Genomic Status in Optimizing Subgroup Patients Selection: a Case Study

Jean-Marie Ledeine

Principal Biostatistician, Bristol-Myers Squibb

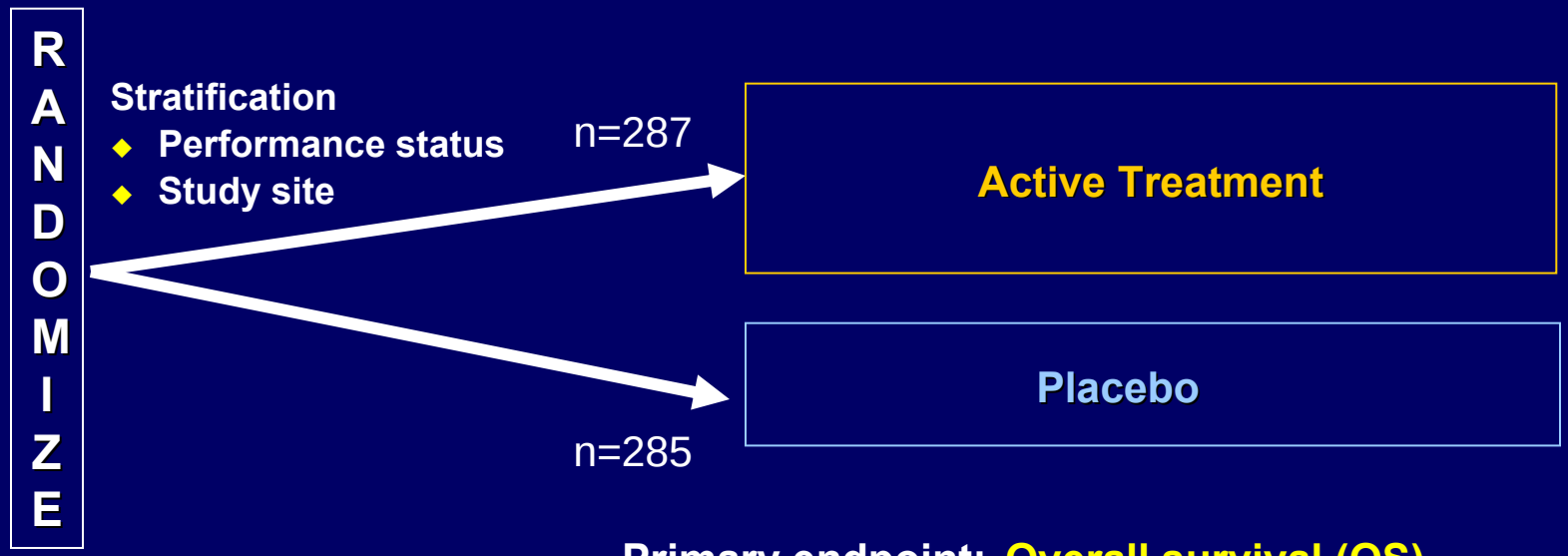
November 30, 2012

Topics

- Original successful submission
- Discovery of a genomic marker (after submission)
- Application of retrospective subset analysis
- Sensitivity analyses (missing genomic marker status)

Original Phase 3 Study - Study Design

Open label study (N=572)

