

EDITORIAL

OPEN



Opportunities and risks associated with external comparator cohorts in clinical drug development in hematology

© The Author(s) 2025

Blood Cancer Journal (2025) 15:199; <https://doi.org/10.1038/s41408-025-01426-2>

In this issue of *Blood Cancer Journal*, Letaillieur et al. explore the use of real-world data (RWD) as a synthetic control arm (SCA) in older patients with diffuse large B-cell lymphoma (DLBCL) and propose this innovative design for future trials [1]. To date, the use of RWD for assessing therapeutic benefit in pre-marketing drug development has largely been confined to benchmarking early phase studies. Such analyses often fail to account for confounders, as shown in a review of 40 EU conditional marketing authorizations where unadjusted, side-to-side comparisons were often used to support therapeutic advantage [2]. By optimizing the use of RWD, earlier certainty about clinical benefit could reduce the risk of marketing withdrawals due to disappointing confirmatory studies. However, caution is required. The risk associated with using RWD in early-phase development—for example, selecting the most promising drug candidate for clinical development—is lower than in confirmatory settings, where biased results could lead to large-scale exposure of patients to ineffective treatments.

The SCA in the study by Letaillieur et al. was drawn from 170 patients from a historical clinical trial (CT, LNH09-7B) and RWD (REALYSA). Its performance was then tested in a hypothetical setting, replacing the internal control arm from SENIOR in a simulated study. SENIOR randomized 249 DLBCL patients ≥ 80 years 1:1 to R-miniCHOP plus placebo or lenalidomide (R²-miniCHOP) with identical 2-year overall survival (OS) of 66% in both arms (95% CI 0.65–1.5) [3]. In the simulation, inverse probability weighing with propensity scores was used to balance confounders between 104 available patients from the experimental arm of SENIOR and the SCA. Using this approach, the authors reached the same conclusion: that R²-miniCHOP was not superior to R-miniCHOP. Results were also similar using RWD alone (REALYSA) to generate the SCA, leading the authors to conclude that careful selection and appropriate analytics may allow RWD to replace internal control arms in selected situations.

While these results are promising, more work is needed before SCAs can adeptly replace internal control arms without increasing uncertainty. With a reported HR of 0.743 (95%CI 0.49–1.12) between the experimental arm of SENIOR and the SCA, the lack of observed difference could be related to statistical power. In fact, based on the included patient numbers and with a 2-year OS of 58% for the SCA vs. 66% for the experimental arm of SENIOR, a post-hoc calculation ($\alpha = 0.05$) yields power of just 30%, assuming the HR of 0.743 is true. In the head-to-head comparison of the SCA and the control arm of SENIOR, the absence of statistical significance at the $P < 0.05$ level also does not fully exclude a clinically relevant difference. With an HR between the internal arm of SENIOR and the SCA of 0.79 (95%CI 0.52–1.20), OS between arms could differ substantially.

Received: 24 August 2025 Revised: 17 October 2025 Accepted: 11 November 2025
Published online: 19 November 2025

Letaillieur et al.'s analyses provide valuable insights needed for the future use of SCA generated from RWD. Considering that with few exceptions [4], survival gains for novel therapies of >10% are rarely seen in DLBCL, there is reasonable concern that any potential bias could risk false outcomes (positive or negative) with the use of SCAs in pivotal studies. Investigators and regulatory authorities need to better understand how to quantify this risk and the conditions required to conclude that the performance of an SCA is equivalent to an internal control. The U.S. Food and Drug Administration and European Medicines Agency guidelines in development for external comparator cohort studies clearly demonstrate that regulatory agencies expect increasing use of innovative trial designs which incorporate RWD [5, 6]. From a societal perspective, these types of trials are highly warranted as they can address costs, timelines, and logistical challenges in research.

Sources of bias in RWD are well-described. Clinical trial participants are likely to have different health-related behaviors compared to general populations. Compared to RWD, clinical trials may also be less likely to include patients with rapidly progressive disease, consistently associated with poorer survival in aggressive lymphoma [7]. Cancer trials are also subject to increasingly selective eligibility criteria [8], with trial ineligibility reliably associated with inferior survival outcomes across lymphoma subtypes [9, 10]. These factors all represent challenges in identifying a relevant external control group from the 'real-world'. While methods like inverse probability weighting do control for measurable, known confounders, unmeasured, uncontrolled confounders will persist. Categorization of continuous variables (e.g., defining lactate dehydrogenase as normal vs. elevated) can also result in loss of informational value and suboptimal matching. Incentivizing the prospective and accurate collection of granular RWD will help to mitigate these sources of bias, as will the promotion of more inclusive and pragmatic clinical trials.

