LETTER

Letter to the Editor Regarding: Indirect Treatment Comparison of Inotuzumab Ozogamicin Versus Blinatumomab for Relapsed or Refractory Acute Lymphoblastic Leukemia, Proskorovsky et al. Adv Ther (2019);36(8):2147–2160. doi:10.1007/s12325-019-00991-w

Jinlin Song · Wei Gao · Jipan Xie · Simran Tiwana

Received: November 25, 2019 / Published online: December 14, 2019
© The Author(s) 2019

Dear Editor,

Two articles in *Advances in Therapy*, Song et al. [1] and Proskorovsky et al. [2], presented indirect treatment comparisons between inotuzumab ozogamicin (InO) and blinatumomab in treating adult patients with relapsed or refractory acute lymphoblastic leukemia (R/R ALL). The same methodology, matching-adjusted indirect comparison (MAIC), was used in both studies. Song et al. concluded that blinatumomab and InO had a similar complete remission (CR) rate, while Proskorovsky et al. concluded that InO had superior efficacy in terms of rates of remission and hematopoietic stem cell transplantation (HSCT) and a favorable trend on event-free survival (EFS) compared to blinatumomab. On examination of the Proskorovsky et al. article, we identified several weaknesses that have led to biased conclusions.

First, Proskorovsky et al. failed to match a key difference between the INO-VATE-ALL and the TOWER trial populations: the number of prior salvage therapies. The INO-VATE-ALL trial population included only patients with no

more than one prior line of salvage therapy, while approximately 23% of the TOWER trial population had two or more lines of prior salvage therapies. Patients treated with more lines of prior salvage therapies have worse outcomes [3]. Specifically, the median overall survival (OS) increased from 7.7 to 8.4 months for patients treated with blinatumomab in TOWER after excluding those with two or more lines of prior salvage therapies. The discrepancy in the number of prior salvage therapies is a fundamental difference between the two trial populations. Without matching the number of prior salvage therapies, bias in favor of InO was introduced in Proskorovsky et al. In contrast, only patients with no more than one line of prior salvage therapy were included in the comparisons by Song et al., which avoided such bias.

Second, Proskorovsky et al. failed to address the differences in the definitions of outcomes between INO-VATE-ALL and TOWER. The outcomes that were compared in Proskorovsky et al. included CR or CR with incomplete hematologic recovery (CR/CRi), HSCT, OS, and EFS. The definition of CRi was inconsistent between the two trials. In INO-VATE-ALL, CRi was defined as CR except with absolute neutrophil counts (ANC) < 1000/μL and/or platelet ≤ 100,000/μL. In TOWER, CRi was defined as CR except with incomplete recovery of peripheral blood counts: ANC > 1000/μL or

J. Song · J. Xie
Analysis Group, Inc., Los Angeles, CA, USA

W. Gao (✉)
Analysis Group, Inc., Boston, MA, USA
e-mail: wei.gao@analysisgroup.com

S. Tiwana
Amgen Inc., Thousand Oaks, CA, USA

platelet $> 100,000/\mu\text{L}$. The definition of CRi in TOWER was more stringent and led to a lower rate of CR/CRi compared to INO-VATE-ALL. This can be observed through the CRi rate for the standard of care arm, i.e., 4.5% in TOWER and 11.9% in INO-VATE-ALL. Therefore, Proskorovsky et al. may bias in favor of InO by comparing CR/CRi using inconsistent definitions.

The comparison of outcome HSCT in Proskorovsky et al. was also questionable. From a clinical perspective, HSCT is driven by multiple factors beyond the effect of the investigational treatment, such as the availability of appropriate donors and the type of transplant. Without being able to adjust for those factors or having a head-to-head controlled study, the HSCT rate was not comparable between the INO-VATE-ALL and TOWER trial populations.