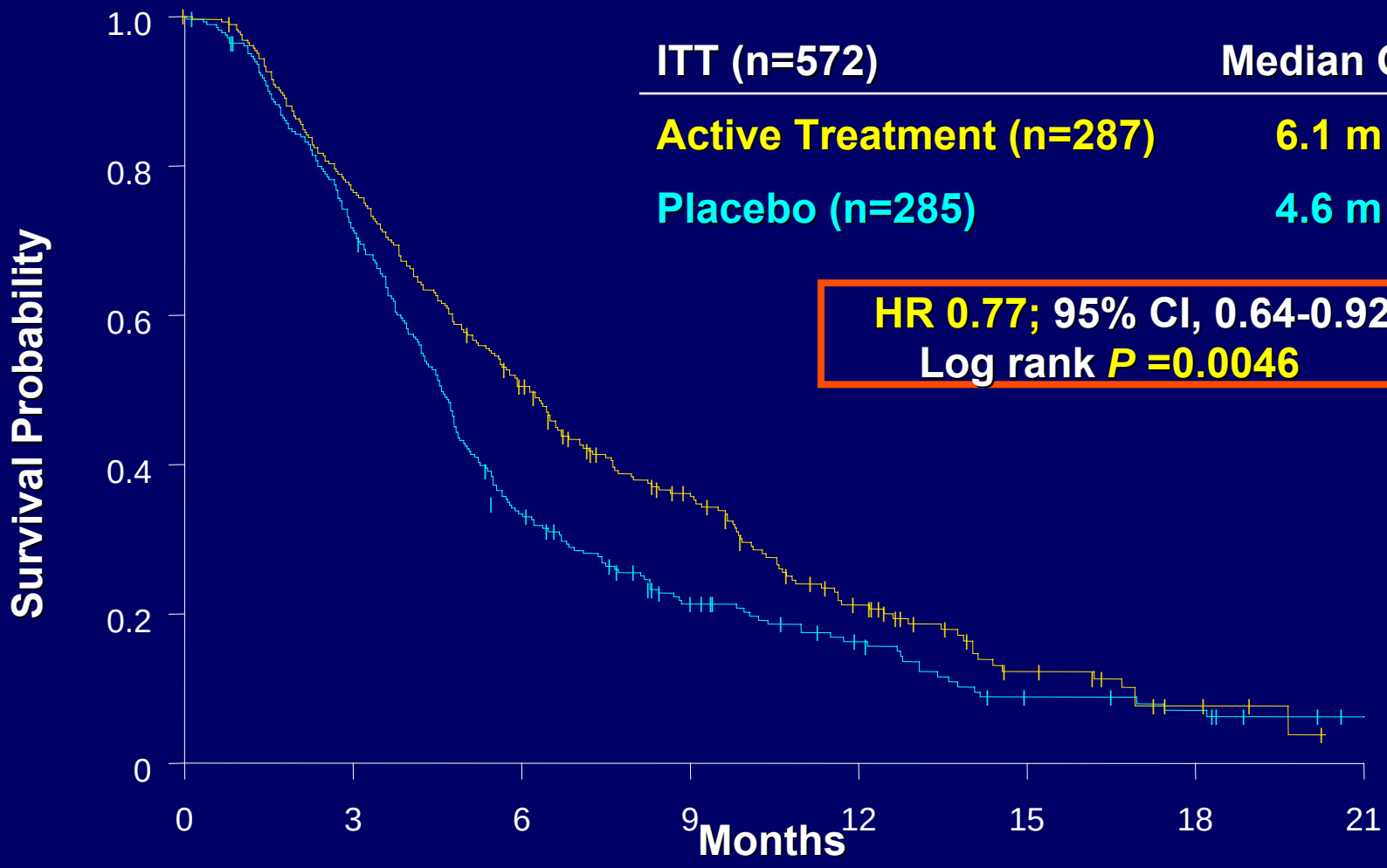


Primary endpoint: Overall survival (OS)

No planned use of any genomic marker

Original Phase 3 Study - Final Results

Primary Endpoint of OS



ITT (n=572)	Median OS
Active Treatment (n=287)	6.1 m
Placebo (n=285)	4.6 m

HR 0.77; 95% CI, 0.64-0.92
Log rank *P* =0.0046

Results of this study supported marketing approval in 2007

Genomic Marker - Emergence of Data

- Increasing evidence that a genomic marker predicts response to this study drug
  - Publications / Presentations
  - Also related to other drugs within same class
- No or minimal response to these drugs in mutation-positive (M+) patients (predictive marker)
- All these data resulted in changes in daily practices
  - Clinicians no longer prescribed drugs to M+ patients
  - US payers and guidance documents evaluated requirements for genomic status testing for treatment decisions

Retrospective Subset Analysis General

- None of these changes were based on any prospective study
- Because of the shift in practice a prospective study was not possible anymore
  - Only "retrospective" subset analyses of completed studies could be done
- FDA defined the basis for "prospective-retrospective" analyses that could address this situation (Advisory Committee meeting in December 2008)

Retrospective Subset Analysis Application

- FDA was approached about using original phase 3 study for a retrospective subset analysis

- FDA accepted because study was:
  - Positive (not a mechanism to salvage a failed trial)
  - Adequate, well-controlled, well conducted
  - Large enough