In addition to the above, there are other less quantifiable risks. Typically, an SCA would be considered for a single-arm trial (SAT) that has impressive efficacy results. However, impressive efficacy can sometimes be the result of selection bias and randomness. If only the SATs with the best observed outcomes are considered for matching with an SCA in a post-hoc manner, this would be a source of bias. Optimally, the plans for SCA should be detailed prospectively prior to knowing SAT results in a statistical analysis plan (SAP). The SAP should include matching variables as well as important prognostic covariates included in adjusted efficacy analyses. For full transparency, the SAP should also describe which RWD sources will be used for the SCA. Selection of those RWD sources should preferably be done with blinding to outcome data from the RWD source to avoid bias related to the Sponsor's knowledge about patient outcomes in the SAT. By requiring preplanned use of an SCA, including how to generate the SCA and adjust for confounders, the risk of systematic bias from post-hoc, data-driven decisions will be lowered.

Endpoint selection when using RWD also merits caution. Without protocolized assessment timepoints, more toxic and disease-related events will be detected in patients undergoing more frequent physician visits and investigations. In the real-world setting, endpoints based on acute events such as time-to-discontinuation (TTD), time-to-next-therapy (TTNT), or OS may therefore be more bias-resistant than endpoints reliant on protocolized clinical or radiologic disease assessment, such as progression-free survival (PFS). However, safety assessment remains challenging; despite a substantially higher rate of Grade 3–4 adverse events for the experimental arm of SENIOR (81% vs. 53%), no comparison of the toxicity rates between the experimental arm and SCA was attempted, underscoring the unresolved challenges of comprehensive benefit/risk assessments with SCAs.

Despite these limitations, there is significant potential to improve the efficiency and quality of evidence generation for regulatory decisions by harnessing RWD, especially in the current digital era of healthcare delivery. Robust RWD should serve as the ultimate control for assessing the impact of new interventions on outcomes in representative populations, particularly in rare diseases where large-scale data are time consuming to collect. RWD provides information largely absent from trials including cost effectiveness, healthcare utilization, and long-term follow up. Yet, in contrast to established global frameworks for transparent clinical trial data management supported by regulatory and research governing bodies, existing RWD policies regarding collection and sharing are highly varied across and within jurisdictions [11, 12]. This significantly diminishes RWD capability and speed of use by researchers and regulators. These barriers largely stem from privacy and legal data custody concerns [13]. Although legitimate, the focus on these mitigatable risks should be balanced against the ethical implications to both existing—and future—patients of failing to incorporate RWD in drug development pathways. Consumer feedback indicates many patients share researchers' concerns regarding data sharing barriers [14, 15].

Regulators must prioritize RWD science, such as that by Letaillier et al., and continue to work with registries globally to create clear guidance on acceptable RWD conduct. Only then can we harness the full benefits of RWD in drug development. Recognizing this need, we have worked with registry leads internationally to form the global Lymphoma Registry Alliance and bring together 50 registries from 30 countries to overcome some of these major issues faced.

Diva Baggio¹, Eliza A. Hawkes^{1,2,3},
Mikkel Runason Simonsen^{4,5} and
Tarec Christoffer El-Galaly^{1,6,7,8}✉

¹Department of Hematology, University College London Hospital, London, UK. ²Olivia Newton John Cancer Research Institute at Austin Health, Melbourne, VIC, Australia. ³School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia.

⁴Department of Hematology, Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark. ⁵Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark.

⁶Departments of Clinical Epidemiology, Molecular Medicine, and Hematology, Aarhus University Hospital, Aarhus, Denmark.

⁷Department of Clinical Medicine, Aarhus University, Aarhus, Denmark. ⁸Division of Clinical Epidemiology, Department of Medicine Solna, Karolinska Institutet, Solna, Sweden. ✉email: tarec.galaly@clin.au.dk

