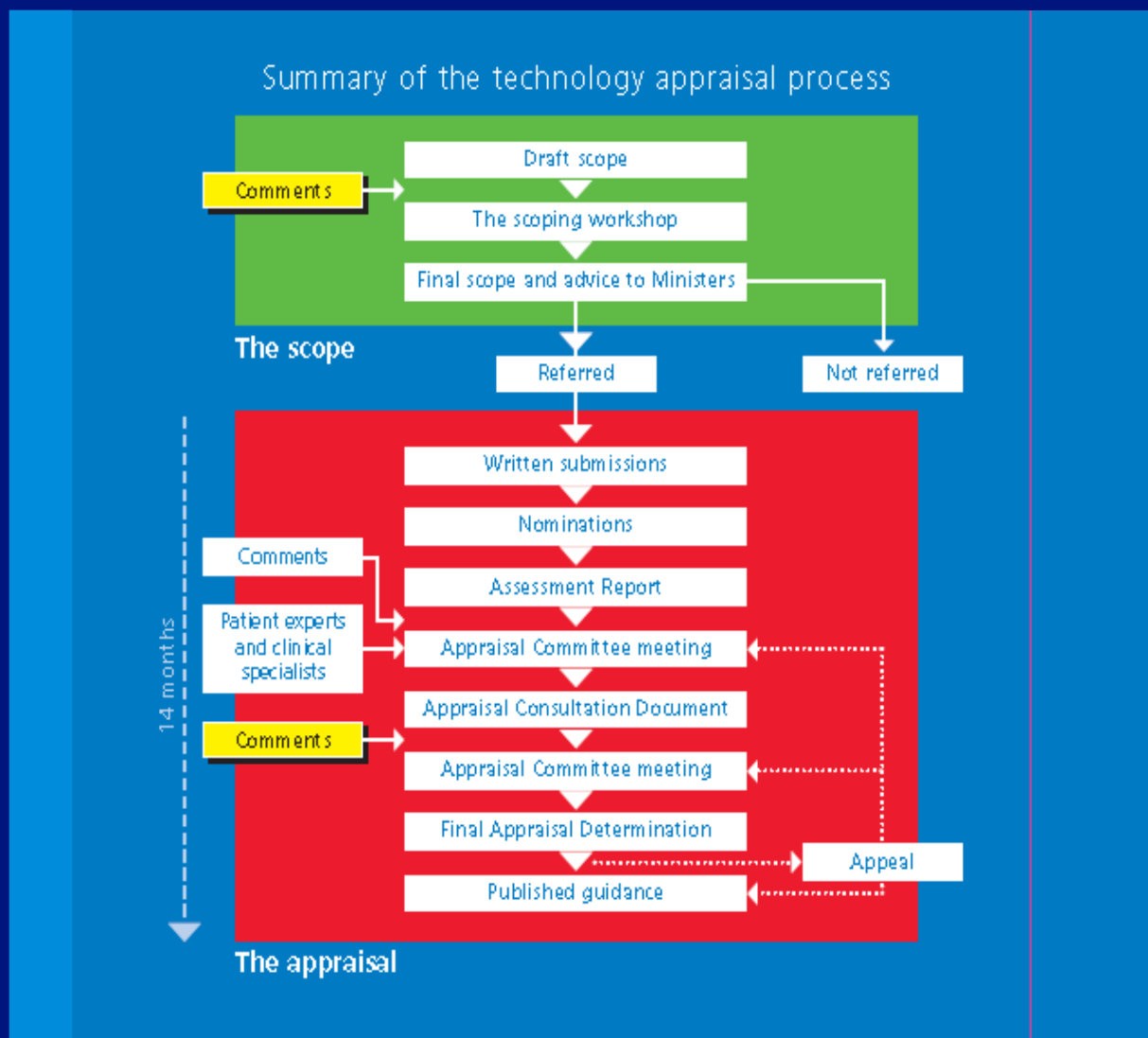


Figure 3-18. Usage of trastuzumab expressed as mg/case (related to mortality in breast cancer in 2000) in E13, France, Germany Italy, Spain and the UK.

NICE Multiple Technology Appraisal Process



What is a Technology Appraisal?

- The purpose of an appraisal is to appraise the health benefits and the costs of a technology and to make recommendations to the NHS in England and Wales
- Appraisals considers the evidence of the health benefits and costs of a technology, this includes:
 - Impact on Quality of Life
 - Probable effects on mortality
 - Estimates of the associated costs specifically to the NHS and Personal Social Services

Who Makes the Decision?

- **The Appraisal Committee makes recommendations to NICE on how the technology should be used.**
- **The Committee is independent of NICE and has members who include:**
 - Doctors
 - Nurses
 - NHS managers
 - Health Economists
 - Statisticians
 - Lay representatives
- **The Committee have expertise in undertaking appraisals but will not necessarily be experts in the specific clinical area.**
- **The Committee will meet at least twice to consider the evidence**

What Evidence is Considered?

Evidence for a technology appraisal is derived from a number of sources:

- Technology assessment carried out by an independent academic group (the ‘Assessment Group’)
- Information (‘evidence’) provided by the consulters
- Participation of selected clinical experts
- Participation of selected patient experts

Key Documents Received from NICE

- **The Assessment Report**
 - This is a systematic review of the published and unpublished evidence on the clinical effectiveness and cost effectiveness of the technology. It DOES NOT make recommendations about how the technology could be used by the NHS
- **The Appraisal Consultation Document (ACD)**
 - This sets out the provisional views of the Appraisal Committee and may change in response to consultation.
- **The Final Appraisal Determination (FAD)**
 - Subject to any appeal by consulters, the FAD will form the Institute's guidance on the use of the appraised technology.

NICE Sunitinib GIST Appraisal Timeline

- Pfizer submission deadline – October 31, 2008
- First appraisal committee meeting – February 11, 2009
 - ACD released on March 2009: *“The Committee is minded not to recommend sunitinib”*
- Second appraisal committee meeting – April 8, 2009
 - Final Appraisal Determination (FAD) released on May 2009: *“Sunitinib is recommended as a treatment option for people with unresectable and/or metastatic malignant gastrointestinal stromal tumours*”