Adjusting the Crossover Effect in Overall Survival Analysis Using a Rank Preserving Structural Failure Time Model:

The Case of Sunitinib GIST Trial

Xin Huang[1] and Qiang (Casey) Xu[2]

[1]Pfizer Oncology Statistics, [2]Columbia University

Outline

- Endpoints in Oncology Trials and Challenges
- The Case of sunitinib GIST Trial
- Crossover Issues in Oncology Clinical Trials
- Apply Rank Preserving Structural Failure Time (RPSFT) Model to sunitinib GIST Case
- Conclusions

Uniqueness of Oncology Trials

The design and conduct of oncology trials are typically more complicated than trials for other diseases

- Life threatening disease
  - Change of treatment upon disease progression
- Non-randomized single arm study
- Single study
- Active randomized controlled trial (RCT)
- Multi-center, multi-national, co-operative group study

Types of Endpoints in Oncology

- Solid tumors
  - Metastatic Disease:
    - Tumor response rate, Time to tumor progression, Progression free survival, Patient Reported Outcomes (PRO), Overall survival (OS)
  - Adjuvant Setting:
    - Time to recurrence, Disease free survival, PRO, OS
- Hematological malignancies
  - Complete remission, Time to recurrence, Recurrence free survival, PRO, OS

Overall Survival (OS)

- OS is defined as the time from randomization until death from any cause, and is measured in the intent-to-treat population
- OS has long been the gold standard representing clinical benefit in oncology
- Survival is considered the most reliable cancer endpoint, and when studies can be conducted to adequately assess survival, it is usually the preferred endpoint
- Bias is not a factor in endpoint measurement
- OS should be evaluated in randomized controlled studies (RCT)

Objective Response Rate (ORR)

- ORR is defined as the proportion of patients with tumor size reduction of a predefined amount and for a minimum time period
- Response duration usually is measured from the time of initial response until documented tumor progression
- Generally, ORR is defined as the sum of partial responses (PR) plus complete responses (CR)
- ORR is a direct measure of drug antitumor activity, which can be evaluated in a single-arm study
- Standardized criteria such as RECIST should be used to ascertain response

Time to Tumor Progression (TTP)

- Time to Tumor Progression (TTP) is defined as the time from randomization until objective tumor progression
- TTP has served as primary endpoint for drug approval
- TTP does not include deaths
- In TTP analysis, deaths are censored, either at the time of death or at an earlier visit representing informative censoring
- The precise definition of tumor progression is important and should be carefully detailed in the protocol

Progression Free Survival (PFS)

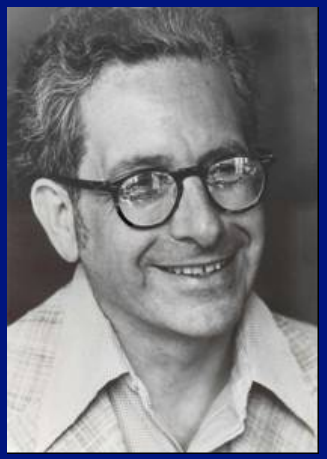
- PFS is defined as the time from the date of randomization to the date of the first documentation of progression or death due to any cause, whichever occurs first
- Compared with TTP, PFS is the preferred regulatory endpoint
- PFS includes deaths and thus can be a better correlate to overall survival
- PFS can reflect tumor growth and be assessed before the determination of a survival benefit
- Its determination is not confounded by subsequent therapy
- However, the formal validation of PFS as a surrogate for survival for the many different malignancies can be difficult

OS, TTP, and PFS Challenges

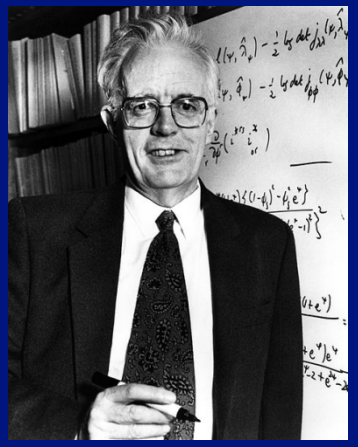
- **OS challenge** – discontinuation/switch in therapies (not randomized) after adverse event/progression (crossover effect)
  - True treatment effect?
- **TTP and PFS Challenges** – can progression be measured reliably?
  - Difference in treatment regimen/schedule between treatment arms – open label
  - Difference in frequency/assessment times between treatment arms – biased estimates, false positive
  - Disagreement between INV and IRC
  - Change of therapy after INV progression assessment
  - Scans not readable
  - Missed schedule
  - Measure some lesions - not all identified target lesions
  - Not all scans available for review by IRC (lagging btw. INV and IRC)

Statistical Methods in Oncology

- Kaplan-Meier estimator
  - an estimator for estimating the survival function from time-to-event data
- Log-rank test
  - a hypothesis test to compare the survival distributions of two samples
- Cox proportional hazards model
  - if the proportional hazards assumption holds then it is possible to estimate the effect parameter(s) without any consideration of the hazard function



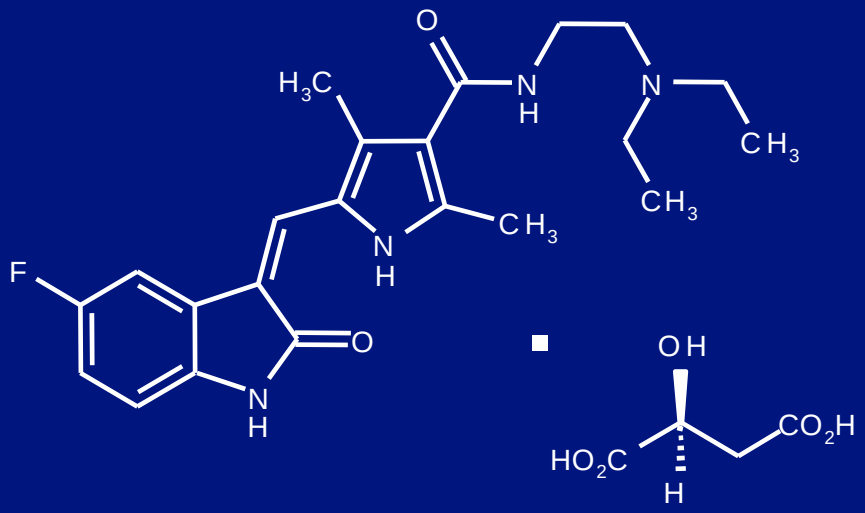
Dr. Paul Meier (1924-2011)



