Response to Letter to the Editor Regarding: Indirect Treatment Comparison of Inotuzumab Ozogamicin versus Blinatumomab for Relapsed or Refractory Acute Lymphoblastic Leukemia

Irina Proskorovsky · Erik Vandendries · Véronique Pagé · Joseph C. Cappelleri · Matthias Stelljes

Received: November 26, 2019 / Published online: December 14, 2019 © The Author(s) 2019

Dear Editor,

My colleagues and I am pleased to submit our response to the Song et al. [1] letter to the editor, which discussed perceived limitations and weaknesses of our comparison of inotuzumab ozogamicin (InO) vs. blinatumomab (Blina) for relapsed/refractory acute lymphoblastic leukemia (ALL).

ALL is a rare hematologic malignancy with a high rate of disease recurrence. No head-to-head studies have compared InO with Blina in adults with relapsed or refractory ALL; therefore, indirect treatment comparisons were used to assess the relative efficacy of these new therapies.

In our analyses, we could not fully adjust for the number of salvage therapies, but we did adjust for Philadelphia chromosome-positive (Ph+) B precursor status. Furthermore, our analysis provides a comparison of additional important outcomes, such as remission rate [defined as complete remission (CR), complete remission with partial hematologic recovery (CRh), or complete remission with incomplete hematologic recovery (CRI)], hematopoietic stem cell transplantation (HSCT), and event-free survival (EFS). This is in contrast with the analysis in Song et al., which compared only overall survival (OS) and CR rates. CR and CRI are important measures of efficacy as patients can achieve minimal residual disease (MRD) negativity and proceed to HSCT, the only potentially curative treatment modality available, thus providing the benefit of long-term survival.

Our detailed responses to each concern raised by Song et al. in the Letter to the Editor are included below.

For the first point, “the discrepancy in the number of prior salvage therapies is a fundamental difference between the two trial populations and without matching the number of prior salvage therapies, bias in favor of InO was introduced in Proskorovsky et al.,” we acknowledge this limitation in several places in the manuscript. The number of prior salvage therapies is a prognostic factor and potential treatment effect modifier. However, since our analyses were anchored, only a treatment effect modification would introduce a potential bias because of the inability to adjust for two or more prior salvage therapies [2]. Forest plots for
OS and the remission rate from the TOWER publication showed that, for OS, the effect of Blina vs. standard of care (SOC) was weaker in third-line plus patients [i.e., the hazard ratio (HR) and 95% confidence interval (CI) were 0.60 (0.39–0.91) for the first line, 0.59 (0.38–0.91) for the second line, and 1.13 (0.64–1.99) for the third line]. For remission rate, the opposite was observed; the odds ratio (OR; 95% CI) for the remission rate for Blina vs. SOC was 2.03 (1.08–3.80) in first line, 3.37 (1.35–8.38) in second line, and 4.10 (1.11–15.12) in third-line salvage.

The effect of Blina vs. SOC for EFS and HSCT in first and later salvage in TOWER was recently published by Dombret et al. [3]. The results showed a stronger treatment effect for Blina vs. SOC for EFS and HSCT for patients in second-line or later salvage. These findings suggest that failing to adjust for two or more lines of prior salvage therapies in the indirect treatment comparison (ITC) analyses by Proskorovsky et al. could have led to an overestimation of the treatment effect for OS and an underestimation of the treatment effect for the remission rate, EFS, and HSCT for InO vs. Blina.

Furthermore, the Song et al. letter did not mention that there were 49 (15.0%) patients with Ph + B precursor ALL included in the INO-VATE-ALL study—no such patients were included in the TOWER study. This discrepancy is an important difference between the two trial populations as INO-VATE-ALL results showed a stronger treatment effect for InO vs. SOC in Philadelphia chromosome-negative (Ph−) compared with Ph + patients. Without accounting for this difference, bias in favor of Blina may have been introduced in Song et al.

The second point raised by Song et al. was in regard to inconsistencies in the definition of remission in TOWER and INO-VATE-ALL. In our original analyses, we compared the rate of CR/CRi from INO-VATE-ALL to the rate of CR/CRi/CRh in TOWER. The proportion of patients in the SOC arm of INO-VATE-ALL achieving CRi [defined as CR except with absolute neutrophil counts (ANC) < 1000/μl and/or platelet counts < 100,000/μl] was 14.8% compared with 9% achieving CRi/CRh in the SOC arm of TOWER.