REFERENCES

1. Letaillier V, Chaillol I, Cherblanc F, Ghesquieres H, Peyrade F, Guidez S, et al. Synthetic control arm 1 from mixed clinical trials and real-world data from LYSA group for untreated diffuse large B cell lymphoma patients aged over 80 years: a bona fide strategy for innovative clinical trials. *Blood Cancer J.* 2025;15:190.
2. Lasch F, Carvalho JRB, Pothet C. Demonstration of Major Therapeutic Advantage From a Review of EU Conditional Marketing Authorizations in Oncology and Hematology. *Clin Pharm Ther.* 2025;117:1098–105.
3. Oberic L, Peyrade F, Puyade M, Bonnet C, Dartigues-Cuillères P, Fabiani B, et al. Subcutaneous Rituximab-MiniCHOP Compared With Subcutaneous Rituximab-MiniCHOP Plus Lenalidomide in Diffuse Large B-Cell Lymphoma for Patients Age 80 Years or Older. *J Clin Oncol.* 2021;39:1203–13.
4. Abramson JS, Ku M, Hertzberg M, Huang HQ, Fox CP, Zhang H, et al. Glocitamab plus gemcitabine and oxaliplatin (GemOx) versus rituximab-GemOx for relapsed or refractory diffuse large B-cell lymphoma (STARGLO): a global phase 3, randomised, open-label trial. *Lancet.* 2024;404:1940–54.
5. U.S. Food and Drug Administration. Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products (Draft Guidance) [Internet]. 2023 [cited 2025 Aug 19]. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-and-conduct-externally-controlled-trials-drug-and-biological-products>.
6. European Medicines Agency. Draft Concept Paper on the Development of a Reflection Paper on the Use of External Controls for Evidence Generation in Regulatory Decision-Making [Internet]. 2025 [cited 2025 Aug 19]. Available from: <https://www.ema.europa.eu/en/development-reflection-paper-use-external-controls-evidence-generation-regulatory-decision-making-scientific-guideline>.
7. Olszewski AJ, Ollila T, Reagan JL. Time to treatment is an independent prognostic factor in aggressive non-Hodgkin lymphomas. *Br J Haematol.* 2018;181:495–504.
8. Loh Z, Salvaris R, Chong G, Churilov L, Manos K, Barraclough A, et al. Evolution of eligibility criteria for diffuse large B-cell lymphoma randomised controlled trials over 30 years. *Br J Haematol.* 2021;193:741–9.
9. Simonsen MR, Haunstrup LM, Severinsen FT, Jensen RK, Brown PDN, Maurer MJ, et al. The impact of trial inclusion criteria on outcomes in DLBCL patients treated with R-CHOP in the first line: a Danish nationwide study. *Leuk Lymphoma.* 2024;65:2173–81.
10. Bennedsen TL, Simonsen MR, Jensen P, Brown P, Josefsson P, Khurana A, et al. Impact of Trial Eligibility Criteria on Outcomes of 1183 Patients With Follicular Lymphoma Treated in the Real-World Setting. *Eur J Haematol.* 2025;114:832–9.
11. Gisslander K, Mohammad AJ, Vaglio A, Little MA. Overcoming challenges in rare disease registry integration using the semantic web - a clinical research perspective. *Orphanet J Rare Dis.* 2023;18:253.
12. Andrew NE, Sundararajan V, Thrift AG, Kilkenny MF, Katzenellenbogen J, Flack F, et al. Addressing the challenges of cross-jurisdictional data linkage between a national clinical quality registry and government-held health data. *Aust N Z J Public Health.* 2016;40:436–42.
13. Bradley CJ, Penberthy L, Devers KJ, Holden DJ. Health Services Research and Data Linkages: Issues, Methods, and Directions for the Future. *Health Serv Res.* 2010;45:1468–88.
14. Hutchings E, Loomes M, Butow P, Boyle FM. A systematic literature review of health consumer attitudes towards secondary use and sharing of health administrative and clinical trial data: a focus on privacy, trust, and transparency. *Syst Rev.* 2020;9:235.
15. Varhol RJ, Man Ying Lee C, Hindmarsh S, Boyd JH, Robinson S, Randall S. Consumer attitudes, barriers and facilitators to sharing clinical data for research purposes: Results from a focus group synthesis. *Heliyon.* 2024;10:e34431.

AUTHOR CONTRIBUTIONS

DB, EAH, MRS, and TEG all contributed to conception, writing, and approval of the final manuscript.

COMPETING INTERESTS

DB: No COIs. EAH: EAH has received research funding from Bristol Myers Squibb/Celgene, Merck KGaA, AstraZeneca, TG therapeutics and F. Hoffmann-La Roche (all paid to institution); has acted as a consultant/advisor for F. Hoffmann-La Roche, Antengene, Bristol Myers Squibb, AstraZeneca, Novartis, Merck Sharpe Dohme, Specialised therapeutics, Sobi, Regeneron, Gilead and Abbvie; has acted as a speaker for Roche, AstraZeneca, Janssen, Regeneron, Abbvie and Genmab and received travel expenses from AstraZeneca and Abbvie. MRS: No COIs. TCEG: No COIs. The views expressed in this editorial are those of the authors and do not necessarily reflect the views or positions of any entities the authors are affiliated to or collaborate with.

DISCLAIMER STATEMENT

The views expressed in this editorial are those of the authors and do not necessarily reflect the views or positions of any entities the authors are affiliated to or collaborate with.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nd/4.0/>.

© The Author(s) 2025