Sir David Cox

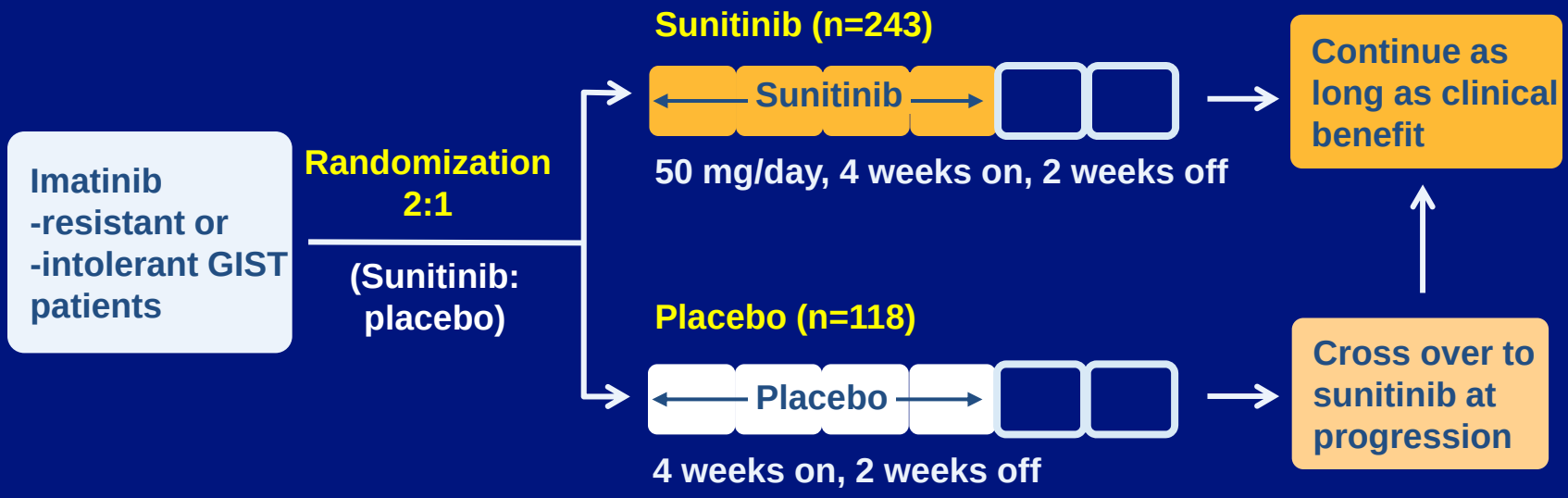
Sutent® (*sunitinib malate*)

- 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

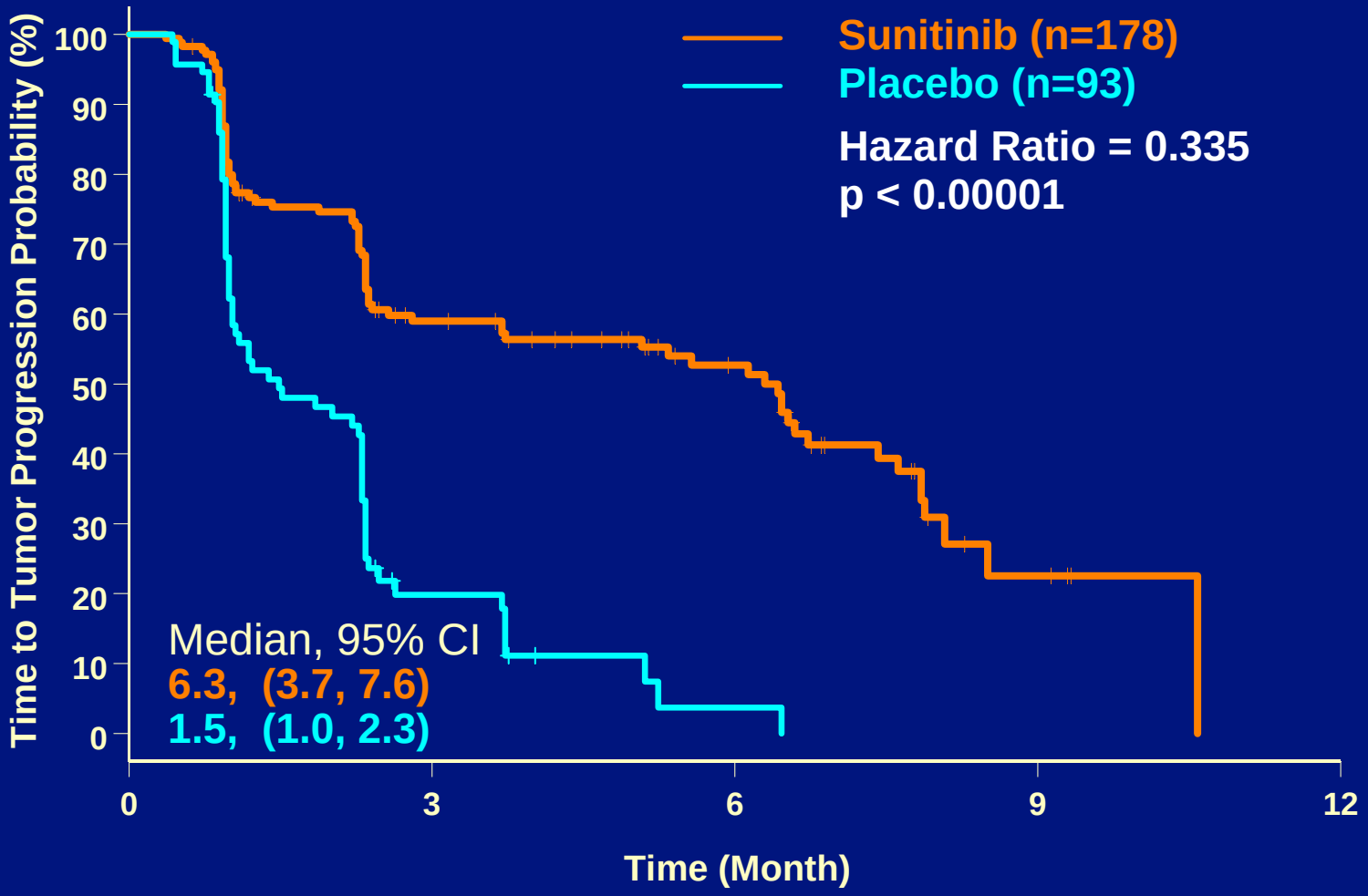
Study Objectives

- Primary
  - TTP (assessed by IRC 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