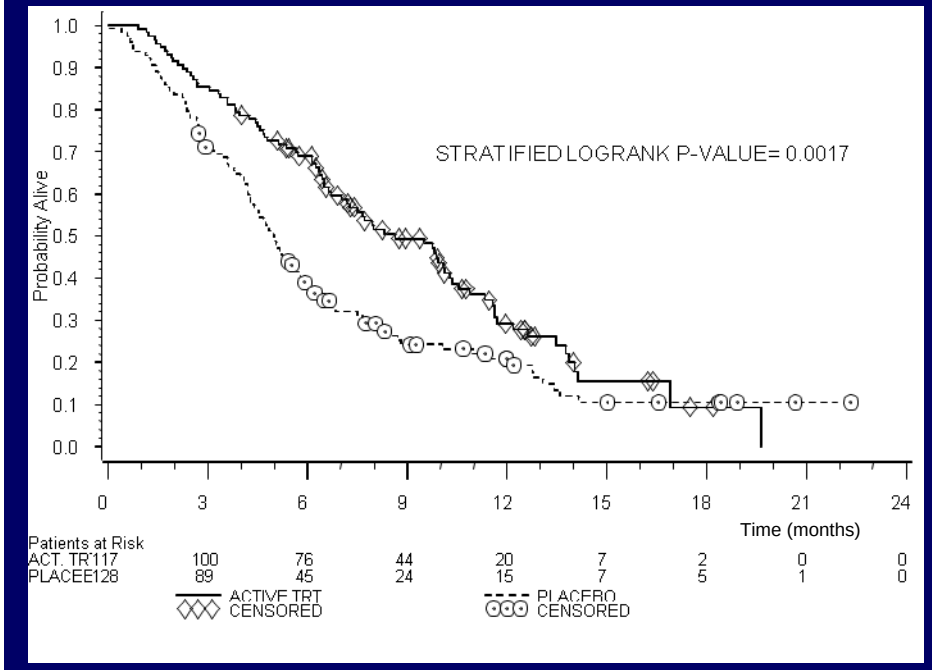
- However FDA asked:
  - To have genomic status for ≥ 90% of subjects
  - To review the SAP in a way that all analyses were as prospectively planned as possible
  - To use a validated assay

Analyses Related to Genomic Status

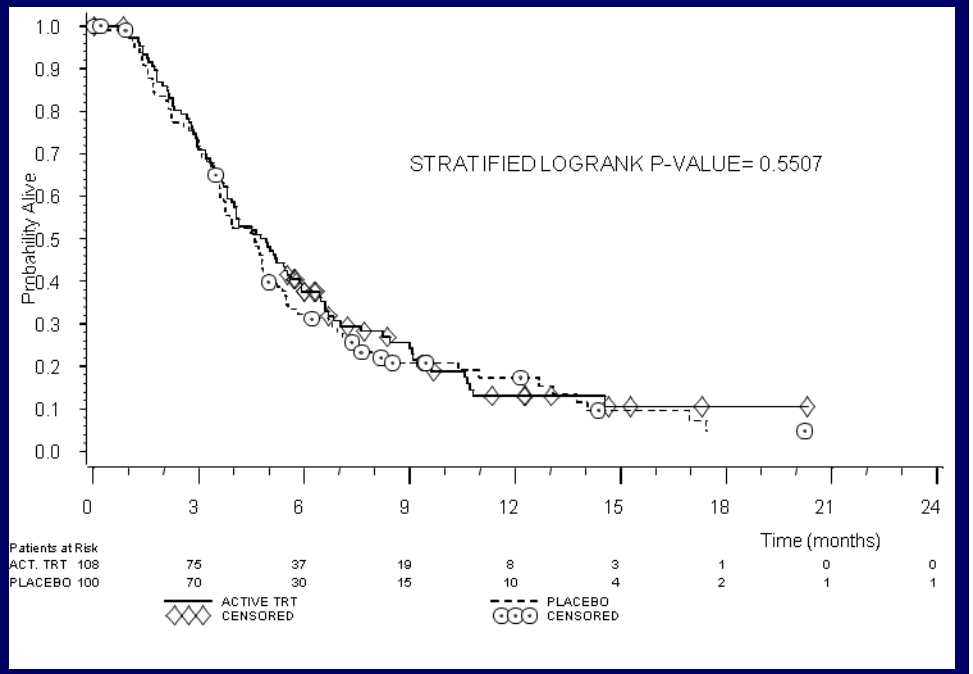
- Demonstrate association between genomic status and treatment
  - Interaction between genomic status and treatment
    - COX model with treatment, genomic status and interaction
    - Challenge: significance level interaction test (0.05, 0.10 or 0.20?)
  - Comparisons between treatments within subsets (M- and M+)
    - HR and log-rank p-values

Primary Efficacy Analysis Results
OS KM Curves

M-



M+



Primary Efficacy Analysis Results (OS)

