Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Stone RM, Mandrekar SJ, Sanford BL, et al. Midostaurin plus chemotherapy for acute myeloid leukemia with a *FLT3* mutation. N Engl J Med 2017;377:454-64. DOI: 10.1056/NEJMoa1614359

Supplement Table of Contents **page**

1. Midostaurin/PlacebioDose Modification

DOSE MODIFICATIONS AND MANAGEMENT OF TOXICITY

Should unanticipated circumstances arise that might require minor variances from the prescribed dosing and schedule of the protocol therapy or recommended supportive care in order to ensure safety and allow patients to continue to receive treatment on study, the Study Chair should be contacted in advance for discussion and approval.

9.0 Midostaurin/Placebo Dose Modifications

9.0.1 Midostaurin/Placebo Induction & Consolidation Therapy Dose Modifications for Hematologic Toxicity

There will be no dose modifications for hematologic toxicity due to midostaurin/placebo during induction and consolidation therapy.

9.0.2 Midostaurin/Placebo Induction & Consolidation Therapy Dose Modifications for Non-Hematologic Toxicity

9.0.2.1 There will be no dose modifications for any grade 1 or 2 non-hematologic toxicity.

9.0.2.2 Pulmonary Toxicity

- **For ≥ grade 3 pulmonary infiltrate**, interrupt midostaurin/placebo for the remainder of the cycle. Resume midostaurin/placebo at the same dose when infiltrate resolves to ≤ grade 1.
- **Missed doses of midostaurin/placebo will not be made up.**

9.0.2.3 Cardiac Toxicity

- **For QTc interval > 450 msecs and ≤ 470 msecs,** check magnesium and potassium levels and correct any abnormalities. If possible, stop any medications that may prolong the QTc interval. Continue midostaurin/placebo at the same dose.
- **For QTc interval > 470 msecs and ≤ 500 msecs,** check magnesium and potassium levels and correct any abnormalities. If possible, stop any medications that may prolong the QTc interval. Decrease midostaurin/placebo to 50 mg once daily for the remainder of the cycle. Resume midostaurin/placebo at the initial dose in the next cycle provided that QTc interval improves to ≤ 470 msecs at the start of that cycle. Otherwise continue midostaurin/placebo 50 mg once daily.
- **For QTc interval > 500 msecs,** check magnesium and potassium levels and correct any abnormalities. Hold or interrupt midostaurin/placebo for the remainder of the cycle, and, if possible, stop any medications that may prolong the QTc interval. If QTc improves to ≤ 470 msecs just prior to the next cycle, resume midostaurin/placebo at the initial dose. If QTc interval is not improved in time to start the next cycle do not administer midostaurin/placebo during that cycle. Midostaurin/placebo may be held for as many cycles as necessary until QTc improves.

- **Missed doses of midostaurin/placebo will not be made up.**

9.0.2.4 Other Non Hematologic Toxicity

If a patient experiences other grade 3/4 non-hematologic toxicity considered at least possibly related to midostaurin/placebo, the midostaurin/placebo will be interrupted until toxicity resolves to ≤ grade 1. If the toxicity resolves prior to day 21, then restart at same dose to complete current cycle. Missed doses of midostaurin/placebo will not be made up.

9.0.2.5 Missed doses of midostaurin/placebo will not be made up.

9.0.3 Midostaurin/Placebo Continuation Therapy Dose Modifications for Hematologic Toxicity

In the presence of grade 4 neutropenia during continuation therapy, midostaurin/placebo must be held until ANC ≥ 1000/μL. Once ANC ≥ 1000/μL, then resume midostaurin/placebo at the previous dose. If neutropenia persists for more than two weeks, then discontinue midostaurin/placebo protocol therapy.

9.0.4 Midostaurin/Placebo Continuation Therapy Dose Modifications for Non-Hematologic Toxicity

9.0.4.1 Cardiac Toxicity

- **For QTc interval > 450 msecs and ≤ 470 msecs**, check magnesium and potassium levels and correct any abnormalities. If possible, stop any medications that may prolong the QTc interval. Continue midostaurin/placebo at the same dose.
- **For QTc interval > 470 msecs and < 500 msecs**, check magnesium and potassium levels and correct any abnormalities. If possible, stop any medications that may prolong the QTc interval. Decrease midostaurin/placebo to 50 mg once daily for the remainder of the cycle. Resume midostaurin/placebo at the initial dose in the next cycle provided that QTc interval improves to ≤ 470 msecs at the start of that cycle. Otherwise, continue midostaurin/placebo 50 mg once daily.
- **For QTc interval > 500 msecs,** check magnesium and potassium levels and correct any abnormalities. Hold or interrupt midostaurin/placebo dose and, if possible, stop any medications that may prolong the QTc interval. If QTc interval improves to ≤ 470 msecs just prior to start of the next cycle, resume midostaurin/placebo at the initial dose. If QTc is not improved in time to start the next cycle, do not administer midostaurin/placebo during that cycle. Midostaurin/placebo may be held for as many cycles as necessary until QTc improves.
- **Missed doses of midostaurin/placebo will not be made up.**
- Dose modifications for QTc are for the remainder of the cycle.

9.0.4.2 Pulmonary Toxicity

- **For ≥ grade 3 pulmonary infiltrate**, interrupt midostaurin/placebo for the remainder of the cycle. Resume midostaurin/placebo at the same dose when infiltrate resolves to ≤ grade 1.
- **Missed doses of midostaurin/placebo will not be made up.**

9.0.4.3 Other Grade 3/4 Non-Hematologic Toxicity

- **For other grade 3/4 non-hematologic toxicities that are considered to be at least possibly related to midostaurin/placebo**, interrupt midostaurin/placebo. Resume midostaurin/placebo at the same dose when toxicity resolves to ≤ grade 2. If midostaurin/placebo is held for more than 28 days, then discontinue midostaurin/placebo

continuation therapy."

- Missed doses of midostaurin/placebo will not be made up.

9.0.4.4 Other Grade 1/2 Non-Hematologic Toxicity

Persistent grade 1 or 2 toxicity during continuation therapy that patients may deem unacceptable may prompt a drug holiday for as many as 28 days. No drug holidays longer than 28 consecutive days will be allowed. **Missed doses of midostaurin/placebo will not be made up.**

9.1 Daunorubicin Hepatotoxicity Dose Modifications

Initial and subsequent daunorubicin doses should be modified as follows for hepatotoxicity:

Total Bilirubin (mg/dL)	% Daunorubicin Dose to Give
≤ 2	100%
> 2 – ≤ 3.0	75% (25% dose reduction)
> 3.0	50% (50% dose reduction)

For patients with evidence of hepatic dysfunction, reassess regularly during remission induction treatment.

9.2 High-Dose Cytarabine Consolidation Therapy Dose Modifications

9.2.1 Contributions of concomitant medications to neurotoxicity should be assessed and other medications discontinued if possible.

9.2.2 For neurotoxicity ≥ grade 2 due to high-dose cytarabine during consolidation therapy, discontinue high-dose cytarabine for the remainder of the cycle. High-dose cytarabine may be considered at the next consolidation therapy cycle with a dose modification from 3 g/m^2 to 2 gm/m^2 if the toxicity has resolved to ≤ grade 1.

9.2.3 For a second occurrence of neurotoxicity ≥ grade 2, high-dose cytarabine should be permanently discontinued.

9.3 Dose Modification for Obese Patients

There is no clearly documented adverse impact of treatment of obese patients when dosing is performed according to actual body weight. Therefore, **all dosing is to be determined solely by the patient's BSA as calculated from actual weight**. This will eliminate the risk of calculation error and the possible introduction of variability in dose administration. **Failure to use actual body weight in the calculation of drug dosages will be considered a major protocol deviation.** Physicians who are uncomfortable with administering chemotherapy dose based on actual body weight should not enroll obese patients on Alliance protocols

2. *FLT3* testing

Testing for activating *FLT3* mutations (internal tandem duplication [ITD], tyrosine kinase domain [TKD] mutations at codons D835 and I836) was done in one of nine academic laboratories: Dr. Thomas Prior and Dr. Guido Marcucci, Ohio State University, Columbus, US (for CALGB, ECOG, SWOG, NCI, NCCTG); Dr. Konstanze Döhner, Ulm University, Ulm, Germany and Dr. Jürgen Krauter, Hannover Medical School, Hannover, Germany (AMLSG and OSHO); Dr. Francesco LoCoco, Tor Vergata, Rome, Italy (GIMEMA); Dr. Christian Thiede, University of Dresden, Dresden, Germany (SAL), Dr. Joop Jansen, Radboud University of Nijmegen, Nijmegen, The Netherlands (EORTC); Dr. Josep Nomdedeu, Hospital de la Santa Creu I Sant Pau, Barcelona, Spain (CETLAM), Dr. Pascual Bolufer, Hospital Universitario la Fe, Valencia, Spain (PETHEMA); and Dr. Andrew Wei, The Alfred Hospital, Melbourne, Australia (ALSG).