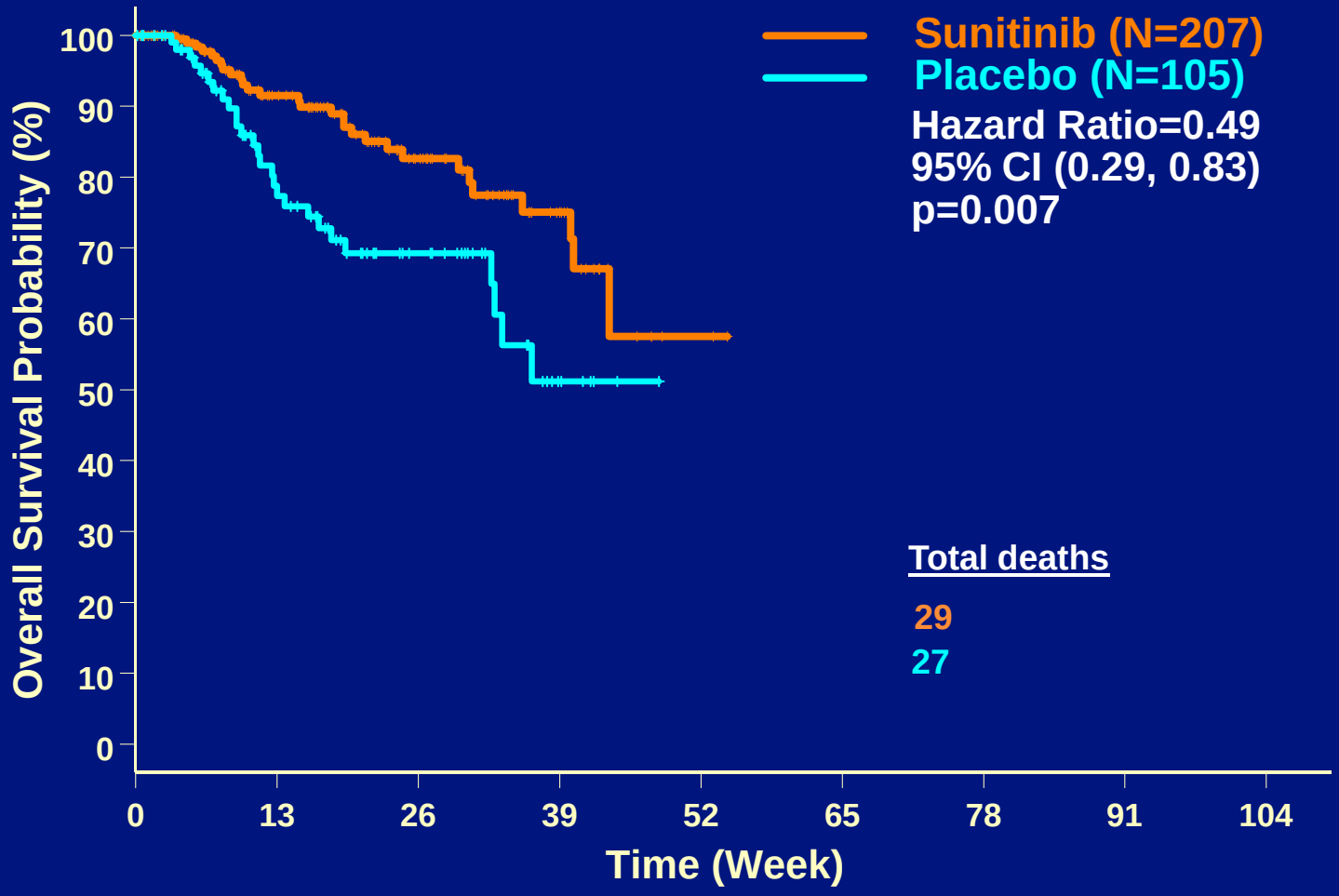
Results from the Primary Endpoint

- 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 (IDMC) 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