For the comparison of EFS in Proskorovsky et al., sensitivity progression-free survival (PFS) was used as a proxy for EFS given that EFS was not directly recorded in INO-VATE-ALL. Data on the sensitivity PFS were not included in prior publications of INO-VATE-ALL. Based on Proskorovsky et al., the sensitivity PFS was defined as the time from randomization to the earliest of death or progressive disease (either objective progression or relapse from CR/CRi). For subjects who did not achieve CR/CRi per investigator's assessment, the sensitivity PFS was assigned as 0. On the other hand, the EFS in TOWER was defined as the time from randomization to the earliest of death or hematological or extramedullary relapse after achieving CR/CR with partial hematologic recovery/CRi. It is unclear how the differences between the sensitivity PFS in INO-VATE-ALL and the EFS in TOWER impact the outcomes.

Third, Proskorovsky et al. matched on different subsets of treatment effect modifiers for different outcomes without providing convincing rationale. For example, duration of first remission, prior number of salvage therapies (0 or 1), and maximum of central/local bone marrow blasts were considered as treatment effect modifiers and were matched for the outcome CR/CRi but not for HSCT, OS, and EFS. Without evidence supporting the selection of effect modifiers for each outcome and the

discrepancy in the selection of effect modifiers across outcomes, their matching strategies may introduce bias and raise concerns regarding the weighted outcomes. When there is limited knowledge or data on what variables are effect modifiers, especially for new treatments, it is critical to match all baseline characteristics that are available in both trials in order to avoid potential bias due to incorrect selection of baseline characteristics. Even when there is clinical rationale that only a subset of baseline characteristics are effect modifiers, an analysis adjusting for all available baseline characteristics can also provide important information on assessing potential bias.

Despite the increasing popularity of MAIC, there is a need to fully understand the appropriate ways to apply this method [4–6]. As discussed above, there are several key methodological limitations in Proskorovsky et al. that led to biases in favor of InO, including failures to match key baseline characteristics, inconsistent definitions of outcomes, and the lack of evidence supporting their selection of matching variables. The results and the conclusion of this study should be interpreted with caution. Song et al., on the other hand, conducted the MAIC analyses more rigorously and addressed these limitations with a more comprehensive and less selected approach to the relevant parameters. Therefore, it provided more objective results for the comparison of blinatumomab and InO.

ACKNOWLEDGEMENTS

Funding. Amgen funded and participated in all aspects of the preparation of this letter. No payment was received by the journal for the publication of this letter.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this letter, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Disclosures. Jinlin Song is an employee of Analysis Group, Inc., which has received funding from Amgen to conduct this analysis. Wei Gao is an employee of Analysis Group, Inc., which has received funding from Amgen to conduct this analysis. Jipan Xie is an employee of Analysis Group, Inc., which has received funding from Amgen to conduct this analysis. Simran Tiwana is an employee of and stockholder in Amgen Inc.

Compliance with Ethics Guidelines. This article is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

Peer Review. Please note, contrary to the journal's standard single-blind peer review process, as a letter this article underwent review by a member of the journal's Editorial Board.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and

your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Song J, et al. Matching-adjusted indirect comparison of blinatumomab vs. inotuzumab ozogamicin for adults with relapsed/refractory acute lymphoblastic leukemia. *Adv Ther.* 2019;36(4):950–61.
2. Proskorovsky I, et al. Indirect treatment comparison of inotuzumab ozogamicin versus blinatumomab for relapsed or refractory acute lymphoblastic leukemia. *Adv Ther.* 2019;36(8):2147–60.
3. Kantarjian H, et al. Blinatumomab versus chemotherapy for advanced acute lymphoblastic leukemia. *N Engl J Med.* 2017;376(9):836–47.
4. Signorovitch JE, et al. Comparative effectiveness without head-to-head trials. *Pharmacoeconomics.* 2010;28(10):935–45.
5. Signorovitch JE, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health.* 2012;15(6):940–7.
6. Gao W, et al. Matching adjusted indirect comparisons: sensitivity analyses and graphical diagnostics. Indirect comparisons of treatment effects for clinical regulatory and health economic evaluations (topic contributed papers), joint statistical meeting, August 2, 2018; Vancouver.