Oral azacitidine preserves favorable level of fatigue and health-related quality of life for patients with acute myeloid leukemia in remission: results from the phase III, placebo-controlled QUAZAR AML-001 trial

Despite relatively high remission rates with intensive chemotherapy (IC), most patients with acute myeloid leukemia (AML) will relapse, and overall survival (OS) in relapsed AML is dismal.[1] In the phase III, placebo-controlled QUAZAR AML-001 trial, oral azacitidine (Oral-AZA [CC-486]) significantly prolonged OS *versus* placebo (*P*=0.0009; median 24.7 *vs.* 14.8 months from randomization) and relapse-free survival (RFS) (*P*=0.0001; 10.2 *vs.* 4.8 months) as maintenance therapy for patients with AML in first remission after intensive chemotherapy (IC), and was associated with a manageable safety profile.[2] Health-related quality of life (HRQoL) and fatigue generally improve over time for patients with AML in remission; an ideal maintenance treatment should prolong survival without compromising HRQoL.[3,4]

The impact of Oral-AZA on patient-reported fatigue and HRQoL, a key secondary endpoint in QUAZAR AML-001, was assessed using the self-administered Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale and EuroQoL EQ-5D-3L instruments. We hypothesized that Oral-AZA treatment would not meaningfully worsen fatigue or overall HRQoL from baseline, and that mean changes from baseline in fatigue and HRQoL scores in the Oral-AZA arm would be comparable (i.e., not inferior) to those in the placebo arm.

Topline HRQoL outcomes of this study are described briefly elsewhere.[2] At study entry, patients reported generally favorable levels of fatigue and overall HRQoL. Mean FACIT-Fatigue and EQ-5D-3L health utility index (HUI) scores remained similar to baseline over time during Oral-AZA treatment, with similar changes between the Oral-AZA and placebo arms.[2] We describe previously unreported HRQoL results from QUAZAR AML-001, including longitudinal analyses using linear mixed-effect models for repeated measures (MMRM), outcomes in patient subgroups defined by prognostic baseline characteristics, and rates of clinically meaningful deterioration in HRQoL scores.

QUAZAR AML-001 was a randomized, double-blind, placebo-controlled phase III trial. Study design and endpoints are reported in detail elsewhere.[2] Briefly, patients aged ≥55 years, with intermediate- or poor-risk cytogenetics at diagnosis, ECOG PS ≤3, and ineligible for transplant, were randomized to Oral-AZA 300 mg or placebo once-daily for 14 days/28-day cycle within 4 months after achieving first CR or CR with incomplete hematologic recovery (CRi) with IC (induction ± consolidation). Patients who relapsed on-study with 5-15% blasts could receive an escalated 21-days/cycle dosing schedule at the discretion of the treating investigator.

The FACIT-Fatigue Scale is a 13-item questionnaire that measures an individual's level of fatigue during daily activities over the previous week. The EQ-5D-3L is a generic instrument that includes a descriptive questionnaire that assesses impairment across five dimensions (mobility, self-care, pain/discomfort, usual activities, anxiety/depression) at three severity levels (none, moderate, severe), and a visual analogue scale (VAS) that asks patients to rate their perceived HRQoL from 0-100. Higher scores indicate lower fatigue (FACIT-Fatigue) and better health state (EQ-5D-3L). Both instruments were completed on day 1 of each cycle and end-of-treatment (EOT). HRQoL-evaluable patients had non-missing assessments at baseline and ≥1 post-baseline visit.

In order to interpret changes from baseline, we used predefined thresholds for clinically meaningful changes within/between treatment arms (i.e., minimally important differences [MID]) and at the individual level (i.e., responder definitions [RD]).[5] Thresholds used to define clinically meaningful improvement and deterioration from baseline, respectively, were score changes of +3/–3 on the FACIT-Fatigue Scale;[6] +0.08/–0.10 on the EQ-5D-3L HUI;[7,8] and +11/–11 on the EQ-5D VAS.[8]