Patients signed the informed consent allowing pre-registration prior to obtaining a diagnostic bone marrow and blood sample. Samples were sent via courier express to one of the above central laboratories. DNA was extracted according to laboratories' standard operating procedures. Results from *FLT3* mutational screen were reported to the investigators within 48 hours of receipt of sample within the laboratory.

FLT3 mutation analysis

FLT3-ITD mutation analysis was performed according to the method described by Thiede et al, with minor modifications. High molecular weight genomic DNA was prepared from mononuclear bone marrow or peripheral blood cells ($5x10^6$ to $1x10^7$) using Qiagen Blood Mini Columns according to the protocol of the manufacturer. DNA was eluted in TE-buffer (10 Mm TRIS, 0.1 mM EDTA, pH 8) and the DNA concentration was assessed by a Nanodrop™ device. PCR conditions were as described previously, the are designed to avoid overamplifcation of the reaction, leading to problems in the accurate quantification. A total of 5 ng of DNA was used in a 50µl PCR reaction. To increase reproducibility, each DNA sample was run in triplicate and PCR products (1µl) were analyzed using Applied Biosystems 3130 or 3730 DNA Analyzers or an equivalent instrument. Collected data were analyzed using the Genscan software v3.0 and GeneMapper v3.7 (Applied Biosystems). *FLT3*-ITD signals were further

analyzed, if a signal longer than the wt *FLT3* signal was detected in all three individual reactions. The areas under the curves were quantified for *FLT3*-ITD and the wild-type allele, respectively. To ensure appropriate sensitivity, one of the signals (*FLT3* wt or a putative ITD signal) had to have an AUC of at least 25000 units. The level of *FLT3*-ITD was expressed as a ratio of the area under the curve for *FLT3*-ITD derived signals (ie. all signals longer than the *FLT3* wt peak) divided by the area under the curve for wild-type *FLT3* signal, a mean was calculated for the three reactions. A sample was defined to be positive for *FLT3*-ITD or *FLT3*-TKD mutations when the mean mutant to wild type ratio was 0.05 or higher. *FLT3*-TKD mutations were analyzed according to the method described by Murphy et al. 2003[1]. Briefly, DNA was amplified and an *Eco* RV digest of the PCR product was performed. The digest was further analyzed using capillary electrophoresis as described above. The same criteria for positivity of sample as described for the ITD samples were applied. To control for a lack of digestion, one primer contains an additional Eco RV recognition sequence, so an undigested sample would yield a band 10 bp longer than expected. Appropriate positive controls (standard set 3; see above), negative controls (wt DNA) and no template controls were run with every batch of samples. All positive samples were subsequently confirmed by DNA sequencing, but not as part of the screening procedure.

Laboratory cross-validation procedure

Several steps were taken to ensure consistency of *FLT3* mutation screening across all participating laboratories. Prior to the study start, a common assay protocol (including monuclear cell preparation, DNA extraction, PCR performance and analysis) was defined by the groups and circulated to all laboratories (see below). In order to demonstrate equivalent performance of this *FLT3* mutation assay, all reference laboratories participated in two pre-validation procedures (in April and May 2007). In order to perform external validation of performance, before the start of patient accrual and during the screening phase of the study, (i.e. between December 2007 until July 2011, on average every 6 months) 8 rounds of cross validation were performed. The detailed results of this comprehensive procedure are going to part of a separate manuscript. In brief; DNA standard samples (set 2, see below) were produced and then shipped to a central facility. Aliquots were distributed in a blinded fashion to all

participating laboratories. Every round of cross validation included 25 samples (10 *FLT3* ITD samples with high and low allelic burden, 5 *FLT3* TKD samples with low allelic burden and 10 *FLT3* wt samples). Results were reported to the central data manager and analyzed, only laboratories showing acceptable performance (no positive result in wt samples, no major quantitative deviation) were allowed to move on with *FLT3* screening.

Standard materials

Three sets of standard material were produced and distributed to the participants: First set (for pre-validation): 10 samples containing dilutions of *FLT3*-ITD positive (MV4;11) and negative cell lines (HL60). Second set (used for cross validation): 50 samples of whole genome (WGA) amplified DNA from patients with different *FLT3*-ITD and TKD mutations as well as wt DNA. By dilution of the mutant samples with wt DNA, the mutant allele burden was adjusted to bracket the critical cut-off points for inclusion (i.e. ratio wt *FLT3* to mutant *FLT3* 0.05) as well as for stratification (i.e. the *FLT3* ITD high vs. low allelic burden ratio 0.7). The third set (used for control of individual assays): aliquoted positive control samples for ITD and TKD (from WGA material) was adjusted to have an allele burden just above the cut-off point of 0.05.

3. Study Conduct, Data Collection and Monitoring

Participating cooperative groups included: Alliance/CALGB (Cancer and Leukemia Group B, US), AMLSG (AML Studiengruppe, Germany), CETLAM (El Grupo Cooperativo de Estudio y Tratamiento de las Leucemias Agudas y Mielodisplasicas, Spain) ECOG (Eastern Cooperative Oncology Group, US), EORTC/HOVON (European Organiszation on Treatment and Research in Cancer/Stichting Hemato-Oncologie voor Volwassenen Nederland, Europe), GIMEMA (Gruppo Italiano Malattie Ematologico Malgne dell'Adulto, Italy), NCIC (National Cancer Institute of Canada Clinical Trials Group), OSHO (Ostdeutsche Studiengruppe Hamatologie/Onkologie), PETHEMA (Programma para el Estudio de la Terapeutica en Hemopatia Maligna, Spain) , SAL (Studienallianz Leukamie, Germany), SWOG (Southwest Oncology Group, US) , the Australian LSG (Leukemia Study Group), and individual sites. Alliance/CALGB was the lead group.

All case report forms were transmitted to the Alliance Statistics and Data Center in Durham, NC and Rochester, MN for data collection. Study conduct was monitored by the Alliance Data and Safety and Monitoring Board (DSMB) on a bi-annual basis convened by the Alliance/CALGB and data analysis was conducted by the Alliance Statistics and Data Center. Data quality was assured by review of data by the Alliance Statistics and Data Center and by the study chairperson following Alliance policies.

4. Additional Statistical Details, including an amended analysis plan accounting for a low event rate and secondary endpoints

Change in original statistical plan because of low event rate. Due in part to a higher than expected HCT rate (25% in CR1 and 57% overall) the event rate reached a plateau (6 deaths in 2014, 4 by May 2015) by which time fewer than 70% of the required events were observed. With sufficient follow-up available to assess the efficacy (median of 52.6 months among survivors), an amendment to perform the primary OS analysis was approved by the Alliance DSMB and NCI-CTEP in May 2015 using a critical value of 0.0239 (one-sided accounting for the alpha spent at the interim analysis (0.5%)). In addition, EFS was promoted to be a key secondary endpoint (with confirmatory testing at the one-sided alpha of 0.025 if the OS analysis is significant). Here, we report the results of this primary analysis; a supportive (final) analysis for the OS endpoint will also be conducted at the end of the trial when the 10-year post-randomization follow-up period has been completed for all patients, or when 509 events are observed, whichever occurs first.

Secondary Endpoints. Event-free survival (EFS) was defined as the time from randomization until the earliest qualifying event, including: failure to obtain a CR on or before 60 days of initiation of protocol therapy (protocol-specified CR); relapse; or death from any cause. Patients alive and event-free at the time of the analysis were censored for this endpoint on the date of last clinical assessment. The definition of CR included neutrophil count >1000/ul and platelet count > 100, 000/ul and a marrow showing less than 5% blasts occurring on or before day 60. Other secondary endpoints included: OS where patients who received an HCT were censored at the time of the transplant (subsequently referred to as the censored OS analysis); CR rate; disease-free survival (DFS) defined as the time from documentation of first CR at any time to the first of relapse or death from any cause in patients who achieved a CR; and HCT rates.