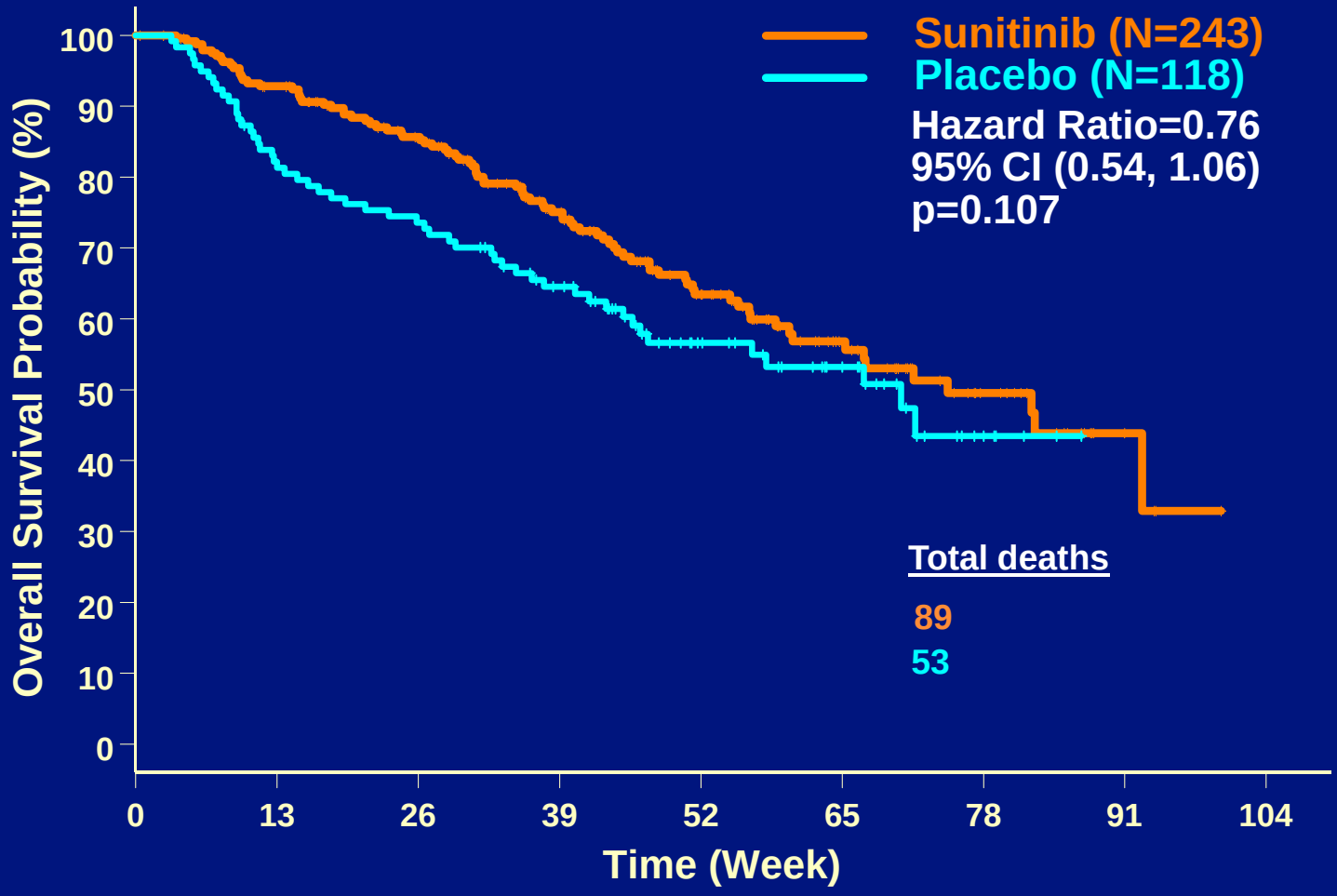
Time to Tumor Progression
(Interim Analysis Based on IRC, 2005)



Overall Survival (NDA, 2005)

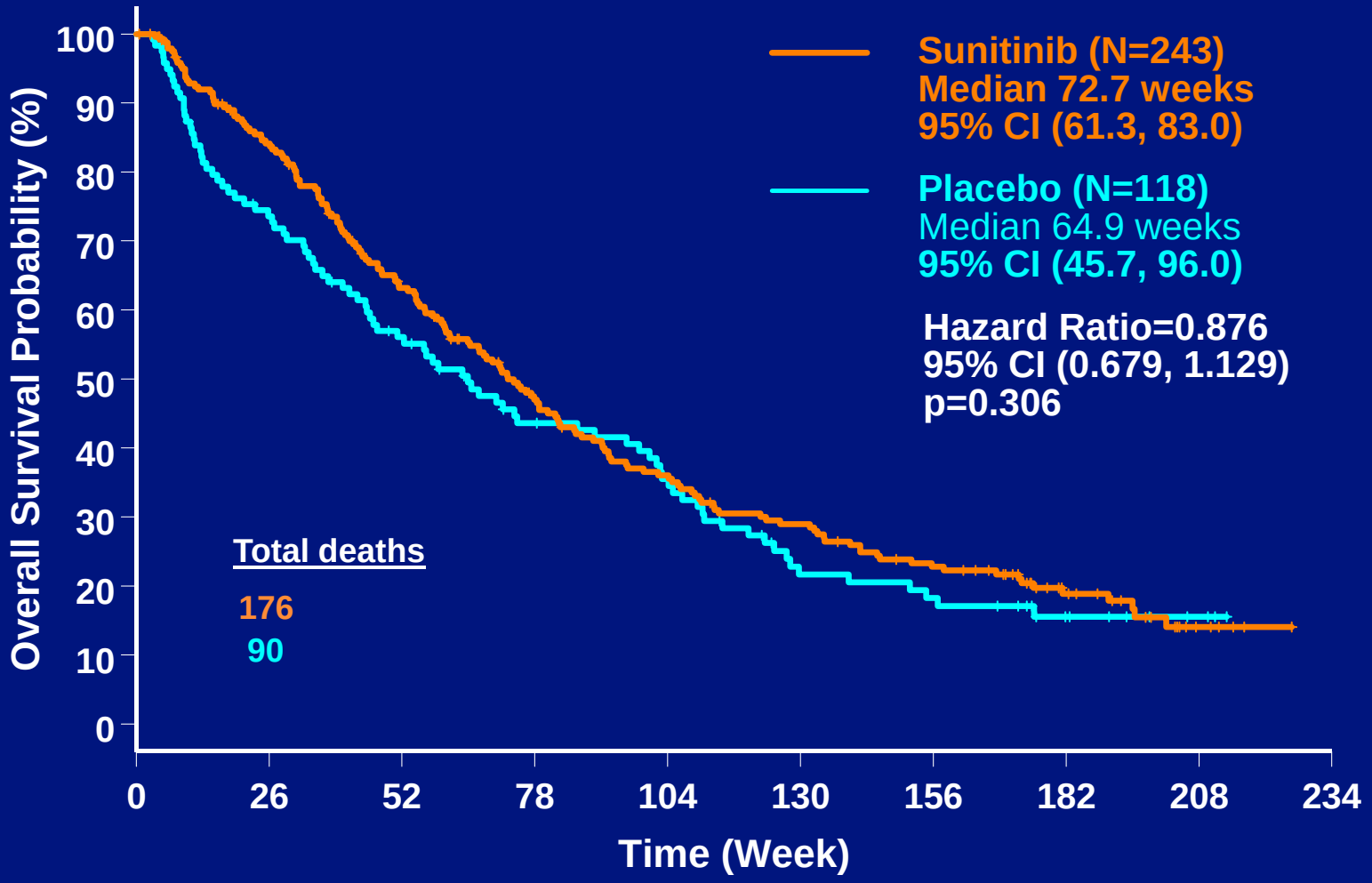


Overall Survival (ASCO, 2006)



