Indirect treatment comparisons: how to MAIC it right?

Emmanuel Bachy

Hospices Civils de Lyon, Hematology Department, Lyon, France

Correspondence: E. Bachy
emmanuel.bachy@chu-lyon.fr

Received: January 11, 2024.
Accepted: February 8, 2024.
Early view: February 15, 2024.

https://doi.org/10.3324/haematol.2023.284534

As more phase II non-comparative clinical trial data are generated with new emerging treatment options, indirect cross-trial and real-life comparisons have grown exponentially. Sometimes they aim to provide anticipated information while waiting for ongoing phase III trials, while most often they try to substitute for head-to-head comparisons that will never be performed by pharma companies or academic consortia. In this issue of *Haematologica*, Maurer and colleagues performed a matching-adjusted indirect comparison (MAIC) between patients from the Lymphoma Epidemiology for Outcomes (LEO) Consortium for Real World Evidence (CReWE) and from a phase I/II study of mosunetuzumab, a CD20xCD3 T-cell engaging bispecific antibody, as a single agent (GO29781 – *clinicaltrials.gov 02500497*).[1,2] Phase III randomized studies are of the highest quality standards for treatment comparison because random allocation can control both for measured confounders (e.g., age, sex, disease stage, treatment line, performance status) and potential unmeasured confounders (e.g., unquantifiable belief by the physician that a particular patient would benefit from a given drug or unmeasured socio-economic determinants). Amidst both end of the spectrum ranging from basic unadjusted or unmatched comparisons and the phase III gold standard stand various statistical approaches trying to mitigate these potential biases (Figure 1A). A MAIC is usually considered when individual patient data (IPD) are available for one group of patients while only aggregated data are available for the comparator (e.g., median or mean values, interquartile range). This is usually the case when trial sponsors (usually pharma companies) want to compare the outcome for patients treated with the product they develop with a competitor drug for which IPD are not available. This is also the case here where an academic consortium compares real-world IPD from a large cohort of patients to a given trial. Basically, a MAIC relies on selecting patients and pondering outcome according to their IPD characteristics to render them as close as possible to aggregated data available from a trial publication (Figure 1B).[3,4] MAIC is usually considered as providing a poor confounding control as opposed to comparisons where IPD are available for both treatment groups (e.g., using propensity score-based matched comparisons or adjustments) but is usually the only option when only aggregated data are available for one treatment group. Furthermore, and like any comparison except randomized trials, MAIC cannot control for unmeasured biases that could confound the comparison of outcomes.

In follicular lymphoma, a disease still considered as incurable for most patients, overall survival is now believed to extend beyond a median 20 years with a significant fraction of patients dying from non-lymphoma-related cause.[5,6] However, disease outcome is highly heterogeneous with some patients experiencing poor survival despite the theoretical indolent nature of the disease.[7,8] Therefore, taking into account disease heterogeneity between FL patient cohorts is critical for cross-trial and other treatment comparisons. The study conducted by Maurer and colleagues shows that after careful weighting of patient data from the LEO CReWE to match clinical characteristics with those from the GO29781 study, overall (80% *vs*. 73%) and complete (60% *vs*. 53%) responses rates (ORR and CRR) were found slightly higher in the trial cohort. Despite these higher response rates, no obvious difference was observed regarding 12-month progression-free survival (PFS) (60% *vs*. 58% in the LEO *vs*. trial cohorts, respectively). However, and as stated by the authors, response status assessment differences between trial (with frequent imaging exams) and routine practice are a major bias for PFS measurement possibly penalizing response duration in the mosunetuzumab trial. Furthermore, 40% of patients in the LEO cohort were treated as part of a clinical trial for the selected index therapy, making it quite different from standard of care (SOC) strategies outside of specific tertiary care centers. Altogether, on one hand, prognosis is likely overestimated in the LEO cohort in contrast with routine practice because of patient selection. But, on the other hand, 40% of them have probably been followed more closely to what

