

TO THE EDITOR:

Hard-wired biases in trials: maintenance azacitidine in patients with acute myeloid leukemia and framework for future trials

Timothée Olivier,^{1,2} Vinay Prasad,² and Bernard L. Marini³

¹Department of Oncology, Geneva University Hospital, Geneva, Switzerland; ²Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA; and ³College of Pharmacy, University of Michigan, Ann Arbor, MI

Hard-wired bias in trial design may skew the results of randomized, double-blind, placebo-controlled trials.¹ Four biases in the QUAZAR trial design illuminate concerns that are broadly applicable to future trial designs.¹

Oral azacitidine is a hypomethylating agent with a different pharmacokinetic and pharmacodynamic profile compared with the injectable azacitidine formulation.² The QUAZAR AML-001 trial (#NCT01757535; clinicaltrials.gov) was a randomized, double-blind, placebo-controlled, phase 3 trial of oral azacitidine, as maintenance therapy after first remission, in acute myeloid leukemia (AML) patients deemed ineligible for allogeneic hematopoietic cell transplantation (alloHCT).³ The trial reported an overall survival advantage with oral azacitidine over placebo, with an increase in median overall survival from 14.8 months to 24.7 months (hazard ratio, 0.69; 95% confidence interval, 0.55-0.86; $P = .0009$). Based on these results, the US Food and Drug Administration approved azacitidine tablets for “continued treatment of patients with acute myeloid leukemia who achieved first complete remission or complete remission with incomplete blood count recovery following intensive induction chemotherapy and are not able to complete intensive curative therapy.”

There are several strengths of the QUAZAR trial. First, the development of an oral formulation of azacitidine may improve care delivery, access, and quality of life (QoL). Along with other oral compounds being approved (venetoclax), this approval may permit an entirely oral regimen for some AML patients. Second, the study focused on patients with intermediate- to high-risk AML who are ineligible for alloHCT, a population that constitutes an unmet need. Third, the inclusion of QoL assessments was a major step forward in ensuring AML trials remain patient focused. However, there are limitations in the trial that are not immediately apparent, and illustrate how hard-wired biases may limit trials results applicability in the real world (Figure 1).

First, inclusion and exclusion criteria affect the generalizability of results and cannot be adjusted for after the fact. As such, they must be carefully considered. In the QUAZAR trial, some included patients likely could have pursued more intensive therapy in contrast with the stated inclusion criteria. Despite such patients being labeled “ineligible for allogeneic transplantation,” demographics reveal favorable characteristics for more aggressive treatment: patients were only a median age of 68 years old, 48% had an ECOG of 0, and 86% of them presented with intermediate cytogenetic risk AML (with 14% of poor cytogenetic risk). Finally, 3% of enrolled patients were considered ineligible because of unfavorable cytogenetics, a biological feature that usually leads one to consider alloHCT. Furthermore, reasons given for ineligibility were “patient decision” in 11% and “other” in 10%, reasons where physician communication may play a role.⁴ Criteria for ineligibility for alloHCT in the trial were not prespecified. It is therefore possible that many patients in the “grey zone,” where the indication or contraindication to alloHCT was not straightforward, were enrolled in this trial. As evidence of this, the QUAZAR trial reported that 32 patients (14%) in the control group and 15 patients (6%) in the experimental arm were treated at relapse (after the trial) with alloHCT, though these patients were initially deemed ineligible for this strategy.³

Submitted 17 December 2021; accepted 7 March 2022; prepublished online on *Blood Advances* First Edition 14 March 2022; final version published 3 June 2022. DOI 10.1182/bloodadvances.2021006857.

Requests for data sharing may be submitted to Timothée Olivier (timothee.olivier@hcuge.ch).

© 2022 by The American Society of Hematology. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