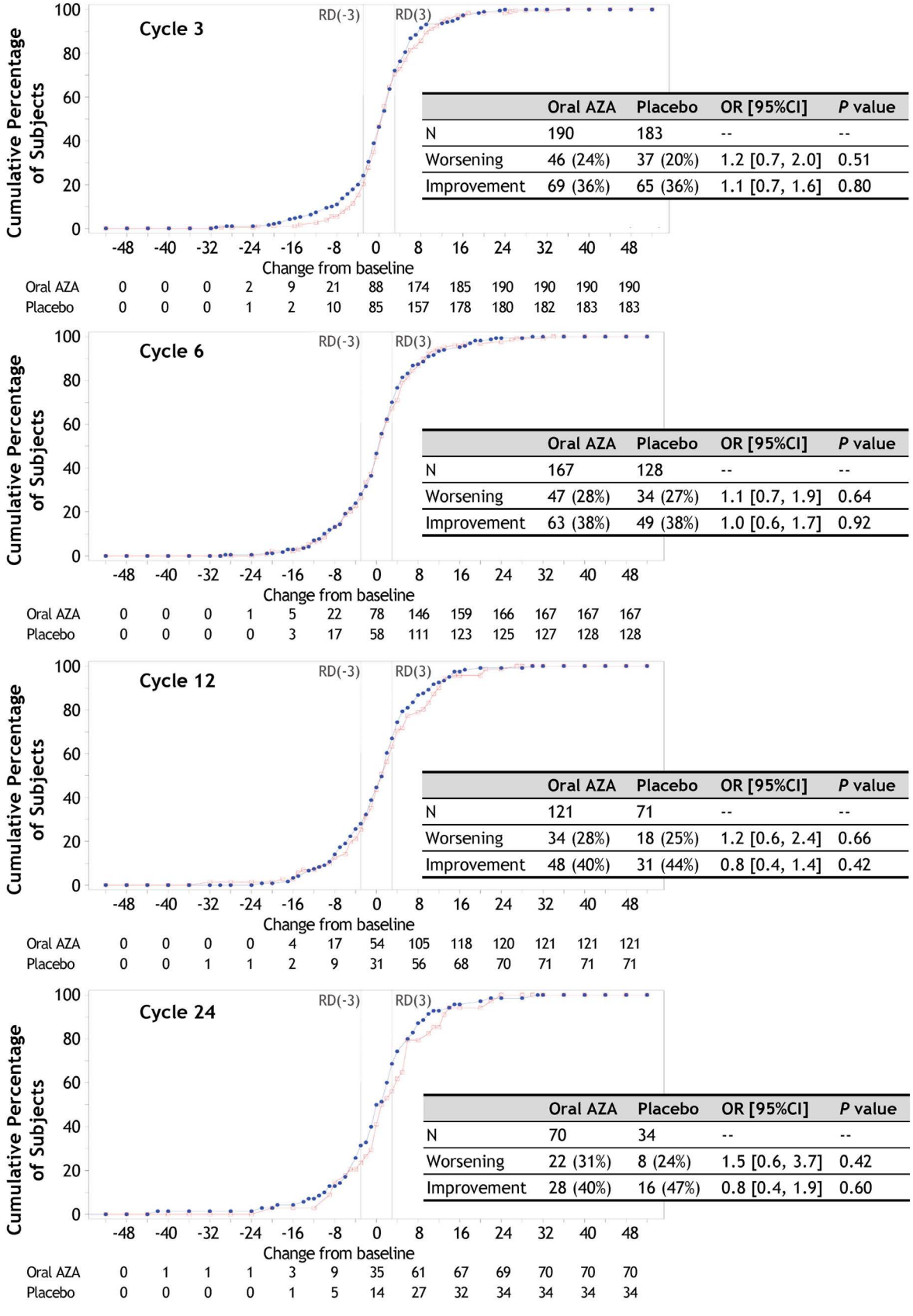
MMRM models were performed to confirm the hypothesized non-inferiority of Oral-AZA and placebo;[9] these models used an unstructured covariance matrix and included the intercept and visit as random effects, and treatment arm, randomization stratification factors,[2] baseline HRQoL score, visit, baseline-by-visit interaction, and treatment-group-by-visit interaction as fixed effects. The dependent variable was change in HRQoL score from baseline. Non-inferiority of Oral-AZA *versus* placebo was demonstrated if the lower bound of the two-sided 95% confidence interval (CI) of the between-group difference in the overall least-squares (LS) mean change from baseline was greater than the MID for deterioration at each assessment.[5,10]

Empirical cumulative distribution frequency (eCDF) curves were generated showing FACIT-Fatigue score changes from baseline for individual patients within each treatment arm at cycles 3, 6, 12, and 24, using the predefined RD for clinically meaningful improvement and deterioration (+3/–3 points). Time to confirmed deterioration was assessed for each patient from the time of randomization until the first of ≥ 2 consecutive visits with a change from baseline surpassing the RD for clinically meaningful deterioration, or until death. Time to confirmed deterioration was estimated using Kaplan-Meier product-limit methods and compared between treatment arms using a stratified Cox proportional hazards regression model with treatment group and baseline score as covariates.