	M-		*M+*	
	Active Treatment (n = 117)	Placebo (n = 128)	Active Treatment (n = 108)	Placebo (n = 100)
Median OS (months) (95% CI)	8.6 (7.0, 10.3)	5.0 (4.3, 5.7)	4.8 (3.9, 5.6)	4.6 (3.6, 4.9)
Hazard ratio (95% CI)	0.63 (0.47, 0.84)		0.91 (0.67, 1.24)	
Log-rank p-value	0.0017		0.5507	
Interaction p-value	0.0703			

OS benefit of drug is observed in the *M-* population only and the p-value for the interaction between genomic status & treatment is low

Sensitivity Analyses

- 453/572 (79%) subjects with genomic marker evaluation
- Raised concerns about validity and robustness of primary findings in observed data population

- To address these concerns:
  - Scrutinize reasons for missing genomic marker
  - Side-by-side tabulations of baseline characteristics in genomic marker evaluated vs. not evaluated
  - Perform sensitivity analyses on OS using different techniques of imputing missing data
    - Multiple imputation (assumes MAR)
      - Applied to missing covariate data
    - Worst/best case scenario

Sensitivity Analyses
Multiple Imputation - Challenges (1 of 2)

- Predefining the variables to be used
  - Variables that may be informative of genomic status or missingness
  - First model including :
    - Patient characteristics (age, gender, countries, sites)
    - Disease characteristics (tumor stage, time since diagnosis)
    - Tumor sample characteristics (type of tumor)
    - Sample handling methods (macro-dissection)
  - Second model also including:
    - OS time / treatment arms… but we want to show a difference between treatments based on OS
  - Large number of variables (model with ~ 70 covariates)
    - Validity of the model?
    - Should we apply a model selection first (e.g., stepwise)?

Sensitivity Analyses
Multiple Imputation - Challenges (2 of 2)

- Technical implementation
  - Variables had different distributions
    - Binomial, categorical, normal, etc.
  - Some pre-predefined variables also presented with missing data
    - Going in two steps?
      - » Imputing missing data for variables in the model first
      - » Imputing missing genomic status data
    - Going in one step
      - » Using the MCMC option of PROC MI
  - Number of imputations
    - Between 5 and 10?
    - Several hundred?
    - Using the relative efficiency (> 99%)

Multiple Imputations Results

	M-	*M+*	
	Hazard ratio (95% CI) (1)	Hazard ratio (95% CI) (1)	Interaction p-value (2)
1ary Analysis (N=453)	0.63 (0.47, 0.84)	0.91 (0.67, 1.24)	0.0703
MI (1ary model, N=572)	0.67 (0.51, 0.87)	0.88 (0.66, 1.16)	0.1822
MI (+ OS time & OS flag)	0.67 (0.52, 0.87)	0.89 (0.67, 1.18)	0.1767
MI (above + trt flag)	0.66 (0.51, 0.86)	0.89 (0.67, 1.19)	0.1357

(1) Hazard ratio of active treatment over placebo, used a stratified Cox model with treatment as unique factor.
(2) Student's t test; degree of freedom as per Rubin (1987). Used a stratified Cox model with treatment, genomic status and interaction.
Note: 13 to 16 MIs were performed to achieve 99% relative efficiency

- **Similar results obtained using other:**
  - **Number of imputations**
  - **Technical implementations than the MCMC**