Chi-squared and Fisher's exact tests were used to compare treatment groups with respect to baseline characteristics, adverse events (AE) patterns, CR rates and HCT rates. All continuous factors were compared

between the arms using a Wilcoxon rank sum test. All time-to-event endpoints (OS, EFS and DFS) were analyzed using Kaplan-Meier curves and stratified (for *FLT3* status) log-rank tests based on the intent-to-treat principle including all randomized patients. The primary efficacy analyses for OS and EFS were performed ignoring the transplant information (per protocol) followed by secondary sensitivity analyses censoring at the time of transplant. Post-transplant survival was computed using landmark survival models. A competing risks analysis, with death from any cause as a competing risk, was used to compute the cumulative incidence curves for relapse among patients achieving a CR, and compared between the arms using the approach of Gray[2]. Forest plots were used to illustrate the comparisons between the arms within subgroups of interest. Univariable and multivariable Cox proportional hazards models for time-to-event endpoints (OS, and EFS) were performed to understand the impact of treatment, gender, baseline white blood cell counts (WBC), age, *FLT3* status, and modified European LeukemiaNet (ELN) classification on outcomes. Multivariable models were constructed using patients with complete data for all predictors. All analyses were done using SAS® version 9.4

5. Analyses using an expanded CR definition

Using an expanded CR definition (CRs during protocol treatment and those in the 30 days following treatment discontinuation), CR rate was significantly higher in patients randomized to midostaurin compared to placebo (68% vs 61%, two-sided Fisher's exact p=0.04, Table S1A). Patients on the midostaurin arm had superior EFS when an expanded definition of CR was utilized: CRs during treatment and those detected in the 30 days following treatment discontinuation (Table S1B).

6. Post-hoc analyses

Given the exploratory nature of this analysis, and the small sample size (fewer events) in many of the subgroups investigated, these data are hypothesis-generating and need to be investigated in subsequent studies. We performed post-hoc analyses investigating the effect of gender, age, WBC, *FLT3* status and modified ELN classification on OS and EFS in univariable and multivariable Cox models. The models describing only the main effects are summarized in Table S3A and S3B. Midostaurin's positive impact on OS and EFS remains strong after adjustment for age, gender, *FLT3* subtype, WBC, and modified ELN category. We also investigated all pairwise interactions in the multivariable model using a backward elimination approach: p-values <=0.10 allowed a factor to stay in the model, and p-values >0.10 eliminated the factor from that model. There were significant interaction effects for OS: gender by treatment; and age by treatment. Significant interaction effects for EFS included: *FLT3* status by ELN classification; WBC by ELN classification; gender by ELN classification, WBC by gender and WBC by treatment. A Forest plot of the overall treatment effect (stratified for *FLT3* status) as well as within males and females for OS and EFS is shown in Figures S4A and S4B. Females did not have an OS benefit with midostaurin while males did; however, both genders demonstrated improved EFS on midostaurin. Further subgroup analysis for treatment comparisons for OS by gender and by *FLT3* status, and for EFS by modified ELN classification by *FLT3* status are shown in Figures S5A and S5B. Men with *FLT3* TKD and women with *FLT3* ITD did not appear to derive OS benefit from midostaurin. Subset analyses of modified ELN subgroups suggest that patients with normal karyotype derive a midostaurin OS benefit (n=375 normal cases, treatment effect two-sided log-rank p=0.006) but that other ELN subgroups do not.

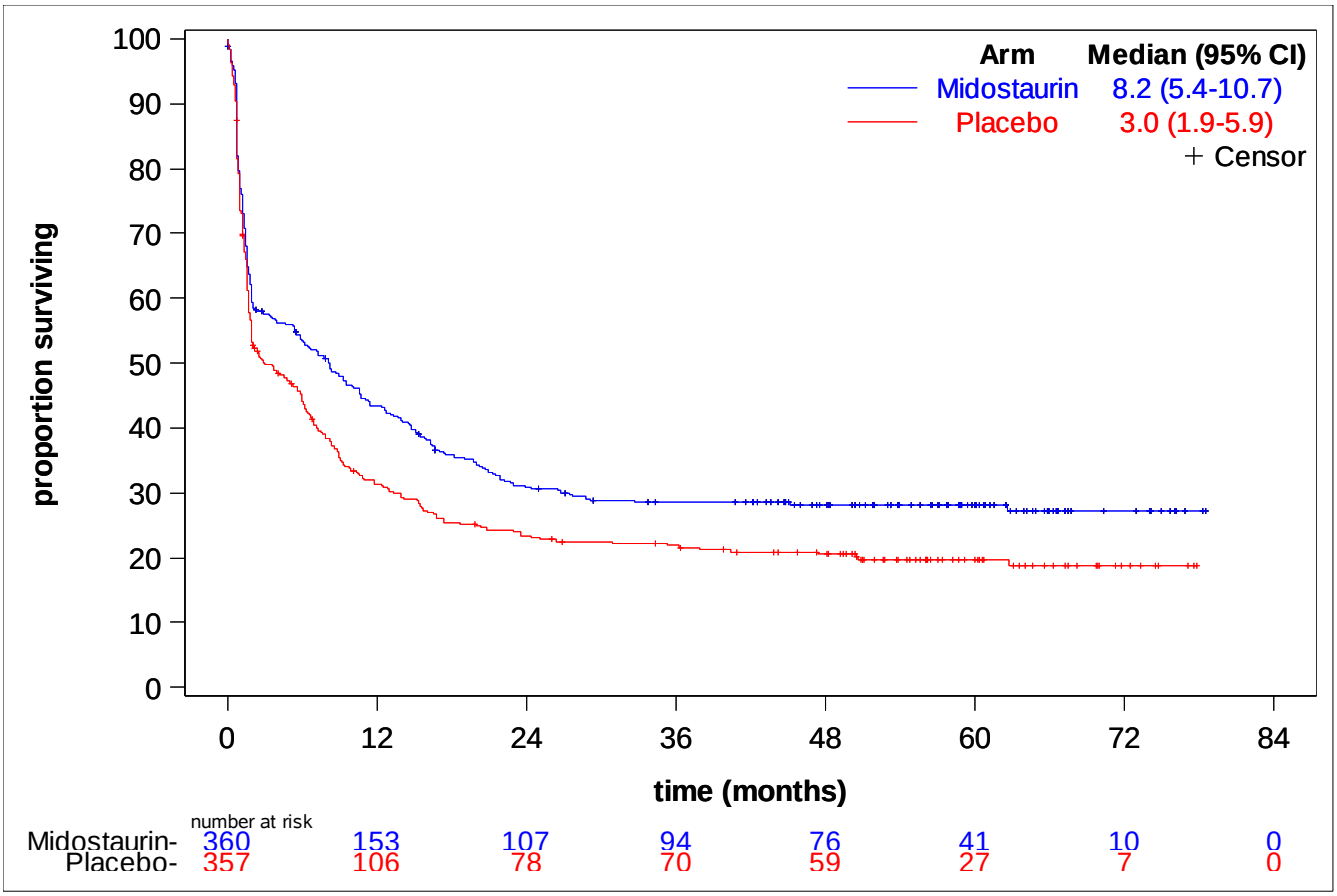
The apparent increased effectiveness of midostaurin in men on OS compared to women needs to be further explored. There is no obvious explanation for this; prior studies with midostaurin showed no difference in pharmacokinetics of the drug according to gender. It is possible that the biology of mutant *FLT3* AML may differ between men and women. Recent data have suggested that gender can have an effect on cancer biology, even in non-hormone sensitive tumors[3]. Interestingly, gender did not influence the EFS results, with men and women both benefitting from midostaurin.

7. Logistical challenges

The clonal and molecular heterogeneity of a given type of cancer poses challenges for therapeutic development of targeted agents particularly in less common neoplasms such as AML. Moreover, *FLT3* mutations are often 'progression' mutations which may occur late in the disease course and are not 'founder' mutations- the eradication of which might be relatively more beneficial due to their presence at an early stage in development of the malignant hematopoietic clone. Nonetheless, mutant *FLT3* was an attractive target for developmental therapeutics in AML. The desire to test a putative *FLT3* inhibitor only in patients with *FLT3* mutant leukemia required a herculean effort in which investigators from around the world needed to screen a large number of routinely diagnosed patients, rapidly assess their molecular status and expeditiously enroll patients onto a prospective randomized trial conducted in multiple centers. Will there be a need to perform similar large trials in other uncommon neoplasms? Rapid molecular screening using more comprehensive techniques such as next-generation sequencing to detect all major recurrent genetic alterations simultaneous could allow early assignment of therapies based on pathophysiologic features thought to be most likely to impact disease biology. Use of novel endpoints, such as determining remission quality by assessment of minimal residual disease, could potentially provide new agents more rapidly for patients with difficult-to-treat cancers.