Table 1. Mixed-effect models for repeated measures analyses: overall least-squares mean changes from baseline within in each arm, between-group differences in overall least-squares mean changes, and prespecified minimally important differences for each assessment.

Assessment	Overall LS mean [95%CI] change from baseline		Difference in overall LS mean change, Oral-AZA *vs.* placebo, mean [95%CI]*	Prespecified MID for clinically meaningful worsening
	Oral-AZA	Placebo		
FACIT-Fatigue scale	–0.60 [–2.19, 0.99]	0.29 [–1.44, 2.02]	–0.89 [–2.37, 0.59]	–3
EQ-5D-3L health utility index	–0.01 [–0.03, 0.01]	0.00 [–0.02, 0.02]	–0.01 [–0.03, 0.01]	–0.10
EQ-5D visual analogue scale	2.64 [–0.59, 5.86]	3.59 [0.01, 7.17]	–0.95 [–4.38, 2.47]	–11

*Mixed-effect models for repeated measures (MMRM) analyses confirmed the noninferiority of Oral-AZA effects on fatigue and overall HRQoL *vs.* placebo, as the lower bounds of the 95% confidence interval (CI) for between-group differences in least-square (LS) mean changes from baseline did not exceed the predefined minimally important difference (MID) for worsening on any assessment. AZA: azacitidine; FACIT: functional assessment of chronic illness therapy.



	Oral AZA	Placebo	OR [95%CI]	*P* value
N	190	183	--	--
Worsening	46 (24%)	37 (20%)	1.2 [0.7, 2.0]	0.51
Improvement	69 (36%)	65 (36%)	1.1 [0.7, 1.6]	0.80

	Oral AZA	Placebo	OR [95%CI]	*P* value
N	167	128	--	--
Worsening	47 (28%)	34 (27%)	1.1 [0.7, 1.9]	0.64
Improvement	63 (38%)	49 (38%)	1.0 [0.6, 1.7]	0.92

	Oral AZA	Placebo	OR [95%CI]	*P* value
N	121	71	--	--
Worsening	34 (28%)	18 (25%)	1.2 [0.6, 2.4]	0.66
Improvement	48 (40%)	31 (44%)	0.8 [0.4, 1.4]	0.42

	Oral AZA	Placebo	OR [95%CI]	*P* value
N	70	34	--	--
Worsening	22 (31%)	8 (24%)	1.5 [0.6, 3.7]	0.42
Improvement	28 (40%)	16 (47%)	0.8 [0.4, 1.9]	0.60

Figure 1. Empirical cumulative distribution frequency curves of observed changes from baseline on FACIT-Fatigue scores for individual patients in the oral azacitidine and placebo arms at cycles 3, 6, 12 and 24. A positive change score includes an improvement from baseline. A change from baseline ≥3 was used to define clinically meaningful improvement and worsening. Odds ratio, 95% confidence interval (CI), and *P* values were estimated using Cochran-Mantel-Haenszel test, stratified by randomization stratification factors. ECDF: empirical cumulative distribution frequency; AZA: azacitidine; FACIT: functional assessment of chronic illness therapy; RD: responder definition.

The FACIT-Fatigue-evaluable population comprised 225 of 238 patients (94.5%) randomized to Oral-AZA and 219 of 234 patients (93.6%) randomized to placebo, and the EQ-5D-3L–evaluable population included 225 and 217 patients, respectively. Baseline demographic and disease characteristics of HRQoL-evaluable patients were balanced between treatment arms (*Online Supplementary Table S1*). FACIT-Fatigue and EQ-5D-3L compliance rates were >95% in both treatment arms at baseline and remained high (>85%) across postbaseline visits except at EOT (~65%), suggesting that HRQoL endpoints were unlikely to be confounded by missing data. Patient-reported FACIT-Fatigue, EQ-5D-3L HUI, and EQ-5D VAS scores were comparable between treatment groups at baseline and similar to reference values from general populations in the United States (FACIT-Fatigue) and