nRisk Sutent	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

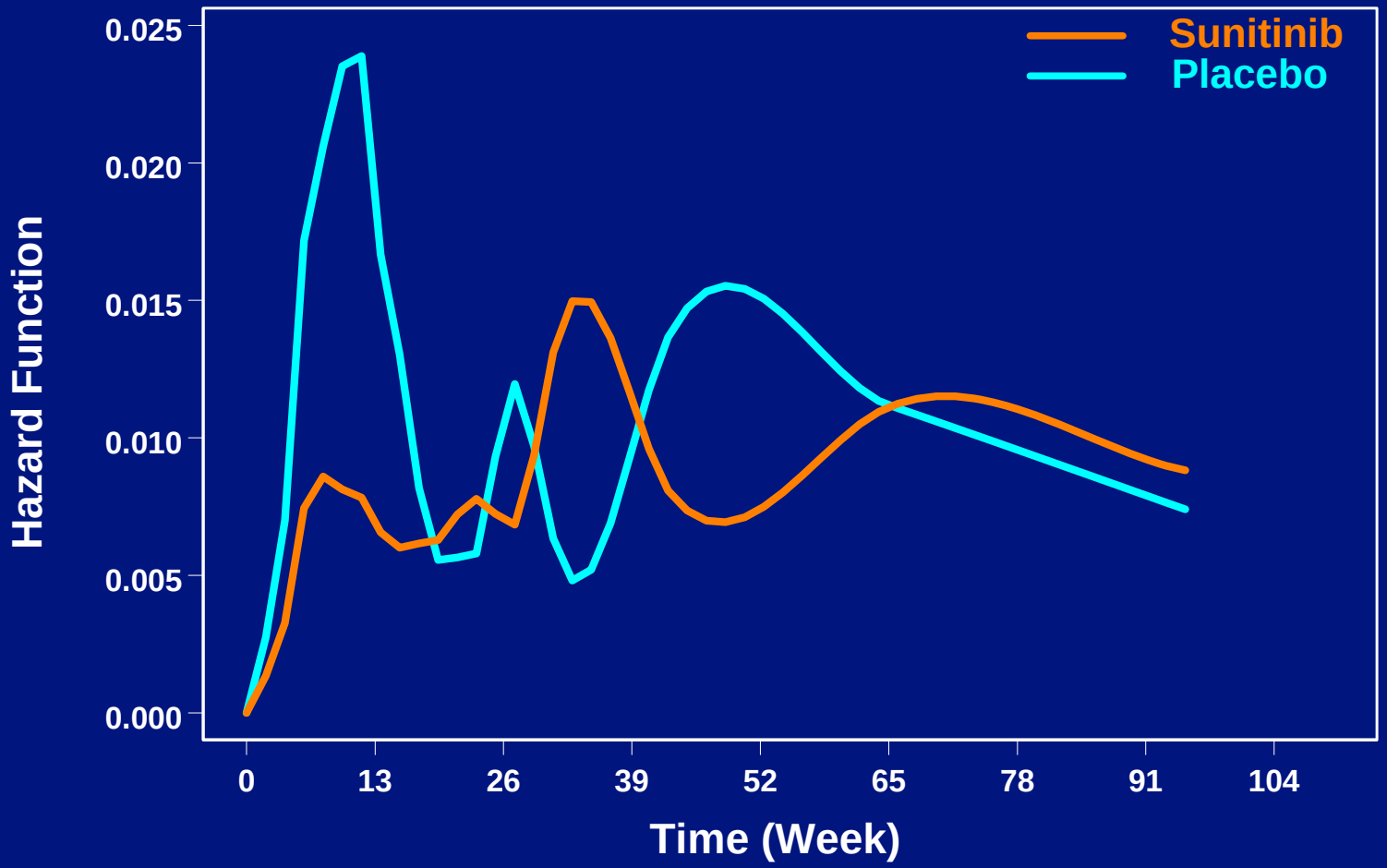
Overall Survival (Final, 2008)



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

Estimated Hazard Functions by Treatment



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

Conventional Analyses

Cox Model	Hazard Ratio (SU/PB), 95% CI and p-value
ITT (naïve)	**0.876 (0.679, 1.129), p=0.306**
ITT (dropping crossover)	0.315 (0.178, 0.555), p<0.0001
ITT (censoring at crossover)	0.825 (0.454, 1.499), p=0.527
ITT (time-dependent treatment)	0.934 (0.520, 1.679), p=0.820

Crossover Issues in Oncology Trials

- Typical oncology trial designs compromise randomization for the overall survival analysis
  - In placebo controlled randomized trials
    - Experimental (E) vs. Placebo (P): P $\rightarrow$ E
  - In active controlled randomized trials, if crossover is allowed
    - E + Standard of Care (S) vs. S: S $\rightarrow$ E + S
    - E vs. S: E $\rightarrow$ S and S $\rightarrow$ E
  - In active controlled randomized trials, if crossover is not allowed
    - E vs. S: E $\rightarrow$ NST and S $\rightarrow$ NST
  - A mix of crossover between study treatments and switch to non-study treatments (NST)

Impact of Crossover to OS Analysis

- A comparison of arms for the treatment effect cannot be made as randomized because of crossover
- 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

Not Have a Treatment Effect on OS

- Higher hurdle to obtain the regulatory approval
  - Bevacizumab breast cancer indication
- Greater challenge to be accepted by the payers
  - The access and reimbursement environment is becoming more challenging as public and private payers scrutinize the value and cost of pharmaceuticals more than ever.
  - Increasingly, robust proof of clinical and/or economic value is being required as a condition of coverage.
  - UK National Institute for Health and Clinical Excellence (NICE) cost effectiveness analyses