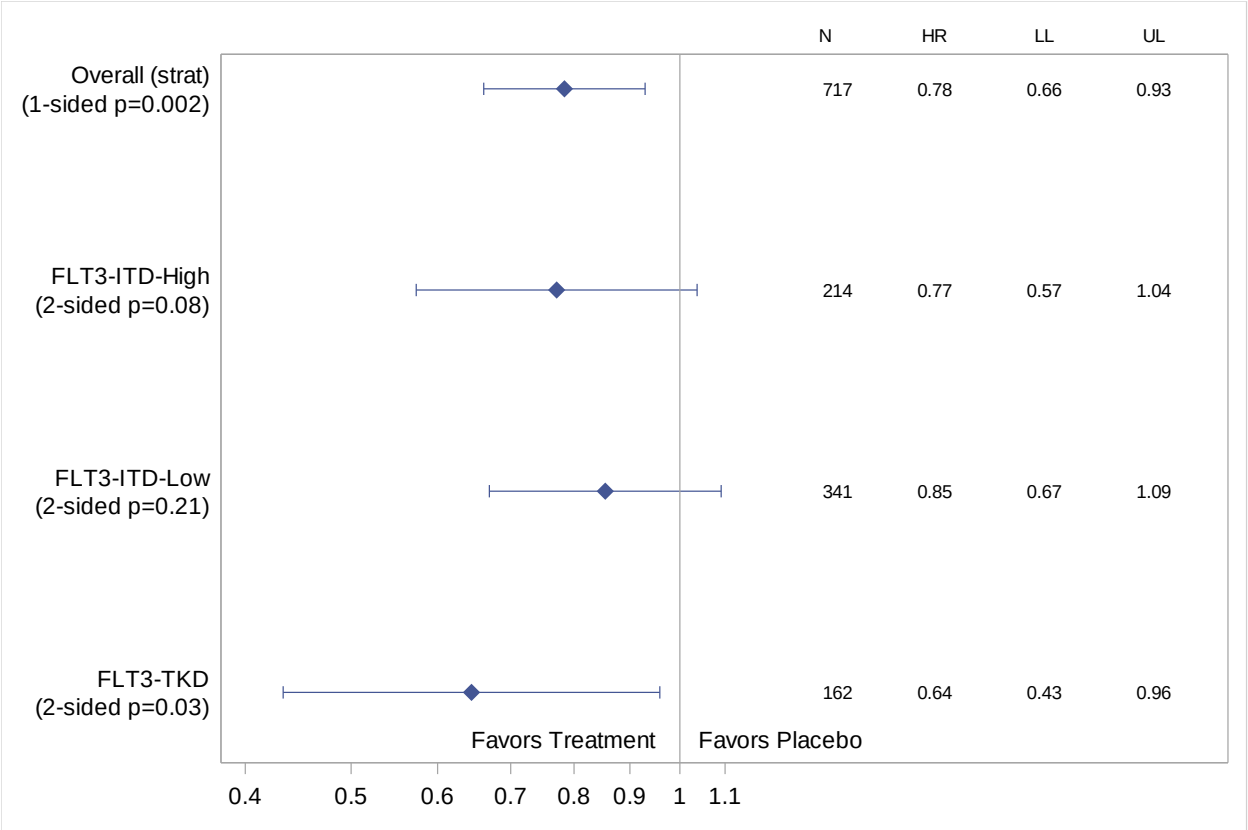
FIGURE S1A. Kaplan-Meier Curves of Event-Free Survival by Arm.

Stratified on *FLT3* subtype, one-sided, log-rank **p=0.002**



Median survival times in months. CI=confidence interval.

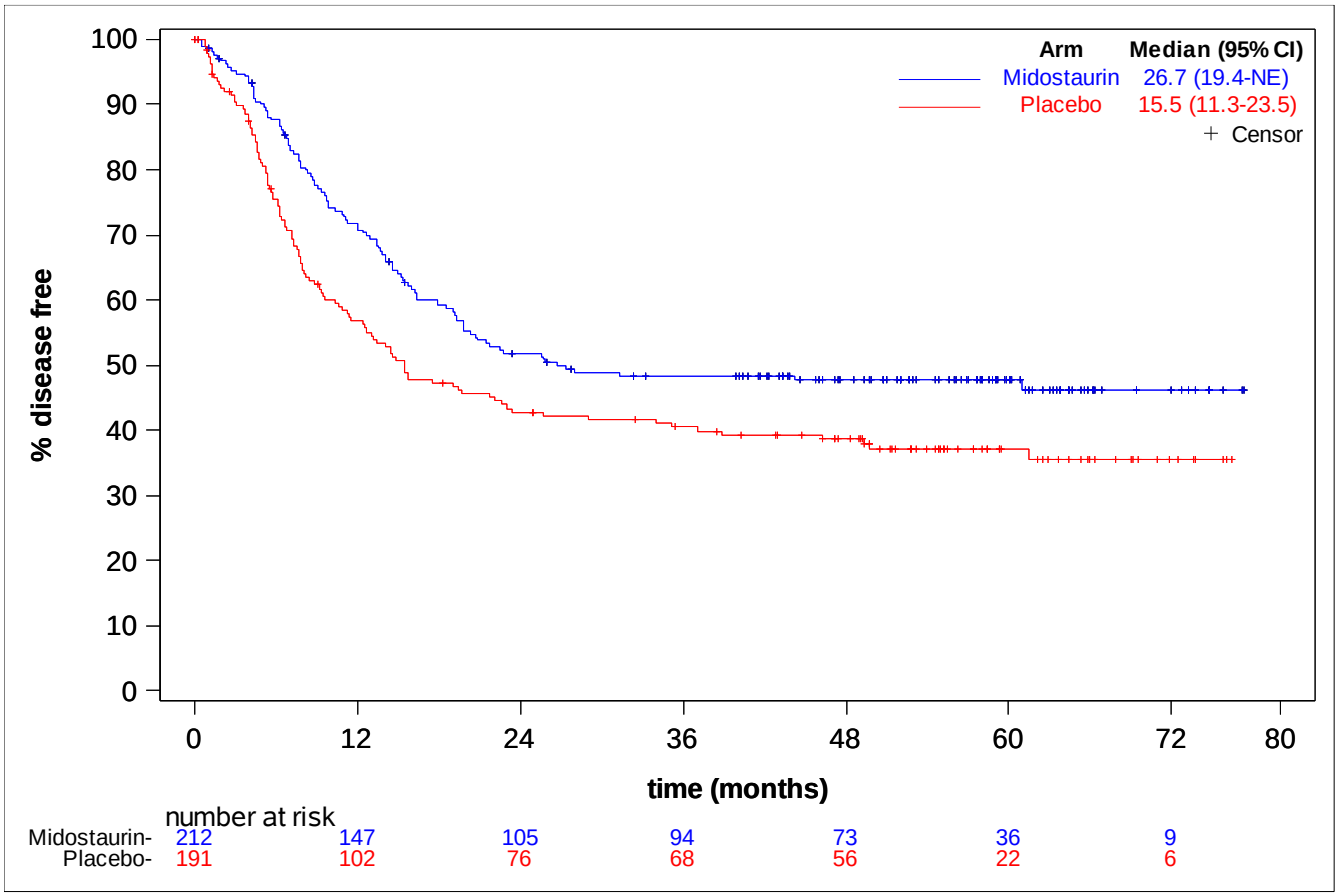
FIGURE S1B. Forest Plot of Event-Free Survival Hazard Ratios, by *FLT3* subgroup.



p=p-value from the Score test
Overall p-value stratified on *FLT3* subtype and gender.
N=number of patients
HR= Hazard ratio
LL=Lower limit of the 95% confidence interval
UL=Upper limit of the 95% confidence interval

FIGURE S2A. Kaplan-Meier Curves of Disease-Free Survival by arm.