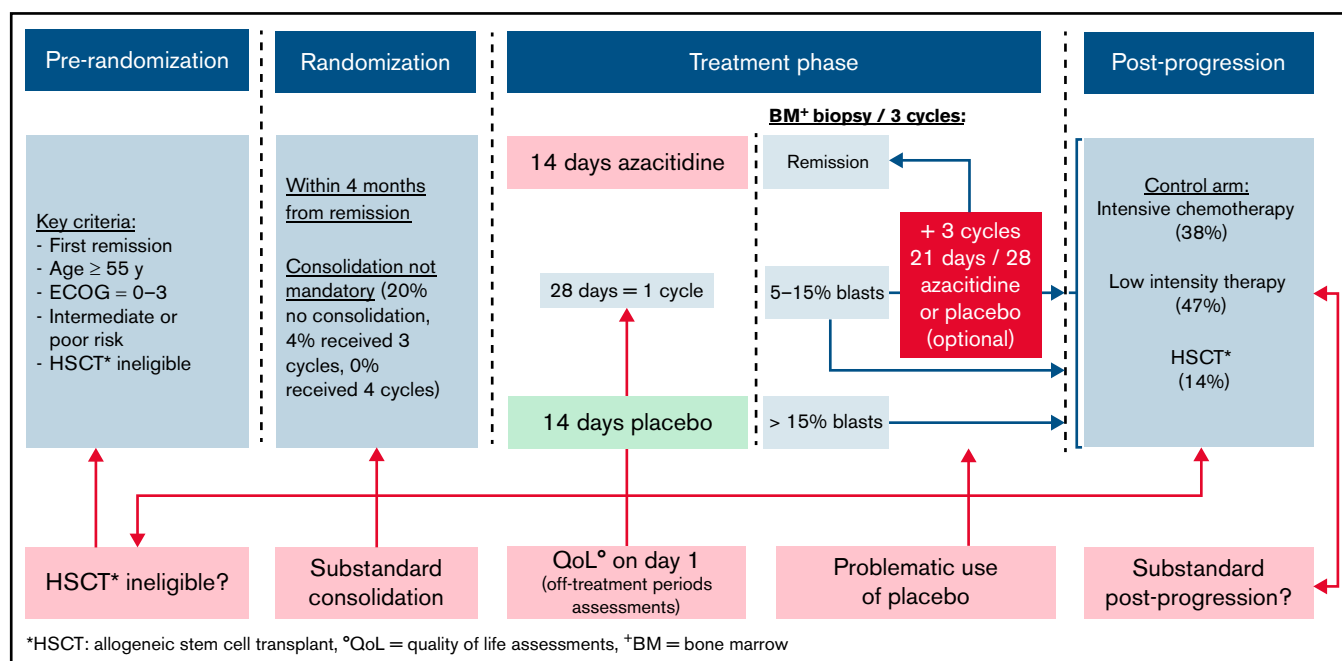


Figure 1. Hard-wired biases in the design of the QUAZAR randomized, double-blind, placebo-controlled, and phase 3 trial.

Second, time constraint before randomization may introduce hard-wired biases also affecting trials external validity. Here, the limited time window (4 months) between achievement of complete remission and randomization may have limited the number of cycles of consolidation and put pressure on early enrollment onto maintenance. Although the optimal number of consolidation cycles for older patients is unknown, studies since the 1980s have demonstrated outcomes are inferior for AML patients who receive no consolidation therapy.^{5,6} In the QUAZAR trial, 20% of patients did not receive any consolidation treatment, 45% received only 1 consolidation, 4% of patients underwent 3 consolidation cycles, and none received 4. Additionally, real-world data from the Connect Myeloid Disease Registry showed a 2.6 mean number of consolidation cycles in patients with intermediate- or poor-risk cytogenetics⁷; consolidation in QUAZAR was inferior to standard practice. A substandard control arm, by definition, artificially favors the experimental drug.

Third, hard-wired bias can be introduced by questionable utilization of placebo, when better options exist. In the QUAZAR trial, drug modification rules allowed, in patients with a relapse suspected based on 5% to 15% blasts cells in the bone marrow biopsy, to increase the “placebo dose intensity” (from 14 to 21 days per 28, plus increasing the dose if previously reduced). Although it may be reasonable, within the experimental arm (azacitidine), to increase the dose with the aim of controlling an early relapse, the same rule cannot be applied in the placebo control arm. A patient taking a placebo who has a rise in blasts should not be given 150% the dose of placebo. The net effect of this rule was to delay the initiation of postprogression treatment in the control arm and increasing the risk of disease-related complications, allowing relapsing patients to receive placebo possibly for up to 3 more cycles until the next bone marrow reassessment. Indeed, in the 17% of patients in the control arm that were continued on placebo while relapsing, frequent ($\geq 10\%$) hematologic adverse events eventually led to discontinuation, and a minority of them (4 out of 35)

presented a least 1 restoration of complete remission. Beyond the ethical issues raised, this design feature further penalizes the control arm.

Fourth, even after trials conclude, hard-wired bias can still occur and distort the results based on postprogression therapy access. Although treating physicians would not intentionally provide substandard care after a trial's end, many trials are run globally, including in countries with limited access to postprogression options. Substandard postprogression therapy has been described in multiple myeloma and renal cell carcinoma trials.^{8,9} Here, in a trial run globally, we hypothesize that postprogression treatment may have been beneath the standard of care. In the placebo arm, 47% of patients received “low intensity” therapy. Salvage low-intensity treatment like hydroxyurea or low-dose cytarabine could have been substandard in such a selected population. Also, when we asked for clarification, the authors of the QUAZAR trial did not specify which therapies the patients received in this subgroup.¹⁰ Because these data were not revealed, substandard postprogression therapy is a possibility, although remaining a hypothesis.