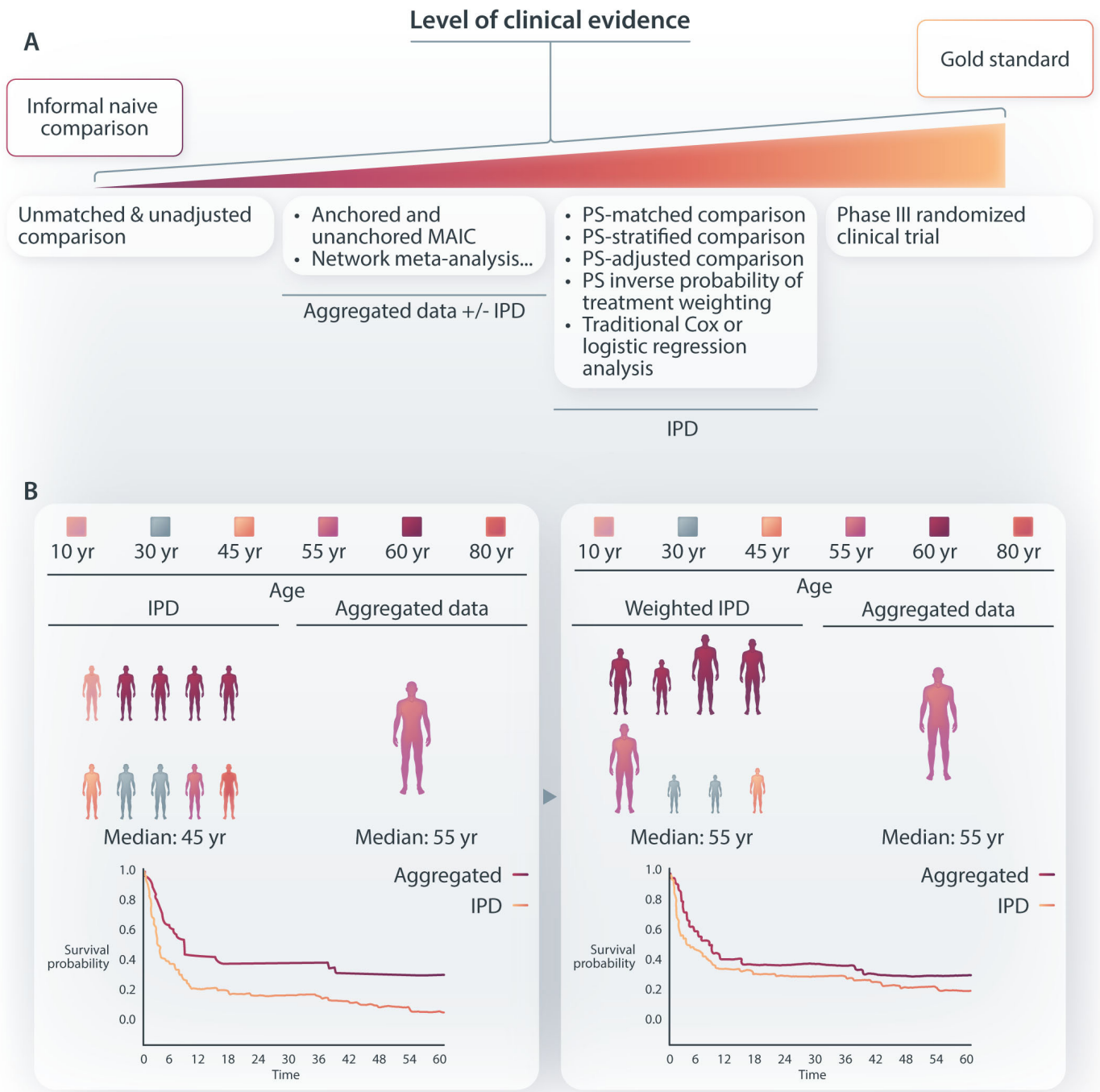

Figure 1. Principle and level of clinical evidence of matching-adjusted indirect comparison. (A) Level of evidence of various treatment comparisons from unadjusted and unmatched comparison (poorest level of evidence) to randomized treatment allocation (gold standard). The list of various statistical approaches to make patient population as comparable as possible is not exhaustive. (B) Matching-adjusted indirect comparison (MAIC) allows for the comparison between aggregated patient data (e.g., based on patient characteristics from a trial publication) and individual patient data (from another trial or a real-life cohort or any other source of individual data). Basically, by removing or pondering patient characteristics to closely match final aggregated data (e.g., patient median age depicted here), the final group of patients from the cohort with available individual patient data (IPD) is rendered as similar as possible to the cohort for which only aggregated data are available (e.g., here, similar median age). This is performed for all variables that are considered as critical confounders for treatment comparison. Finally, outcome is compared in the 2 matched populations; here, the prognosis of the IPD is depicted as better after matching (blue line), while, by definition, the survival of the aggregated data population is left unchanged after matching (red line). Depicted data and survivals are for illustration only and are not based on true or relevant values or weights. PS: propensity score; yr: year.

was performed in the GO29781 trial and more accurately than what is usually performed outside a trial setting. It is also important to notice that among those patients from the LEO cohort, 11 received either another bispecific antibody or a chimeric antigen receptor T-cell therapy and 38 another novel agent as monotherapy or in combination with an anti-CD20 monoclonal antibody. This explains, at least in part, the good outcome of patients from the LEO cohort and is critical to appreciate the small differences observed in the MAIC according to the various scenarios tested.

Beyond information about efficacy of this bi-specific antibody, many lessons are to be learned from this study for any hematologist aiming at developing a critical appraisal of MAIC conclusions as they become increasingly used to support early phase study results for therapies still under development. First, results are strongly dependent upon which scenario and constraints are applied to the model. Second, a careful analysis of which variables have