Sensitivity Analyses
Deterministic Scenarios

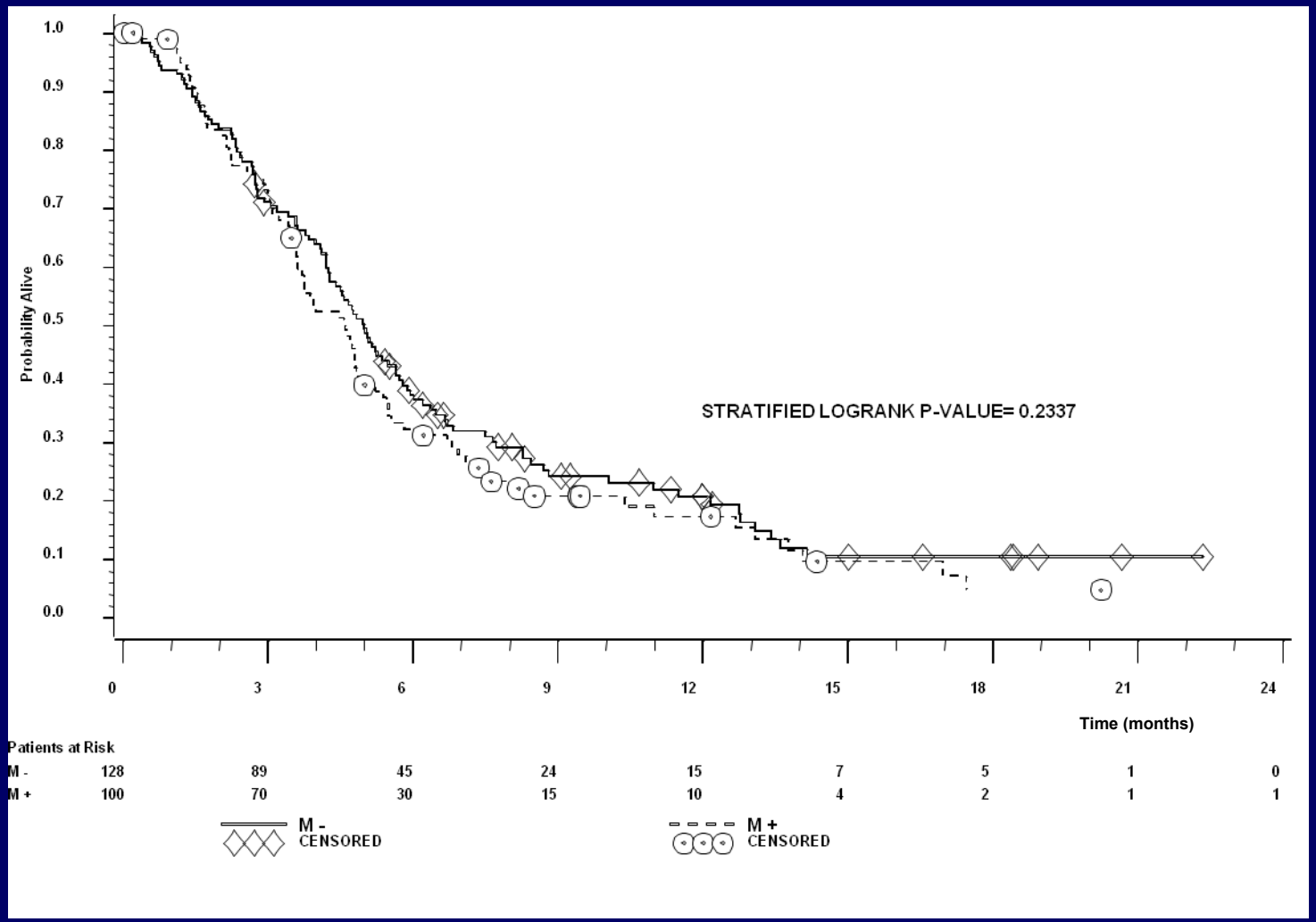
- Four deterministic scenarios pre-defined
  - A) M- for all missing data
  - B) M- for all subjects who died, M+ otherwise
  - C) M- if the OS was “short” , M+ otherwise
  - D)
    - In the active treatment arm
      - M- if the OS was “short”, M+ otherwise;
    - in the placebo arm
      - M- if the OS was “long”, M+ otherwise
- Reverse situations (+ 4 cases)
- Mixture of worst/best cases
- Challenges: definition of “short” (obs. proportion of M-)

Worst/Best Case Scenario Results

	M-	*M+*	
Scenarios	Hazard ratio (95% CI)	Hazard ratio (95% CI)	Interaction p-value
A	0.68 (0.54, 0.87)	0.91 (0.67, 1.24)	0.1091
A reversed	0.63 (0.47, 0.84)	0.87 (0.68, 1.10)	0.1037
B	0.70 (0.55, 0.88)	0.88 (0.65, 1.19)	0.2420
B reversed	0.61 (0.45, 0.81)	0.90 (0.70, 1.14)	0.0454
C	0.72 (0.56, 0.92)	0.83 (0.63, 1.08)	0.3572
C reversed	0.60 (0.46, 0.79)	0.96 (0.74, 1.25)	0.0162
D	0.82 (0.63, 1.05)	0.66 (0.51, 0.87)	0.2924
D reversed	0.53 (0.41, 0.69)	1.22 (0.93, 1.60)	<0.0001

D) In the active treatment arm, M- if the OS was "short", M+ otherwise; in the placebo arm M- if the OS was "long", M+ otherwise

Genomic marker prognostic?
In Placebo Arm OS by *Genomic* Status



Conclusions

- Marker discovered late in development generated challenges:
  - Application of 'Prospective-Retrospective' analysis
  - Address missing genomic marker data
- Experimental agent:
  - Benefit only observed in the M- population
  - Genomic marker predictive of outcome (not prognostic)
  - Primary OS results supported by various sensitivity analyses
  - Consistent effect observed across secondary endpoints

Questions