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

Standard Approaches to Analyze OS

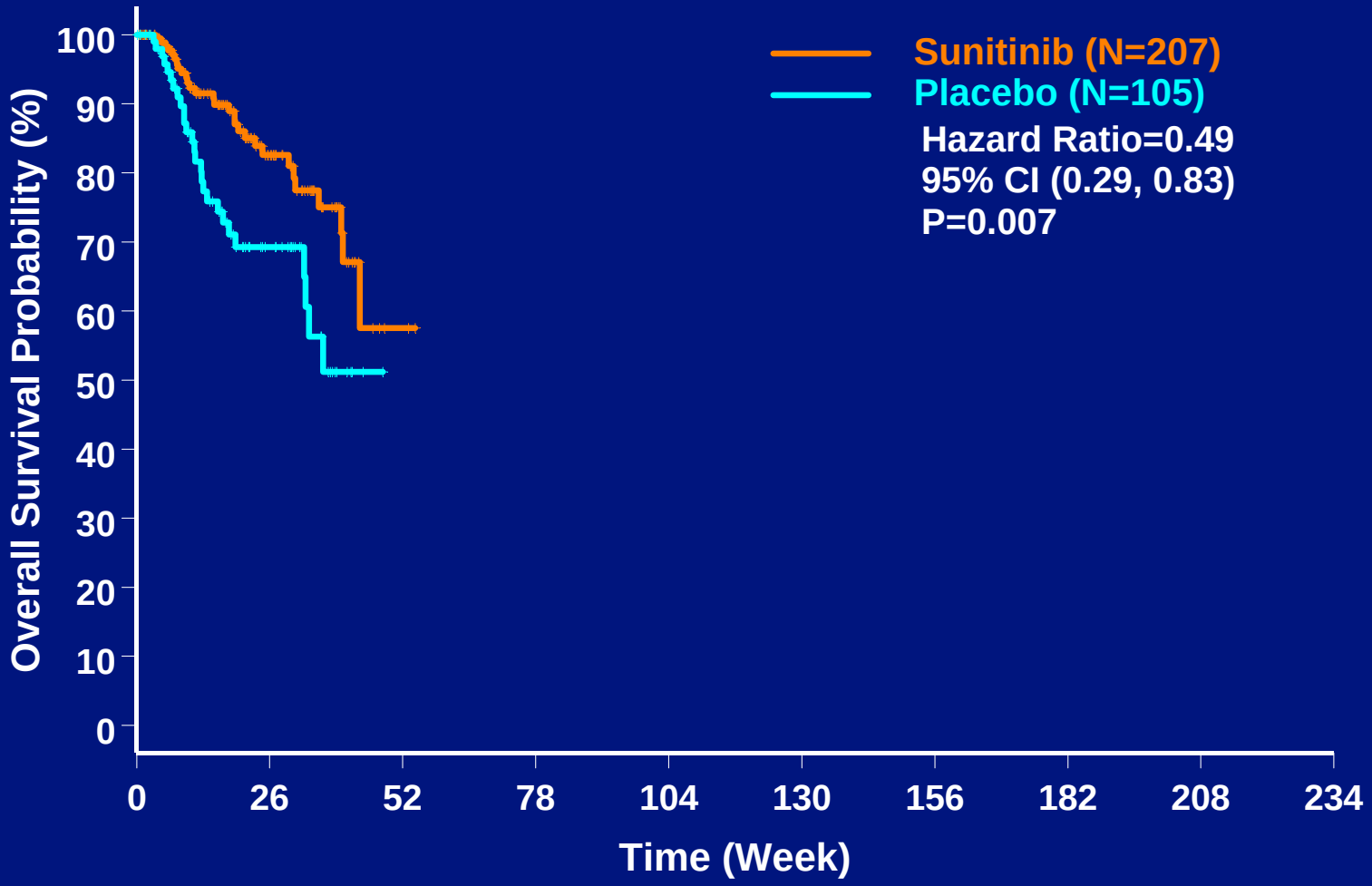
- ITT Analysis
  - Analyze patients as randomized and makes no adjustment for crossover
- Per-protocol Analysis
  - Drop patients who cross over
- On-Treatment Analysis
  - Censor patients when they cross over
- Time Dependent Treatment Analysis using the Cox model