Stratified on *FLT3* subtype, two-sided log-rank p=0.01

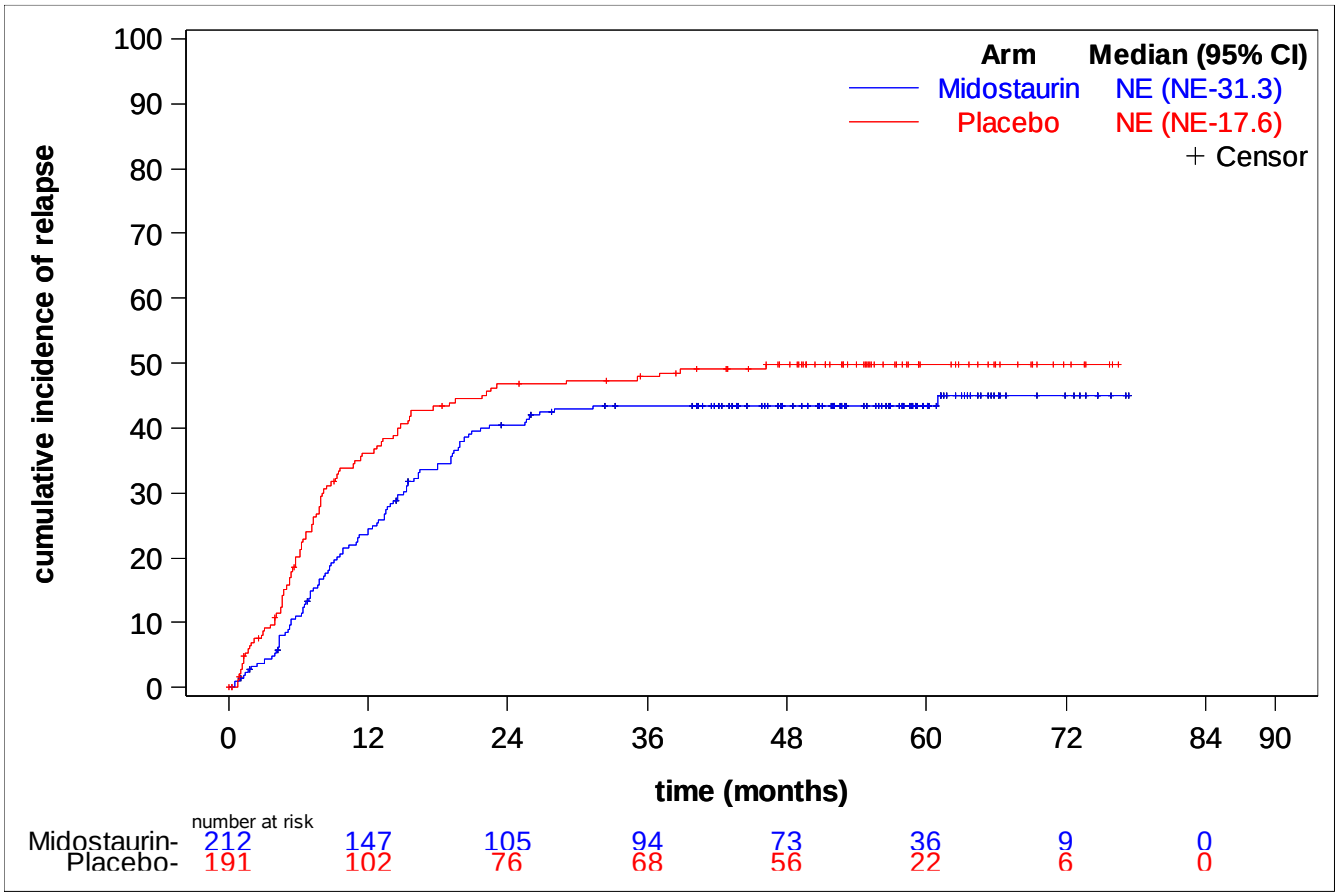


Median survival times in months. CI=confidence interval.

DFS based on protocol specified CR.

Figure S2B. Cumulative Incidence of relapse treating death as a competing risk (protocol CRs only)

Stratified on *FLT3* subtype, two-sided log-rank p=0.13

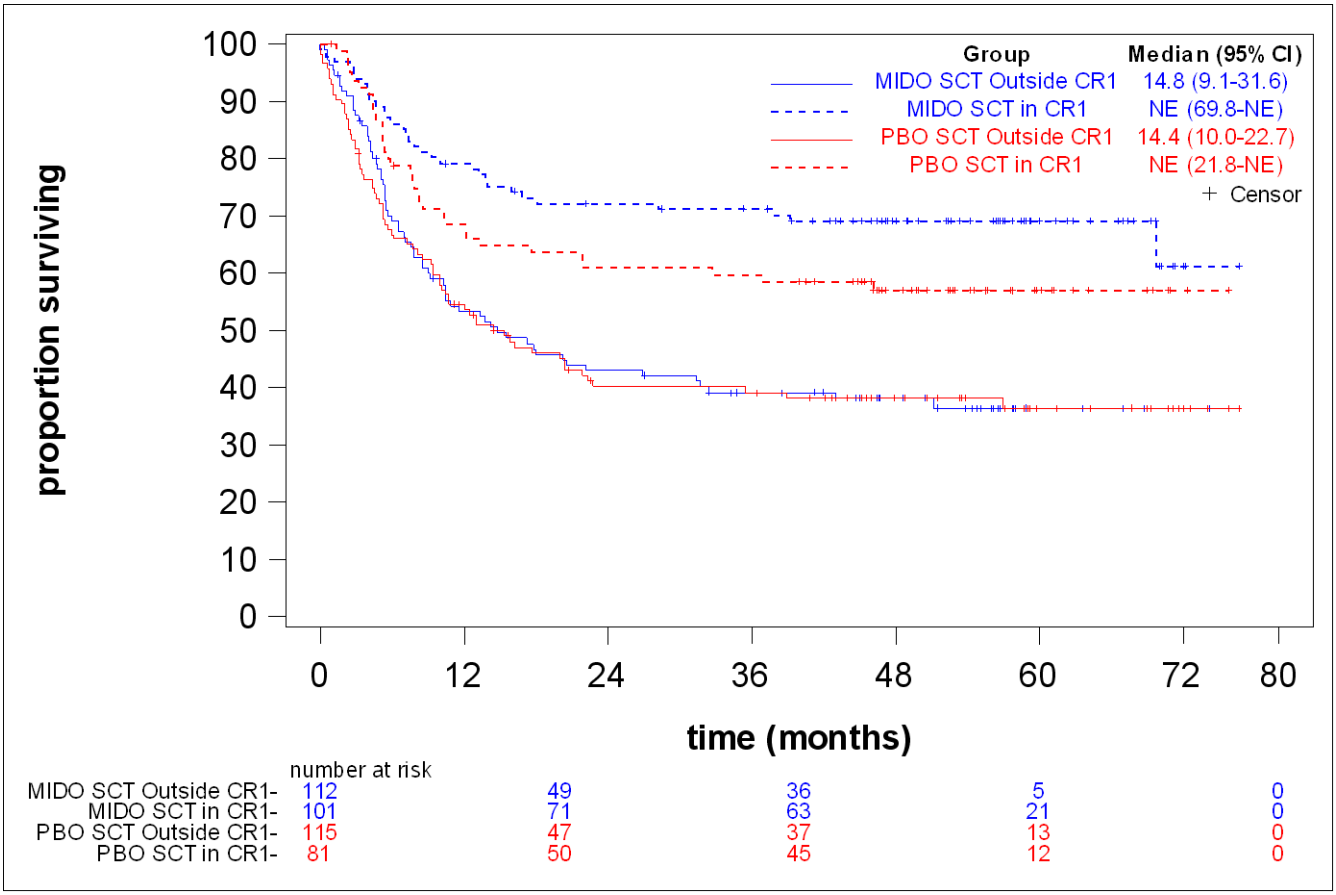


Median time to relapse in months. CI=confidence interval.

FIGURE S3A. Post-transplant Kaplan-Meier survival curves by arm and timing of transplant.