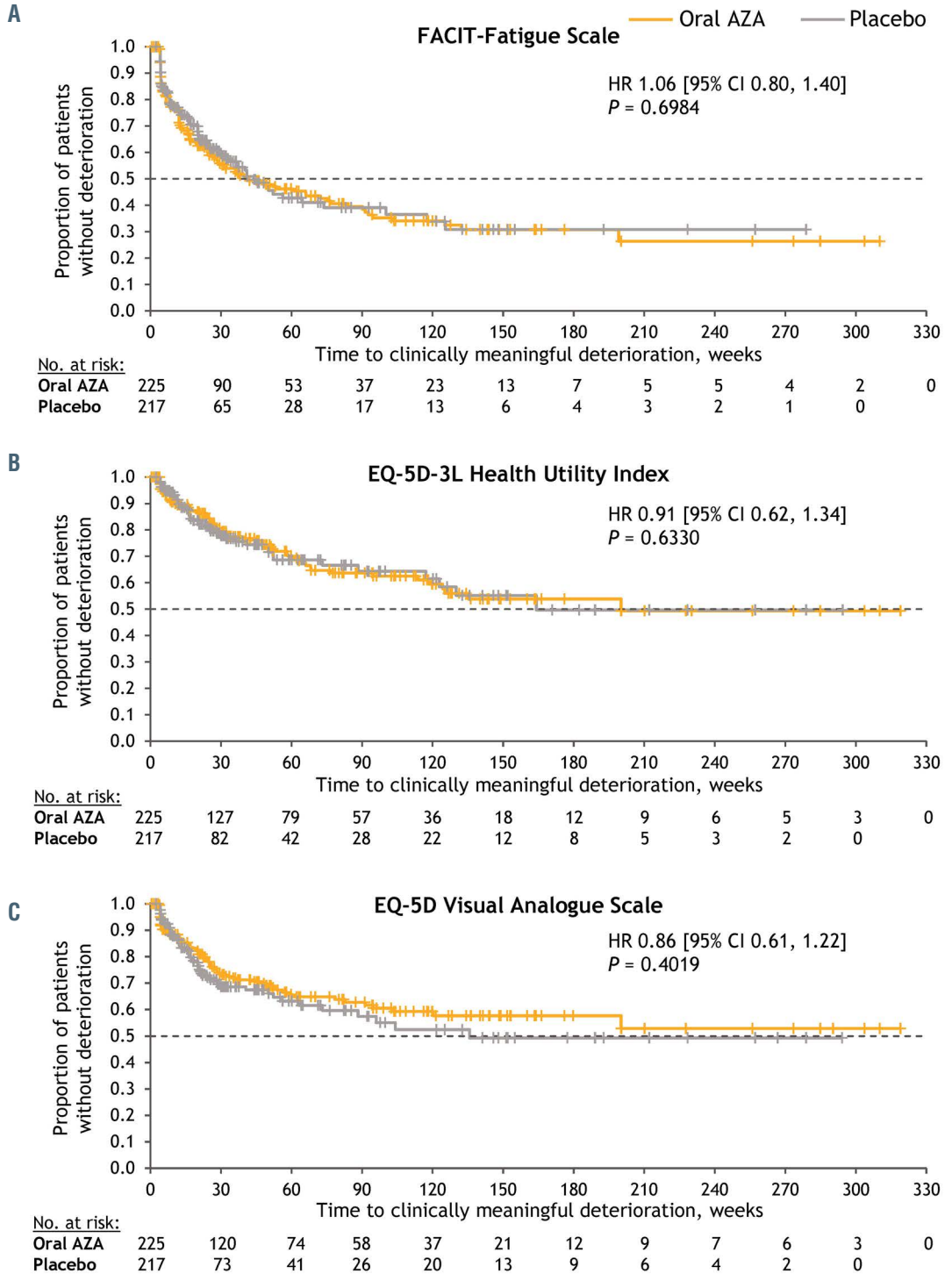


Figure 2. Kaplan-Meier estimated times to confirmed deterioration from baseline. (A) FACIT-Fatigue scale. (B) EQ-5D-3L health utility index. (C) EQ-5D visual analogue scale scores. Time to definitive deterioration was defined as time from randomization to clinically meaningful deterioration sustained from ≥2 consecutive assessment visits. cAZA: azacitidine; FACIT: functional assessment of chronic illness therapy; CI: confidence interval; HR: hazard ratio.