been selected (or omitted), and the reason why they have been incorporated or not, should be thoroughly examined. Third, inclusion/exclusion criteria, outcome definition, balance in the average baseline characteristics after matching and sensitivity analyses should be key in interpreting results of MAIC. Finally, extreme caution should always be applied when drawing conclusions based on such indirect comparisons and one should remember that only well-controlled randomized study can balance unmeasured confounders. Regarding all these critical points, this academy-concepted study will be highly informative to the reader beyond results of the comparison itself and despite all limitations of the statistical approach.

Disclosures

Honorarium/Ad Board from Kite, a Gilead Company, Bristol Myers Squibb, Novartis, Pfizer, Incyte, ADC Therapeutics, Roche, Takeda; fees from Kite, a Gilead Company, Bristol Myers Squibb, Novartis, Pfizer; funding from Amgen, BMS.

References

1. Maurer MJ, Casulo C, Larson MC, et al. Matching-adjusted indirect comparison from the Lymphoma Epidemiology of Outcomes Consortium for Real World Evidence (LEO CReWE) study to a clinical trial of mosunetuzumab in relapsed or refractory follicular lymphoma. Haematologica. 2024;109(7):2177-2185.
2. Budde LE, Sehn LH, Matasar M, et al. Safety and efficacy of mosunetuzumab, a bispecific antibody, in patients with relapsed or refractory follicular lymphoma: a single-arm, multicentre, phase 2 study. Lancet Oncol. 2022;23(8):1055-1065.
3. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. Pharmacoeconomics. 2010;28(10):935-945.
4. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940-947.
5. Batlevi CL, Sha F, Alperovich A, et al. Follicular lymphoma in the modern era: survival, treatment outcomes, and identification of high-risk subgroups. Blood Cancer J. 2020;10(7):74.
6. Sarkozy C, Maurer MJ, Link BK, et al. Cause of death in follicular lymphoma in the first decade of the rituximab era: a pooled analysis of French and US cohorts. J Clin Oncol. 2019;37(2):144-152.
7. Casulo C, Byrtek M, Dawson KL, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the National LymphoCare Study. J Clin Oncol. 2015;33(23):2516-2522.
8. Maurer MJ, Bachy E, Ghesquieres H, et al. Early event status informs subsequent outcome in newly diagnosed follicular lymphoma. Am J Hematol. 2016;91(11):1096-1101.