MIDO vs PBO in CR1: Stratified on *FLT3* subtype, two-sided log-rank p=0.07

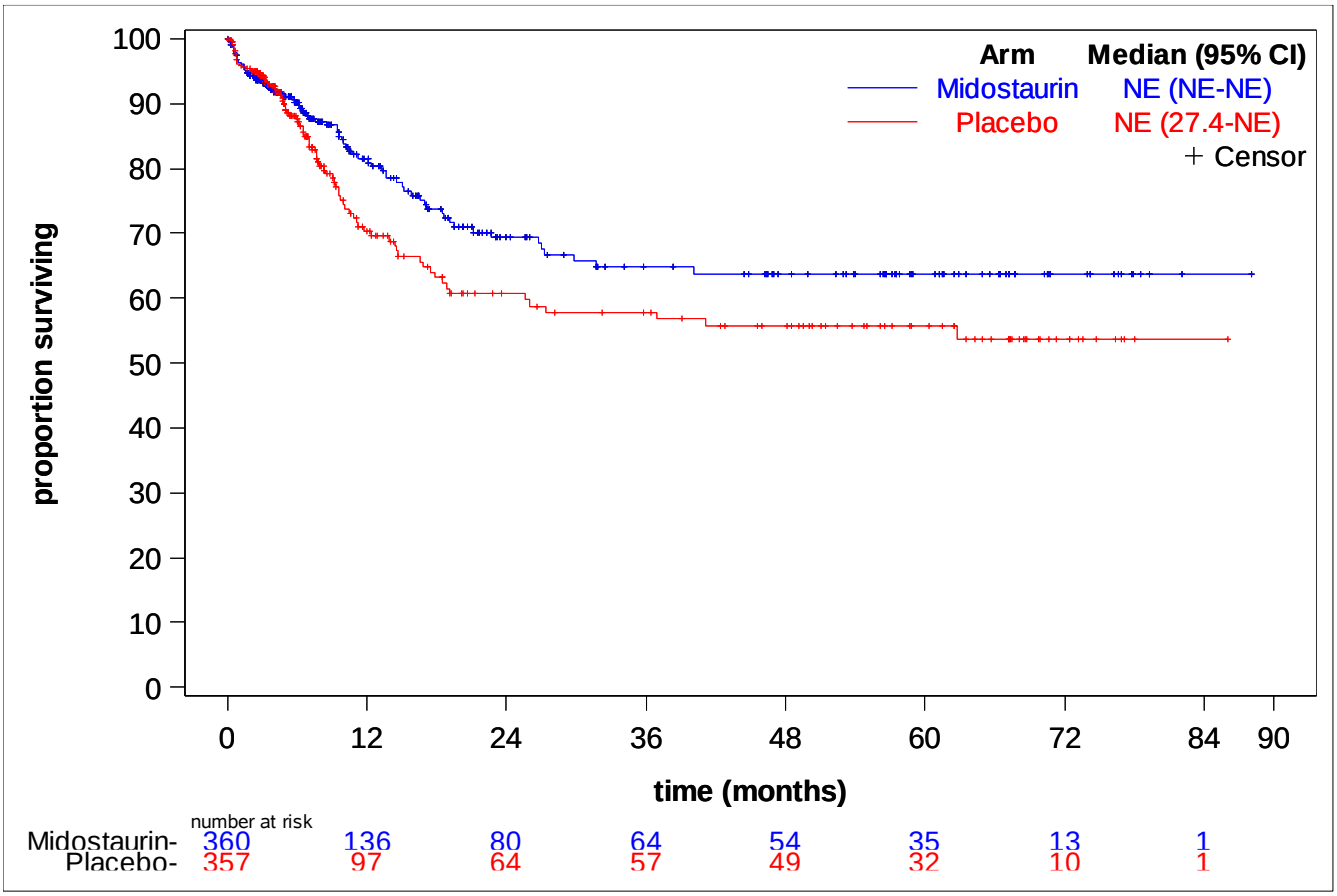
MIDO vs PBO outside CR1: Stratified on *FLT3* subtype, two-sided log-rank p=0.85



Median survival times in months. CI=confidence interval.

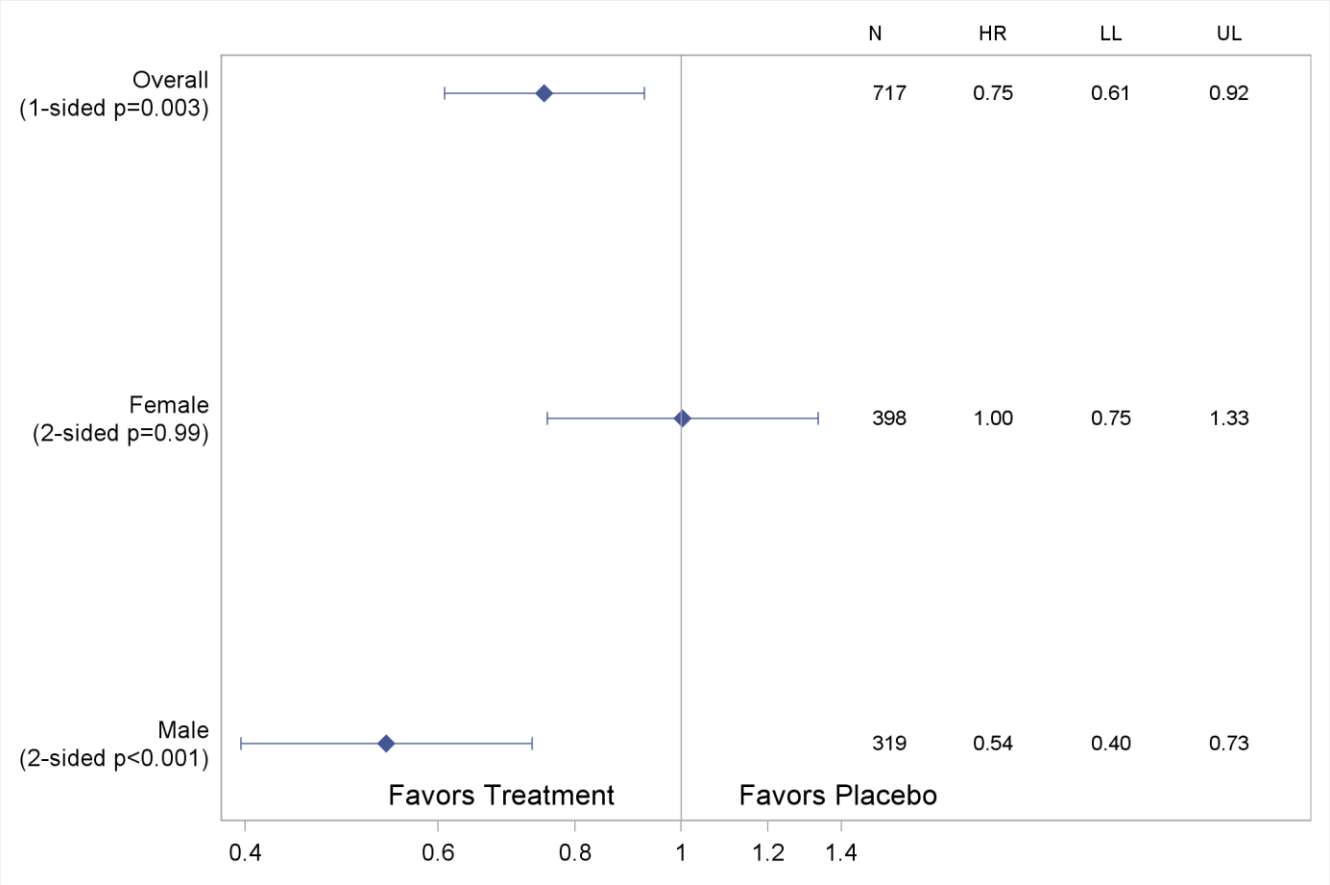
FIGURE S3B. Kaplan-Meier curve of Overall Survival, censoring at the time of transplant.

Stratified on *FLT3* subtype, two-sided, log-rank p=0.08



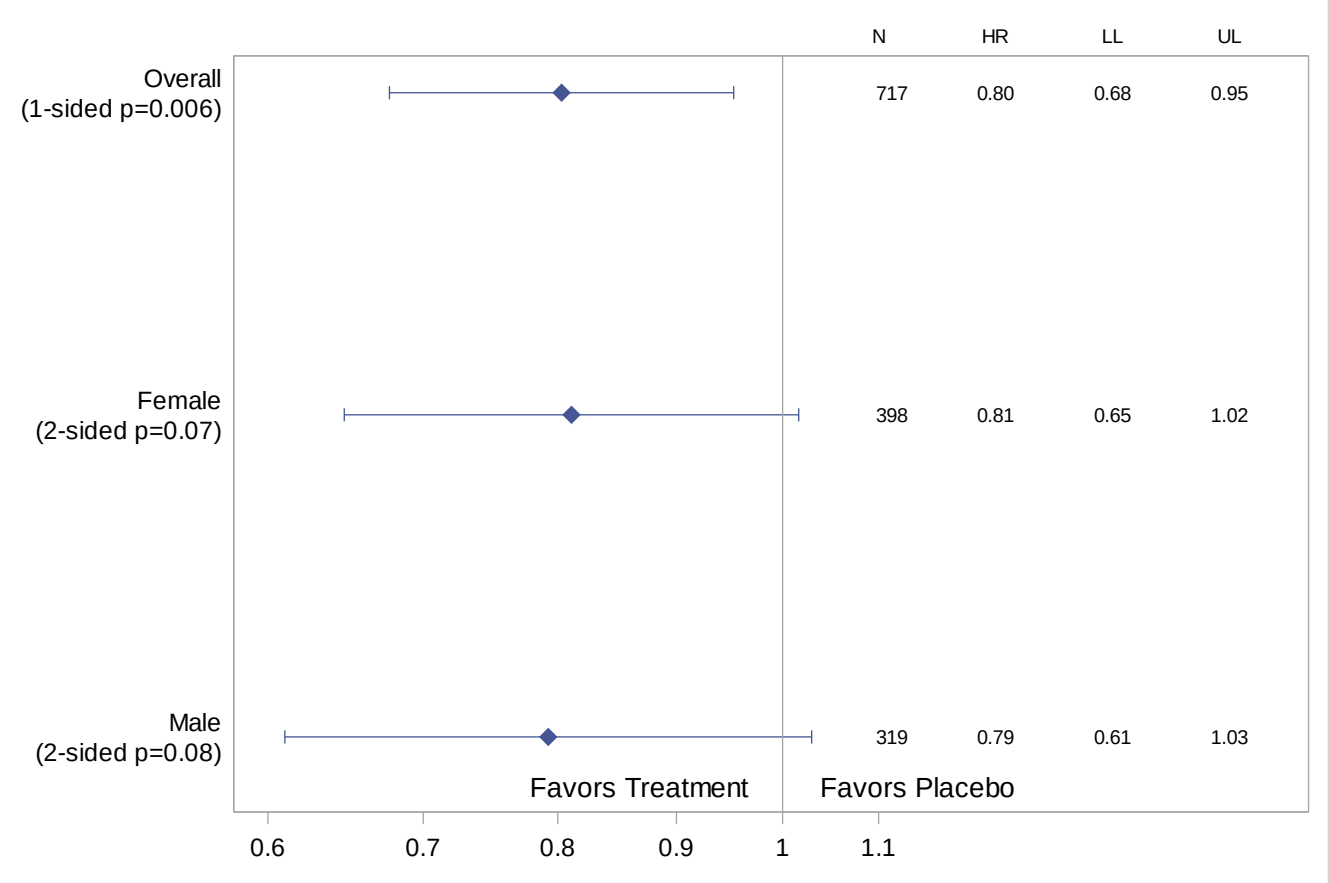
Median survival times in months. CI=confidence interval.