Germany (EQ-5D-3L) (*Online Supplementary Table S2*).[2,11,12] Median treatment durations for HRQoL-evaluable patients were 12 cycles and 7 cycles in the Oral-AZA and placebo arms, respectively.

As reported previously, there were no clinically meaningful differences in observed mean changes from baseline FACIT-Fatigue or EQ-5D-3L HUI scores within treatment arms, or between the Oral-AZA and placebo arms, at any postbaseline visit.[2] Longitudinal MMRM analyses confirmed the non-inferiority of Oral-AZA effects on fatigue and overall HRQoL relative to placebo, as the lower bounds of the 95% CI for between-group differences in LS mean changes from baseline did not exceed the predefined MID for worsening on any instrument (Table 1).

In subgroup analyses, observed mean HRQoL scores generally remained similar to baseline over time within each arm. Mean changes in FACIT-Fatigue, EQ-5D-3L HUI, and EQ-5D VAS scores were comparable between treatment arms within patient subgroups defined by cytogenetic risk at diagnosis (intermediate/poor), response after induction (CR/CRi), receipt of consolidation chemotherapy (yes/no), ECOG PS score (0-1/2-3), age (<65/65-74/≥75 years), and HRQoL domain score (<25th/25th-74th/≥75th percentile). Overall, 45 HRQoL-evaluable patients experienced relapse with 5-15% blasts and received Oral-AZA for 21 days/cycle. Escalated Oral-AZA dosing was not associated with clinically meaningful differences in changes from baseline in mean FACIT-Fatigue, EQ-5D-3L HUI, or EQ-5D VAS scores at any visit compared with 14-day Oral-AZA dosing.

eCDF curves detailing individual FACIT-Fatigue changes from baseline in the Oral-AZA and placebo arms at cycles 3, 6, 12, and 24 generally overlapped, with similar proportions of patients reporting clinically meaningful improvement or deterioration at each visit (Figure 1). Proportions of patients with clinically meaningful deterioration for each measure were low in both treatment arms, and rates were similar between arms on each instrument at almost all post-baseline visits (*Online Supplementary Figure S1*); deterioration rates were significantly higher in the Oral-AZA arm at cycle 19 (EQ-5D VAS) and cycle 29 (FACIT-Fatigue), but these may have occurred by chance as these analyses did not include any adjustments for multiple testing. Times to confirmed deterioration were similar between the Oral-AZA and placebo arms on each instrument (Figure 2). Estimated median times to confirmed deterioration were 41 weeks for Oral-AZA and 44 weeks for placebo on the FACIT-Fatigue (hazard ratio [HR]: 1.06; 95% CI: 0.80-1.40); 200 and 164 weeks, respectively, on the EQ-5D-3L HUI (HR: 0.91; 95% CI: 0.62-1.34); and not reached *versus* 136 weeks on the EQ-5D VAS (HR: 0.86; 95% CI: 0.61-1.22). Similar findings were observed when censoring patients at the time of death.

While improving survival is the primary goal of AML treatment, systematic evaluation of the impact of treatment on HRQoL is essential because prolonged survival may be less meaningful if accompanied by drug-related HRQoL decrements. To our knowledge, QUAZAR AML-001 is the first placebo-controlled study to prospectively investigate the impact of long-term maintenance therapy on HRQoL for patients with AML in remission post-IC. At study entry, these older patients (median age 68 years[2]) reported generally favorable levels of fatigue and overall HRQoL that were comparable to levels in general populations.[11,12] Mean FACIT-Fatigue and EQ-5D-3L scores during Oral-AZA treatment remained at or above baseline levels at almost all post-baseline assessments, and longitudinal MMRM analyses confirmed the noninferiority of Oral-AZA relative to placebo for preserving HRQoL. These HRQoL data are also consistent with the reported manageable safety profile and acceptable tolerability of Oral-AZA in QUAZAR AML-001.[2]

A potential limitation of this study was that HRQoL assessments were conducted on day 1 of each 28-day treatment cycle, allowing for 14 days of recovery after each 14-day dosing period. Additionally, patients in both arms had to undergo routine hospital visits, testing, and marrow collections, which could potentially negatively affect HRQoL outcomes compared with an "observation-only" approach during AML remission.