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

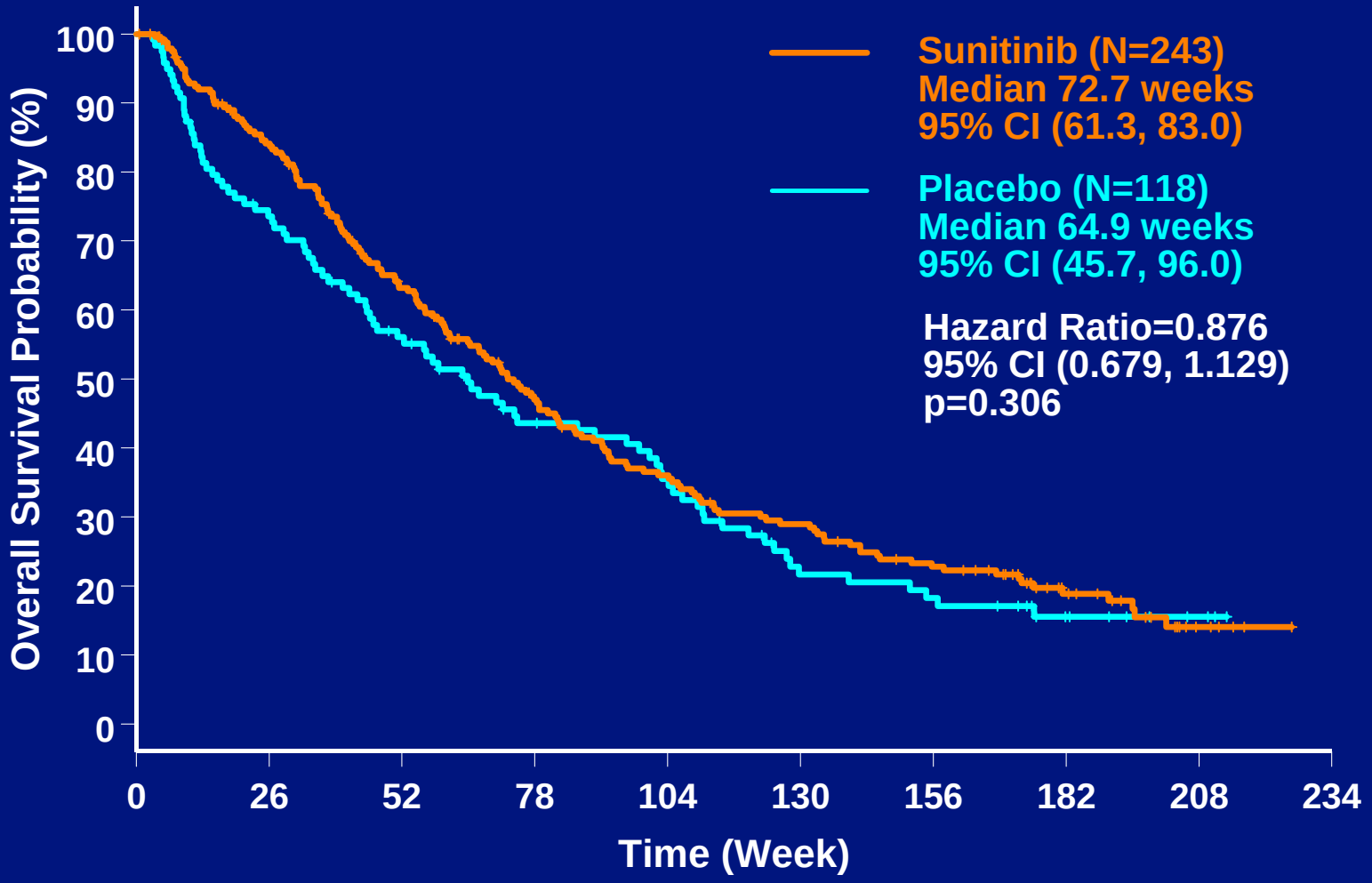
A Novel Solution to the Crossover Issue

- Prefer intent-to-treat (ITT) analysis and desire to compare treatment groups as randomized
- Estimate the unbiased 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) may be used to remove bias induced by crossover

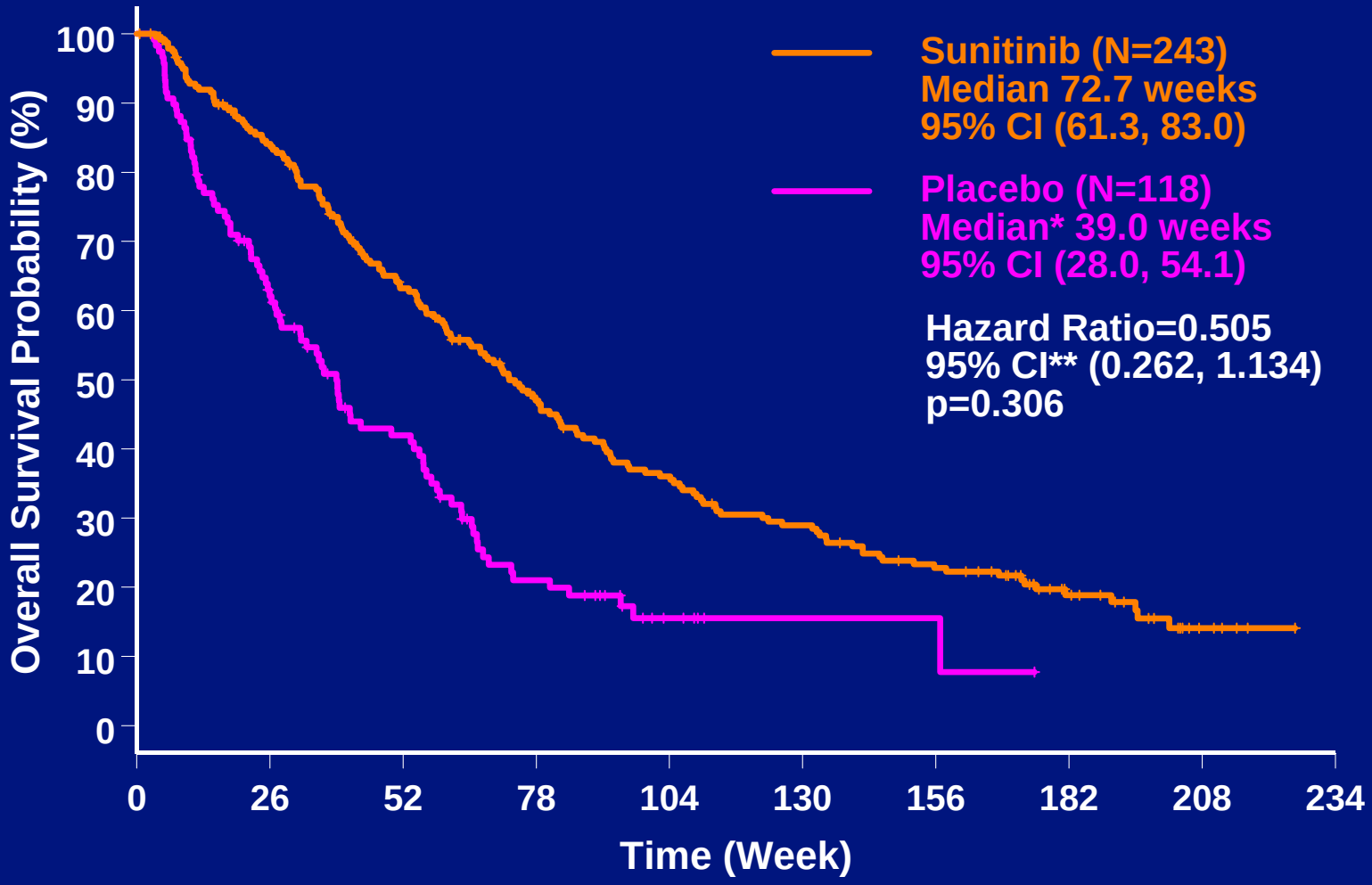
Overall Survival (NDA, 2005)



Overall Survival (Final, 2008)