FIGURE S4A. Forest plot of Overall Survival HRs by gender.



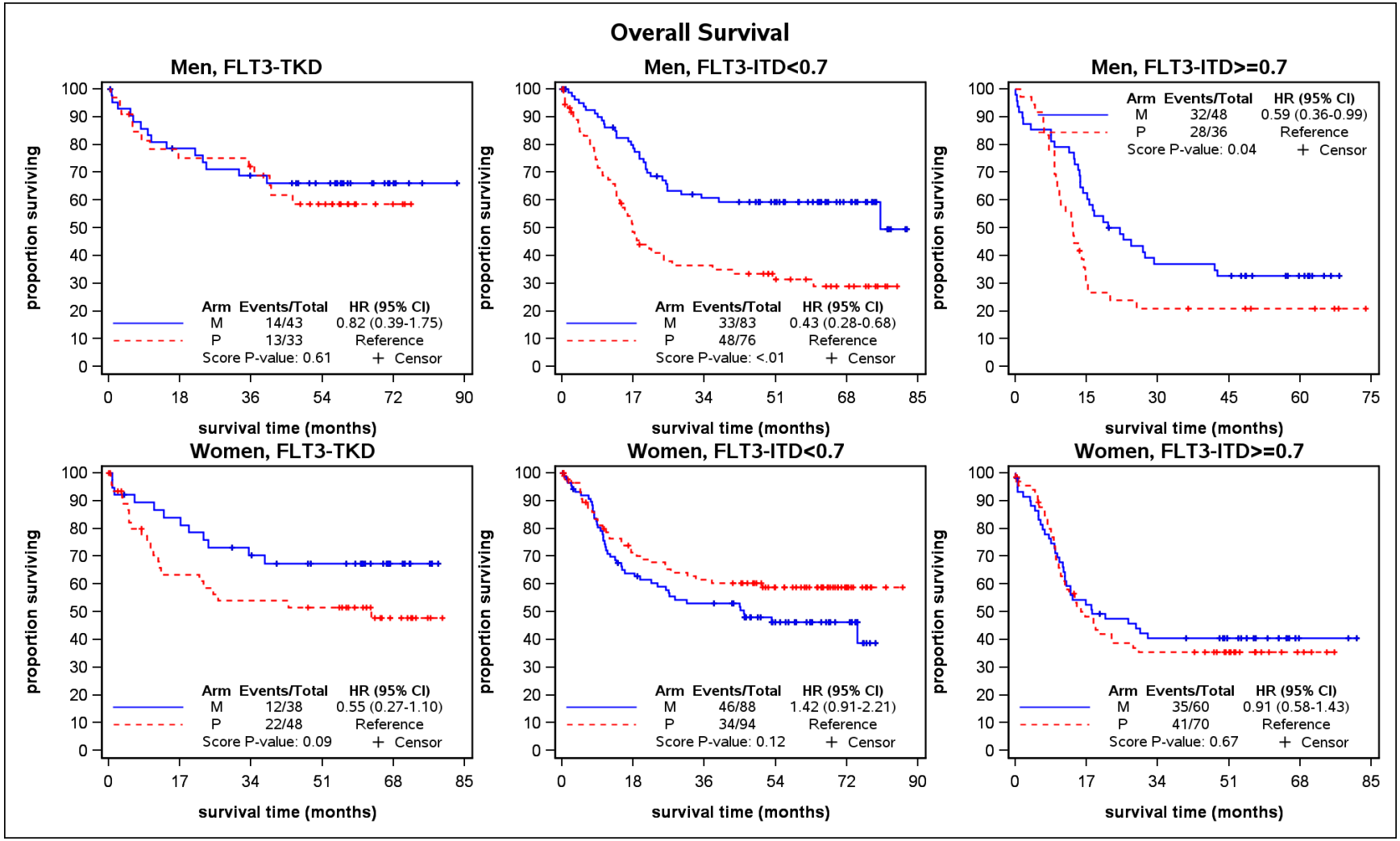
p=p-value from the Score test
p-values for gender subgroups were stratified on *FLT3* subtype.
Overall p-value stratified on *FLT3* subtype and gender.
N=number of patients
HR= Hazard ratio
LL=Lower limit of the 95% confidence interval
UL=Upper limit of the 95% confidence interval

FIGURE S4B. Forest plot of Event-Free Survival HRs by gender.



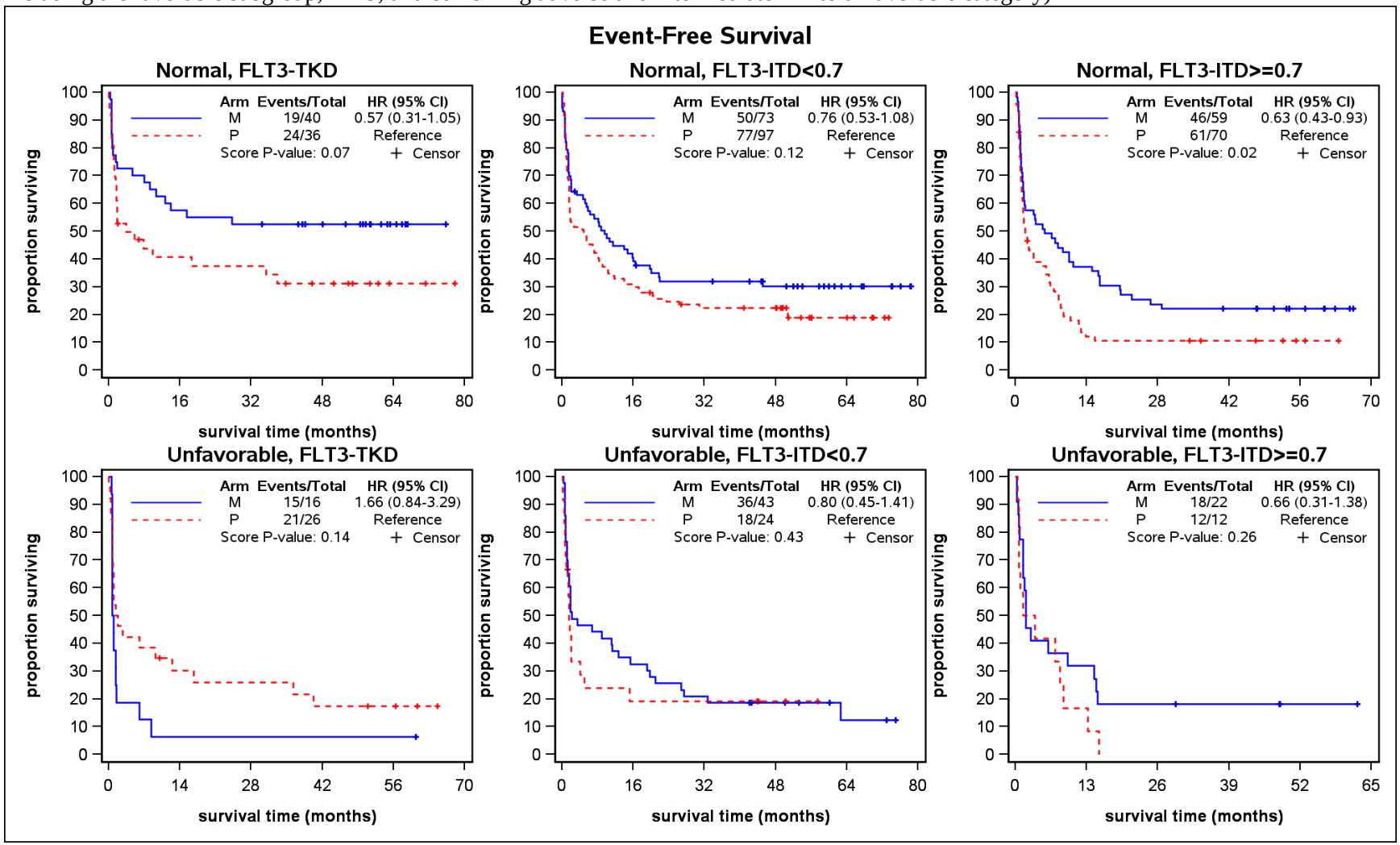
p=p-value from the Score test
p-values for gender subgroups were stratified on *FLT3* subtype.
Overall p-value stratified on *FLT3* subtype and gender.
N=number of patients
HR= Hazard ratio
LL=Lower limit of the 95% confidence interval
UL=Upper limit of the 95% confidence interval