Oral-AZA administration offers a number of potential benefits, including optimal convenience for patients, no injection-site reactions, fewer clinic visits and lower associated costs, and treatment flexibility for long-term use. Findings from QUAZAR AML-001 show that Oral-AZA significantly improves OS and RFS without compromising fatigue or overall HRQoL for patients with AML in remission.

Gail J. Roboz,[1,2] Hartmut Döhner,[3] Christopher Pocock,[4] Hervé Dombret,[5] Farhad Ravandi,[6] Jun Ho Jang,[7] Dominik Selleslag,[8] Jiří Mayer,[9] Uwe M. Martens,[10] Jane Liesveld,[11] Teresa Bernal,[12] Ming Chung Wang,[13] Peiwen Yu,[14] Ling Shi,[14] Shien Guo,[14] Ignazia La Torre,[15] Barry Skikne,[16,17] Qian Dong,[16] Julia Braverman,[16] Salem Abi Nehme,[15] C. L. Beach[16] and Andrew H. Wei[18]

[1]Weill Cornell Medical College, New York, NY, USA; [2]New York Presbyterian Hospital, New York, NY, USA; [3]Ulm University Hospital, Ulm, Germany; [4]Kent & Canterbury Hospital, Canterbury, UK; [5]Hôpital Saint-Louis, Assistance Publique - Hôpitaux de Paris (AP-HP) and Institut de Recherche Saint-Louis, Université de Paris, Paris, France; [6]The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [7]Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [8]AZ Sint-Jan Brugge-Oostende AV, Bruges, Belgium; [9]University Hospital Brno, Brno, Czech Republic; [10]SLK-Kliniken GmbH, MOLIT Institute for Personalized Medicine, Heilbronn, Germany; [11]Wilmot Cancer Institute, University of Rochester, New York, NY, USA; [12]Hospital Universitario Central de Asturias, Oviedo, Spain; [13]Chang Gung Medical Foundation, Kaohsiung, Taiwan; [14]Evidera, Waltham, MA, USA; [15]Celgene, a Bristol-Myers Squibb Company, Boudry, Switzerland; [16]Bristol Myers Squibb, Princeton, NJ, USA; [17]University of Kansas Medical Center, Kansas City, KS, USA and [18]The Alfred Hospital and Monash University, Melbourne, Vicoria, Australia

Correspondence:

GAIL J. ROBOZ - gar2001@med.cornell.edu

doi:10.3324/haematol.2021.279174

Received: May 7, 2021.

Accepted: September 15, 2021.

Pre-published: September 23, 2021.

Disclosures: GJR reports consultancy or advisory board or data and safety monitoring committee of AbbVie, Actinium, Agios, Amphivena, Amgen, Argenx, Array Biopharma, Astex, Astellas, AstraZeneca, Bayer, Bristol Myers Squibb, Celgene, Celltrion, Daiichi Sankyo, Eisai, Epizyme, GlaxoSmithKline, Helsinn, Janssen, Jasper Therapeutics, Jazz, Mesoblast, MEI Pharma (IDMC Chair), Novartis, Orsenix, Otsuka, Pfizer, Roche/Genentech, Sandoz, Takeda (IRC Chair), Trovagene; and research support from Cellectis. HD reports personal fees from Abbvie, Agios, Astellas, Astex Pharmaceuticals, Helsinn, Janssen, Oxford Biomedicals, and Roche; grants and personal fees from Amgen, Celgene, Jazz Pharmaceuticals, and Novartis; and grants from AROG Pharmaceuticals, Bristol Myers