Nevertheless, in preparing this response letter, we conducted additional analyses in which we re-defined the remission rate in INO-VATE-ALL to match TOWER (i.e., CR was defined as platelet count of > 100,000/μl and ANC of > 1000/μl, CRi was defined as CR except with incomplete recovery of peripheral blood counts: platelet count of > 100,000/μl or ANC of > 1000/μl, and CRh was defined as platelet count of > 50,000/μl and ANC of > 500/μl). The observed remission rates in INO-VATE-ALL—per TOWER definition (i.e., CR/CRi/CRh)—in the new analyses were 72.6% and 27.2% for InO and SOC, respectively, compared with the previously published remission rates of 73.8% and 30.9%. The newly estimated treatment effect from the matching-adjusted indirect comparison (MAIC) for InO vs. Blina, expressed as OR (95% CI), was 2.94 (1.17–7.41) (Table 1) compared with 2.81 (1.12–7.05) in the original publication from Proskorovsky et al. This new, supplemental finding is relevant and supports our previous claim that a small discrepancy in the original definition of CRi in INO-VATE-ALL vs. CRi/CRh in TOWER was unlikely to bias the relative effect of InO vs. SOC, and thus the MAIC-derived treatment effect of InO vs. Blina, supporting the robustness of our results.

Similarly, for the comparison of EFS, since our analyses were anchored, the difference in the definitions is unlikely to bias the derived relative treatment effect for InO vs. Blina, given the same considerations as for remission rate.

Song et al. also questioned the comparison of HSCT rates by Proskorovsky et al. stating that HSCT is driven by multiple factors beyond the effect of the investigational treatment. We acknowledge that subsequent therapies, including HSCT, can be driven by factors beyond the effect of the investigational treatment. However, the most important point of anchored indirect comparison is not a comparison of actual rates of HSCT between InO vs. Blina, but relative treatment effects between studies (i.e., InO vs. SOC in INO-VATE-ALL compared with Blina vs. SOC in TOWER). As such, both arms within the same trial would be impacted in a similar manner by the multiple factors mentioned by Song et al. (i.e., the availability of appropriate donors and the type
Therefore, if HSCT was not driven by treatment effect in INO-VATE-ALL, we would then see similar rates of transplant between InO and SOC with an OR of 1; however, such is not the case in INO-VATE-ALL [48.2% and 22.2% received HSCT in InO and SOC, respectively, resulting to an OR (95% CI) of 3.25 (2.01–5.26)].

In contrast, HSCT rates in TOWER were similar between Blina (24.0%) and SOC (23.9%), with an OR (95% CI) of 1.01 (0.62–1.63). A key function of treatment in relapsed or refractory ALL is the ability of patients to bridge to transplant; among those transplanted, InO showed better efficacy. As discussed in our original publication, in INO-VATE-ALL, the separation in OS curves between InO and SOC accelerated at 15 months, indicating a longer term survival benefit, consistent with the higher HSCT rates among patients in the InO arm. For the TOWER study, the OS curves for Blina and SOC converged around this same time point for the intention-to-treat population, consistent with the identical HSCT rates observed among patients in the Blina and SOC arms. However, due to limited follow-up in the TOWER study, making the assessment of long-term OS data is difficult.

Furthermore, for the comparison of OS, similar considerations apply (i.e., OS can be impacted by subsequent therapies and other factors outside of the trial), but Song et al. conducted not only anchored but also unanchored MAIC analyses for OS.

As to the third point raised by Song et al.—adjusting only for select variables in anchored MAIC for each outcome in the analyses by Proskorovsky et al.—the selection of potential treatment effect modifiers was carefully approached; we examined available forest plots from the INO-VATE-ALL and TOWER studies and obtained clinical input. In anchored analyses, adjusting for all available patient characteristics (i.e., prognostic factors and treatment effect modifiers) can lead to a substantial reduction in the effective sample size (ESS), which translates to a loss of precision of the estimated treatment effect without an additional bias reduction [4]. The National Institute for Health and Care Excellence Decision Support Unit recommendation clearly states that for anchored comparison, only treatment effect

| Table 1 | Summary of MAIC results with remission rate as per TOWER definition |
|---------|------------------------------------------------------------------|
|         | ITT Adjusted ITCs | Adjusted ITCs |
|         | TOWER Blina vs. SOC | INO-VATE-ALL InO vs. SOC | InO vs. SOC, NMA | InO vs. SOC, MAIC |
| Remission rate per TOWER definitions | | | |
| OR (95% CI) | 2.40 (1.51–3.80) | 7.09 (4.36–11.55) | 7.06 (3.17–15.74) | 2.96 (1.51–5.77) | 2.94 (1.17–7.41) |
| Difference (95% CI) | 19.28 (9.90–28.67) | 45.40 (35.73–55.07) | 45.28 (29.42–61.15) | 26.12 (12.60–39.60) | 26.00 (7.60–44.40) |
| Remission rate published by Proskorovsky et al. | | | |
| OR (95% CI) | 2.40 (1.51–3.80) | 6.30 (3.89–10.21) | 6.75 (3.04–14.95) | 2.63 (1.35–5.12) | 2.81 (1.12–7.05) |
| Difference (95% CI) | 19.28 (9.90–28.67) | 42.92 (33.12–52.71) | 44.40 (28.42–60.38) | 23.64 (10.10–37.20) | 25.12 (6.60–43.70) |

Blina blinatumomab, CI confidence interval, InO inotuzumab ozogamicin, ITC indirect treatment comparison, ITT intent to treat, MAIC matching-adjusted indirect comparison, NMA network meta-analysis, OR odds ratio, SOC standard of care.
modifiers should be included in the adjustment. In these analyses, adjusting for the full set of patient characteristics (i.e., prognostic factors plus treatment effect modifiers) led to 76% and 73% reduction in the ESS from the original sample size for InO (ESS = 40) and SOC (ESS = 44) in MAIC analyses, making the resulting comparisons between InO and Blina dependent on a small number of individuals in the NOVATE-ALL study and making the comparisons unstable.