FIGURE S5A. Kaplan-Meier Overall Survival Curves by gender and *FLT3* subtype (two-sided Score test p values)



M=Midostaurin
P=Placebo
p=p-value of the two-sided Score test
HR= Hazard ratio
CI=confidence interval

FIGURE S5B. Kaplan-Meier Event-Free Survival Curves by modified ELN classification and *FLT3* subtype (two-sided score test p values; not including the favorable subgroup, n=29; and combining adverse and intermediate II into unfavorable category)



M=Midostaurin
P=Placebo
p=p-value of the two-sided Score test
HR= Hazard ratio
CI=confidence interval

Table S1A. CR using expanded definition

	MIDO (N=360)	PBO (N=357)	p *
All CRs within 30 days of treatment discontinuation	244	216	
Rate	68%	61%	0.04
Time to CR, median (range), days**	37 (20-192)	36 (20-108)	

* two-sided Fisher's Exact p; **Kaplan-Meier estimates
MIDO=Midostaurin
PBO=Placebo

TABLE S1B. Event-Free Survival Outcomes, using two definitions of CR

	N	events	median (months) (95% CI)	p [3]	estimate of 4 year event-free rate (%, 95% CI)	HR (95% CI)	p [4]
Using protocol-defined CR[1]							
MID	360	256	8.2 (5.4, 10.7)	**1-sided p=0.002**	28.2% (23.6%, 33.0%)	0.784 (0.661, 0.929)	**1-sided p=0.002**
PBO	357	280	3.0 (1.9, 5.9)		20.6% (16.4%, 25.0%)		
Using expanded CR definition[2]							
MID	360	241	11.4 (8.9, 15.3)	**2-sided p<0.001**	32.6% (27.8%, 37.5%)	0.729 (0.612, 0.868)	**2-sided p<0.001**
PBO	357	270	6.2 (4.7, 7.6)		23.3% (19.0%, 28.0%)		

[1] protocol specified CR: CRs occurring on or before 60 days of starting therapy.

[2] using expanded CR definition: protocol specified CRs plus those within 30 days of exiting protocol therapy.

[3] log-rank p-value, stratified on *FLT3* subtype.

[4] Score test p-value, stratified on *FLT3* subtype

CI=confidence interval

TABLE S2A. Transplant rates by arm

Transplant rates (all transplants)					
Arm	N	transplanted	rate	(95% CI)	p *
Midostaurin	360	213	59%	(54%, 64%)	0.26
Placebo	357	196	55%	(50%, 60%)	
Overall	717	409	57%	(53%, 61%)	
Transplant rates (transplants in protocol-defined CR1 only)					
Arm	N	transplanted	rate	(95% CI)	p *
Midostaurin	360	101	28%	(23%, 33%)	0.10
Placebo	357	81	23%	(18%, 27%)	
Overall	717	182	25%	(22%, 29%)	

* two-sided Fisher's Exact p

TABLE S2B. Transplant Timing by Region

	North America n=236	Outside North America n=481
Timing of transplant	n (%)	n (%)
During CR1	68 (28.8%)	114 (23.7%)
Outside CR1	46 (19.5%)	181 (37.6%)
Not transplanted	122 (51.7%)	186 (38.7%)

two-sided Fisher's Exact **p <0.001**

TABLE **S3A**. Univariable and Multivariable Cox Models for Overall Survival

	Univariable Models[1]				**Multivariable Model[2] (n_{eval}=541)**		
Modeling the effect of:	**n_{eval}**	**Hazard Ratio**	**95% CI**	**p-value[3]**	**Hazard Ratio**	**95% CI**	**p-value[3]**
Age at study entry (years)	717	1.011	(1.001, 1.021)	0.03	1.013	(1.002, 1.025)	0.03
Gender (Female vs Male)	717	0.919	(0.747, 1.131)	0.42	0.802	(0.630, 1.021)	0.07
FLT3 subtype (ITD<0.7 vs TKD)	717	1.370	(1.020, 1.840)	0.04	1.396	(1.004, 1.939)	0.05
FLT3 subtype (ITD≥0.7 vs TKD)		2.189	(1.616, 2.964)	<0.01	2.106	(1.480, 2.997)	<0.01
WBC, counts x $10^3/\mu L$ (increments of 10,000 counts)	707	1.029	(1.012, 1.046)	<0.01	1.016	(0.996, 1.037)	0.12
ELN Favorable vs Normal	547	0.543	(0.287, 1.026)	0.06	0.691	(0.347, 1.374)	0.29
ELN Intermediate II vs Normal		1.117	(0.829, 1.506)	0.47	1.259	(0.926, 1.712)	0.14
ELN Adverse vs Normal		1.949	(1.292, 2.941)	<0.01	2.165	(1.433, 3.271)	<0.01
Treatment, MIDO vs PBO	717	0.783	(0.636, 0.963)	**0.02**	0.724	(0.568, 0.922)	<0.01

[1] Univariable Cox models were constructed for each predictor.
[2] One multivariable containing all predictors was constructed, with each predictor adjusted for all others in the model.
[3] Two-sided Wald χ^2 p-values.
n_{eval} =number evaluable
CI=confidence interval
MIDO=Midostaurin
PBO=Placebo

TABLE **S3B.** Univariable and Multivariable Cox Models for Event-Free Survival

	Univariable Models[1]				**Multivariable Model[2] (n_{eval}=541)**		
Modeling the effect of:	**n_{eval}**	**Hazard Ratio**	**95% CI**	**p-value[3]**	**Hazard Ratio**	**95% CI**	**p-value[3]**
Age at study entry (years)	717	0.999	(0.991, 1.007)	0.84	0.999	(0.990, 1.008)	0.87
Gender (Female vs Male)	717	1.217	(1.025, 1.445)	0.03	1.137	(0.932, 1.388)	0.21
FLT3 subtype (ITD<0.7 vs TKD)	717	1.352	(1.073, 1.704)	0.01	1.223	(0.943, 1.585)	0.13
FLT3 subtype (ITD≥0.7 vs TKD)		1.616	(1.264, 2.067)	<0.01	1.514	(1.140, 2.010)	<0.01
WBC, counts x $10^3/\mu L$ (increments of 10,000 counts)	707	1.017	(1.003, 1.032)	0.02	1.018	(1.001, 1.035)	0.04
ELN Favorable vs Normal	547	0.620	(0.380, 1.012)	0.06	0.737	(0.438, 1.241)	0.25
ELN Intermediate II vs Normal		1.200	(0.939, 1.534)	0.14	1.331	(1.032, 1.718)	0.03
ELN Adverse vs Normal		2.406	(1.702, 3.401)	<0.01	2.739	(1.928, 3.892)	<0.01
Treatment, MIDO vs PBO	717	0.787	(0.664, 0.932)	<0.01	0.739	(0.607, 0.899)	<0.01

[1] Univariable Cox models were constructed for each predictor.
[2] One multivariable containing all predictors was constructed, with each predictor adjusted for all others in the model.
[3] Two-sided Wald χ^2 p-values.
n_{eval} =number evaluable
CI=confidence interval
MIDO=Midostaurin
PBO=Placebo

10. References

1. Murphy KM, Levis M, Hafez MJ, Geiger T, Cooper LC, Smith BD, Small D, Berg KDJ. Detection of FLT3 internal tandem duplication and D835 mutations by a multiplex polymerase chain reaction Mol Diagn. 2003; 5:96-102.
2. Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. Annals of Statistics 1988; 16: 1141-54.
3. Yuan Y, Liu K, Chen H, et al. Comprehensive characterization of molecular differences in cancer between male and female patients. Cancer Cell 2016; 29: 711-722.