Squibb, Pfizer, and Sunesis. HD reports grants and personal fees from Celgene, Amgen, Incyte, Novartis, Jazz Pharmaceuticals, Daiichi Sankyo, Servier, and Astellas; and personal fees from Pfizer, Cellectis, Menarini, Otsuka, AbbVie, Janssen, Shire-Baxalta, Celyad, Agios, and Immunogen. FR reports honoraria and consulting fees from Bristol Myers Squibb and Celgene; and research funding from Bristol Myers Squibb. DS reports honoraria from Novartis, Celgene, Amgen, Janssen-Cilag, AbbVie, Alexion, GSK, MSD, Pfizer, Sanofi , Takeda, Incyte, and Teva; consultancy for Novartis, Celgene, Amgen, Janssen-Cilag, AbbVie, Alexion, GSK, MSD, Pfizer, Sanofi, Takeda, Incyte, and Teva; and speakers' bureau participation for Novartis, Celgene, Amgen, MSD, Takeda, and Teva. JM reports research funding from Celgene. UMM reports consultancy for Bristol Myers Squibb, Merck, Amgen, Roche, and Celgene; and travel accommodations/expenses from Bristol Myers Squibb, Amgen, Pierre-Fabre, and Celgene. JL reports participation in DSMB for Onconova. PY, LS, and SG are employed by Evidera. BS is employed by Bristol Myers Squibb. ILT, QD, JB, SAN, and CLB are employed at and have equity ownership in Bristol Myers Squibb. AHW reports study-related fees and personal fees from Celgene; royalties from Walter and Eliza Hall Institute of Medical Research; grants from the Medical Research Future Fund; grants and personal fees from Servier, AbbVie, Novartis, Celgene, Astra Zeneca, and Janssen; and personal fees from Astellas, Pfizer, Macrogenics, and Amgen. CP, JHJ, TB, and MCW report no conflicts of interest.

Contributions: the sponsors collected and analyzed data in conjunction with all authors. The lead author wrote the initial draft of the manuscript. All authors revised the manuscript and reviewed and approved the final version for submission.

Acknowledgments: additional support on an early draft of the manuscript was provided by Sheila Truten and Brian Kaiser from Medical Communication Company, Inc. (Wynnewood, PA, USA), funded by Bristol Myers Squibb and in accordance with Good Publication Practice guidelines

Funding: this study was sponsored and funded by Celgene, a Bristol-Myers Squibb Company.

Data sharing statement: BMS policy on data sharing may be found at https://www.bms.com/researchers-and-partners/independent-research/data-sharing-request-process.html

References

1. Brandwein JM, Saini L, Geddes MN, et al. Outcomes of patients with relapsed or refractory acute myeloid leukemia: a population-based real-world study. Am J Blood Res. 2020;10(4):124-133.
2. Wei A, Dohner H, Pocock C, et al. Oral azacitidine maintenance for acute myeloid leukemia in first remission. N Engl J Med. 2020; 383(26):2526-2537.
3. Alibhai SM, Breunis H, Timilshina N, et al. Quality of life and physical function in adults treated with intensive chemotherapy for acute myeloid leukemia improve over time independent of age. J Geriatr Oncol. 2015;6(4):262-271.
4. Timilshina N, Breunis H, Tomlinson GA, et al. Long-term recovery of quality of life and physical function over three years in adult survivors of acute myeloid leukemia after intensive chemotherapy. Leukemia. 2019;33(1):15-25.
5. Gerlinger C, Schmelter T. Determining the non-inferiority margin for patient reported outcomes. Pharm Stat. 2011;10(5):410-413.
6. Patrick DL, Gagnon DD, Zagari MJ, Mathijs R, Sweetenham J, Epoetin Alfa Study G. Assessing the clinical significance of health-related quality of life (HrQOL) improvements in anaemic cancer patients receiving epoetin alfa. Eur J Cancer. 2003;39(3):335-345.
7. Kvam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. Eur J Haematol. 2011;87(4):330-337.
8. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007;5:70.
9. Ashbeck EL, Bell ML. Single time point comparisons in longitudinal randomized controlled trials: power and bias in the presence of missing data. BMC Med Res Methodol. 2016;16:43.
10. U.S. Department of Health and Human Services - Food and Drug Administration. Non-inferiority clinical trials to establish effectiveness: guidance for industry. In: Services UDoHaH, ed, 2016.
11. Montan I, Lowe B, Cella D, Mehnert A, Hinz A. General population norms for the functional assessment of chronic illness therapy (FACIT)-fatigue scale. Value Health. 2018;21(11):1313-1321.
12. Szende A, Janssen B, Cabases J, editors. Self-reported population health: an international perspective based on EQ-5D [Internet]. Dordrecht (NL): Springer, 2014.
13. National Cancer Institute (NCI). Surveillance, Epidemiology, and End Results (SEER) Program. Cancer Stat Facts: Leukemia - Acute Myeloid Leukemia (AML). [cited March 1, 2021]; Available from: https://seer.cancer.gov/statfacts/html/amyl.html.