However, in response to this letter, we conducted additional exploratory analyses for OS and EFS that adjusted for additional baseline patient characteristics (i.e., age, region, previous allogeneic HSCT, Ph + B precursor, duration of first remission, salvage treatment phase, maximum of central/local bone marrow blasts, hemoglobin, and platelets). Results of these analyses are presented in Table 2; they show that the point estimates for the treatment effect of InO vs. Blina for EFS and OS were similar to the previously published results. However, as expected, including additional variables in the adjustment reduced the precision and widened the 95% CIs around the estimated treatment effects.

In conclusion, the analyses by Proskorovsky et al. are not without limitations; however, they are articulated in the discussion section of the paper. In our opinion, the questions raised by Song et al. regarding certain aspects of the ITC analyses are specious. In addition, MAIC analyses by Song et al. are incomplete. They compare only select outcomes (i.e., CR and OS) and omitted comparisons of other important clinical end points measured in both studies, such as remission rate and HSCT, without robust rationale.

Analyses by Proskorovsky et al. are robust and provide a full comparison of end points that are relevant to patient outcomes in relapsed or refractory ALL, supplying objective results for the comparison of InO and Blina.

We believe our responses are informative and useful in rebutting Song et al.’s assertions. All inquiries may be addressed to me as the Corresponding Author. Thank you for

| Table 2 ITC results of inotuzumab vs. blinatumomab depending on level of adjustment |
|---------------------------------|---------------------------------|---------------------------------|
|                                 | InO vs. Blina                   | InO vs. Blina                   |
|                                 | MAIC                            | STC                            |
|                                 | EM adjusted                     | EM adjusted                    |
|                                 | EM + PF adjusted                | EM + PF adjusted               |
| EFS                             |                                 |                                |
| HR (95% CI)                     | 0.73 (0.47–1.13)                | 0.76 (0.44–1.34)               |
| RMST difference (95% CI)        | 3.60 (1.37–5.83)                | 3.07 (0.40–5.74)               |
| RMST ratio (95% CI)             | 2.27 (1.09–4.74)                | 2.23 (0.91–5.47)               |
| OS                              |                                 |                                |
| Overall HR (95% CI)             | 0.96 (0.61–1.50)                | 0.90 (0.51–1.58)               |
| 0–15 month HR (95% CI)          | 1.01 (0.64–1.62)                | 0.92 (0.51–1.64)               |
| 15+ month HR (95% CI)           | 0.24 (0.02–3.17)                | 0.32 (0.02–5.66)               |
| RMST difference (95% CI)        | 0.30 (–2.86 to 3.46)            | 0.98 (–2.96 to 4.92)           |
| RMST ratio (95% CI)             | 1.02 (0.72–1.43)                | 1.09 (0.72–1.65)               |

*Blina* blinatumomab, *CI* confidence interval, *EFS* event-free survival, *EM* effect modifier, *HR* hazard ratio, *InO* inotuzumab ozogamicin, *MAIC* matching-adjusted indirect comparison, *NA* not applicable, *OS* overall survival, *PF* prognostic factor, *RMST* restricted mean survival time, *STC* simulated treatment comparison

△Adis
consideration of our work for publication in Advances in Therapy.

Sincerely,
Irina Proskorovsky, MSc
Senior Research Leader, Evidence Synthesis, Modeling & Communication

ACKNOWLEDGEMENTS

Funding. Pfizer 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. Irina Proskorovsky is a salaried employee of Evidera, which has received funding from Pfizer Inc. to conduct this analysis and develop associated publications. Véronique Pagé is a salaried employee of Evidera, which has received funding from Pfizer Inc. to conduct this analysis and develop associated publications. Erik Vandendries is an employee of and shareholder in Pfizer Inc., which funded this study. Joseph C. Cappelleri is an employee of and shareholder in Pfizer Inc., which funded this study. Matthias Stelljes reports research support from Pfizer during the conduct of the study and other consulting fees and honoraria from Pfizer, Amgen, and Jazz Pharmaceuticals.

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 response 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. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. Med Decis Mak. 2018;38(2):200–11.

3. Dombret H, Topp MS, Schulh AC, et al. Blinatumomab versus chemotherapy in first salvage or in later salvage for B-cell precursor acute lymphoblastic leukemia. Leuk Lymph. 2019;60(9):2214–22.

4. Phillippo DM, Ades AE, Dias S, et al. NICE DSU Technical Support Document 18: methods for population-adjusted indirect comparisons in submissions to NICE. Report by the Decision Support Unit. 2016.