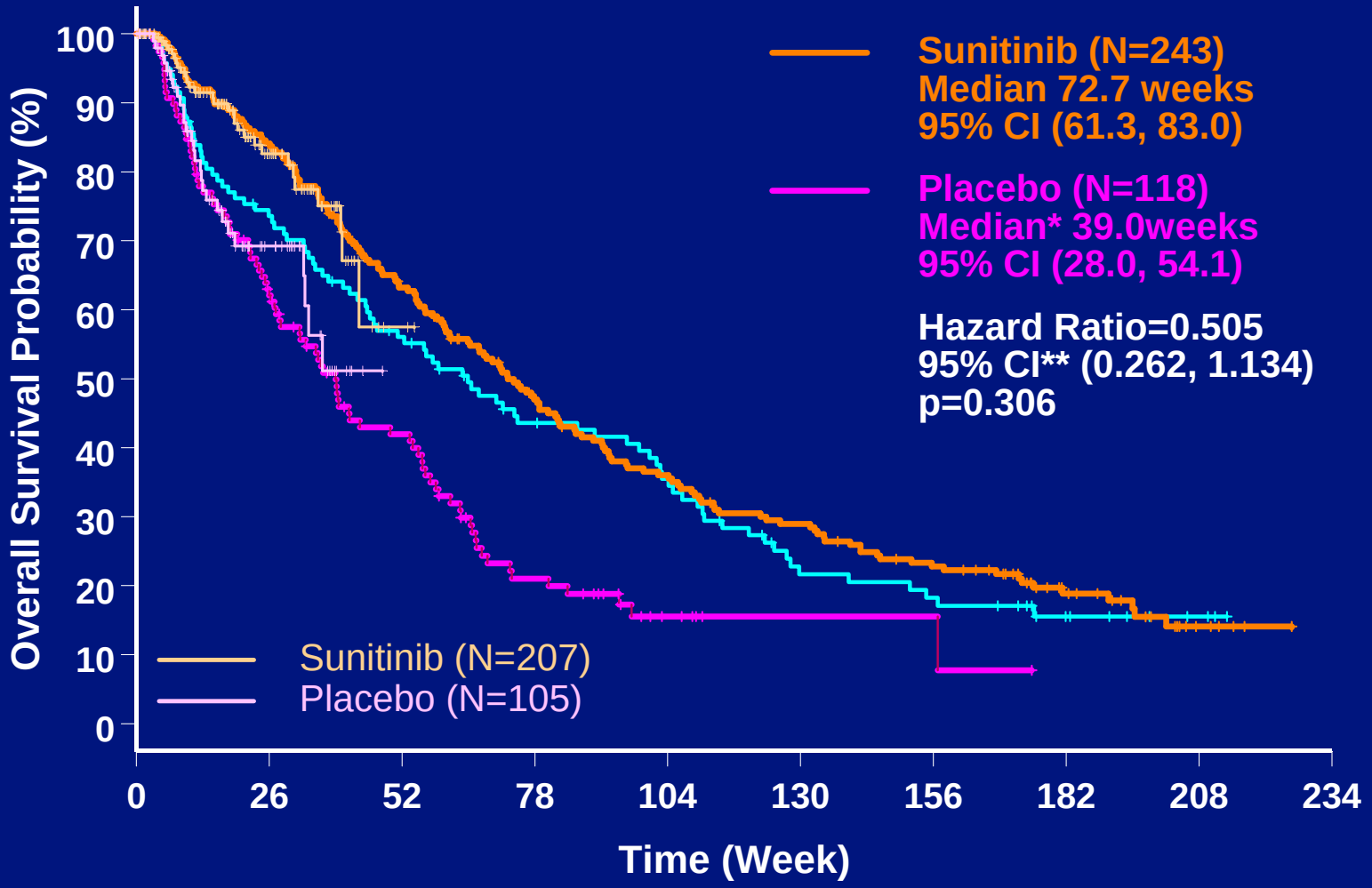


Overall Survival (Final, 2008)
Crossover Adjusted by RPSFT



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

Overall Survival (Final, 2008)
Crossover Adjusted by RPSFT



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

Overall Survival Results

	Median OS (weeks; 95% CI)			
	Sunitinib	**Placebo**	**HR (95% CI)**	**P-value**
Interim (blinded phase) based on ITT	Not reached	Not reached	0.491 (0.290-0831)	0.007
Final (blinded + open label) based on ITT	72.7 (61.3-83.0)	64.9 (45.7-96.0)	0.876 (0.679-1.129)	0.306
Final (blinded + open label) based on ITT using RPSFT Method	72.7 (61.3-83.0)	39.2 (28.0-54.1)	0.505 (0.262-1.134)	0.306

The End of the Story

- The final GIST OS result was 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 ……"

Other Applications of RPSFT in Cost Effectiveness Analyses

Therapy and Tumor Type	Method
Sunitinib for the second-line treatment of GIST (2009)	**RPSFT**
Pazopanib for the first-line treatment of RCC (2010)	**RPSFT** **IPCW***
Everolimus for the second-line treatment of RCC (2010)	**RPSFT** **IPCW**
Trastuzumab + anastrozole for postmenopausal women with HER2+ and HR+ BC (2010)	**RPSFT**
Letrozole and anastrozole vs. tamoxifen as adjuvant therapy in postmenopausal women with early BC (2011)	**RPSFT** **IPCW**

***IPCW – Inverse probability of censoring weighting methods**

Conclusions

- The benefit demonstrated in TTP in sunitinib GIST trial is likely to translate into an OS benefit in Imatinib-resistant or intolerant GIST patients
- 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

References

- Robins JM, and Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics*. - Theory and Methods 1991; 20(8): 2609-2631.
- Robins J M, Mark SD, and Newey WK. Estimating exposure effect by modeling the expectation of exposure conditional on confounders. *Biometrics* 1992; 48: 279-495.
- Mark SD, and Robins JM. Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. *Statistics in Medicine* 1993; 12: 1605-1628.
- White IR, Babiker AG, Walker S, Darbyshire JH. Randomization-based methods for correcting for treatment changes: examples from the Concorde trial. Statistics in Medicine 1999; 18:2617-2634.
- Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. 2007