Finally, including QoL analysis as a prespecified endpoint was a strength in the QUAZAR trial: avoiding any impairment in QoL is particularly important when considering a maintenance strategy. The trial reported that “overall health-related QoL was preserved during CC-486 treatment.” However, the timing of QoL assessments may have matter: assessments in the experimental group (receiving oral azacitidine for 2 weeks followed by 2 weeks without treatment) were completed on day 1 of each cycle. These QoL assessments are referring to the patient condition on the present day, the past 24 hours, or up to the 7 past days from the day of assessment. In other words, patients were asked to evaluate their QoL during off-treatment periods: this may have underestimated the burden of the therapy. Pertinent to quality of life, the trial included regular bone marrow biopsies in the control arm, which are not routinely performed in real-world practice. The bone marrow examination was done on day 1 every 3 cycles, the same day as the QoL assessments of these cycles. This

may have led to a detrimental impact in the control arm QoL, therefore potentially reducing between arms differences in QoL. Lastly, the report did not provide a compliance table nor statistical method to handle missing QoL data; therefore, informative censoring cannot be ruled out in the QoL analysis.⁵

It is unclear whether oral azacitidine is the standard of care in the maintenance setting. Although we understand the optimism that the QUAZAR trial results have, particularly in a poor prognosis setting, our work is primarily aimed to illustrate how hard-wired biases, within a double-blind, randomized, placebo-controlled trial, may have impacted the reported benefit of the experimental drug.¹ These biases may occur at each step of a trial, from its initial conception until after the protocol ceases with postprogression therapy. Institutional review board and regulatory agencies should systematically assess for these biases, and scrutiny must be applied before incorporating such agents into clinical practice.

Acknowledgments: This project was funded by Arnold Ventures, LLC through a grant paid to the University of California, San Francisco.

Contribution: B.L.M., V.P., and T.O. contributed to the conception; T.O. wrote first draft of manuscript; all authors reviewed and revised the manuscript; and all authors provided final approval of the manuscript.

Conflict-of-interest disclosure: V.P. discloses research funding from Arnold Ventures; royalties from Johns Hopkins Press and Medscape; honoraria from Grand Rounds and lectures from universities, medical centers, nonprofits, and professional societies; consulting for United Healthcare; speaking fees from Evicore; and Plenary Session podcast has Patreon backers. The remaining authors declare no competing financial interests.

ORCID profiles: T.O., 0000-0002-6936-5783; B.L.M., 0000-0003-0436-8948.

Correspondence: Timothée Olivier, Department of Oncology, Geneva University Hospital, 4 Gabrielle-Perret-Gentil St, Geneva, Switzerland; e-mail: timothee.olivier@hcuge.ch.

References

1. Prasad V, Berger VW. Hard-wired bias: how even double-blind, randomized controlled trials can be skewed from the start. *Mayo Clin Proc.* 2015;90(9):1171-1175.
2. Garcia-Manero G, Gore SD, Cogle C, et al. Phase I study of oral azacitidine in myelodysplastic syndromes, chronic myelomonocytic leukemia, and acute myeloid leukemia. *J Clin Oncol.* 2011;29(18):2521-2527.
3. Wei AH, Döhner H, Pocock C, et al; QUAZAR AML-001 Trial Investigators. Oral azacitidine maintenance therapy for acute myeloid leukemia in first remission. *N Engl J Med.* 2020;383(26):2526-2537.
4. Albrecht TL, Eggly SS, Gleason MEJ, et al. Influence of clinical communication on patients' decision making on participation in clinical trials. *J Clin Oncol.* 2008;26(16):2666-2673.
5. Büchner T, Urbanitz D, Hiddemann W, et al. Intensified induction and consolidation with or without maintenance chemotherapy for acute myeloid leukemia (AML): two multicenter studies of the German AML Cooperative Group. *J Clin Oncol.* 1985;3(12):1583-1589.
6. Cassileth PA, Harrington DP, Hines JD, et al. Maintenance chemotherapy prolongs remission duration in adult acute nonlymphocytic leukemia. *J Clin Oncol.* 1988;6(4):583-587.
7. Roboz G, Abedi M, Thompson MA, et al. Treatment patterns and survival outcomes of patients with acute myeloid leukemia who achieved remission in the connect myeloid disease registry. *Blood.* 2021;138(suppl 1):279.
8. Mohyuddin GR, Koehn K, Abdallah AO, Goodman AM, Prasad V. Reporting of postprotocol therapies and attrition in multiple myeloma randomized clinical trials: a systematic review. *JAMA Netw Open.* 2021;4(4):e218084.
9. Sharp J, Khaki AR, Prasad V. Use of second-line immunotherapy in control arms of randomized clinical trials in kidney cancer: a systematic review. *JAMA Netw Open.* 2021;4(9):e2124728.
10. Oral Azacitidine Maintenance for Acute Myeloid Leukemia. Oral azacitidine maintenance for acute myeloid leukemia. *N Engl J Med.* 2021;384(13):e51.