CO213

Review of index line selection methods in hemato-oncology externally controlled trials: A secondary analysis of the Hermans et al. systematic review

Michael Wallington, BA[1], Iman Fakih, MSc[2], Yong Chen, PhD[3], Ann-Sophie Demers, MA[4], Yong Chen, PhD[5]

[1]Pfizer UK, London, United Kingdom, [2]STATLOG, Montreal, QC, Canada, [3]University of Pennsylvania, Philadelphia, PA, USA, [4]STATLOG, Québec, QC, Canada, [5]Pfizer, Collegeville, PA, USA

Objectives

- Our objective was to perform a secondary analysis of recent systemic review of hemato-oncology ECAs[1] and evaluate the index line selection methods used.

Background

- External control arms (ECAs) are being increasingly used as comparator groups for hemato-oncology clinical trials.
- However, little guidance exists on how to select an index treatment line (i.e., “time-zero”) for retrospective control patients.
- Selection of index line is an important consideration as it can introduce bias of efficacy estimates.

Methods

- This study was a secondary analysis of a 2024 systematic review by Hermans et al.,[1] which identified 32 real-world data (RWD)-derived hemato-oncology ECAs published 01 Jan 2000 – 23 Oct 2023.
- We imposed additional criteria to only select ECAs that can result in multiple eligible treatments lines:
  1. Patients are retrospectively selected into the ECA
  2. Patients are required to have at least 1 prior LOT or are relapsed/refractory (R/R)
  3. Authors must have access to individual-level data
- General study characteristics and details on index selection methodology were extracted and summarized.
- Index line selection methods used in the ECAs were visualized using schematics and classified as **“recommended”** and **“not recommended”** based on whether they introduce bias or not, per the literature.[2-8]
- The literature indicates that bias is introduced when line selection is predicated on the “future” knowledge of a patient’s number of lines, such as choosing the last line or a random line from all eligible lines.[5,8]

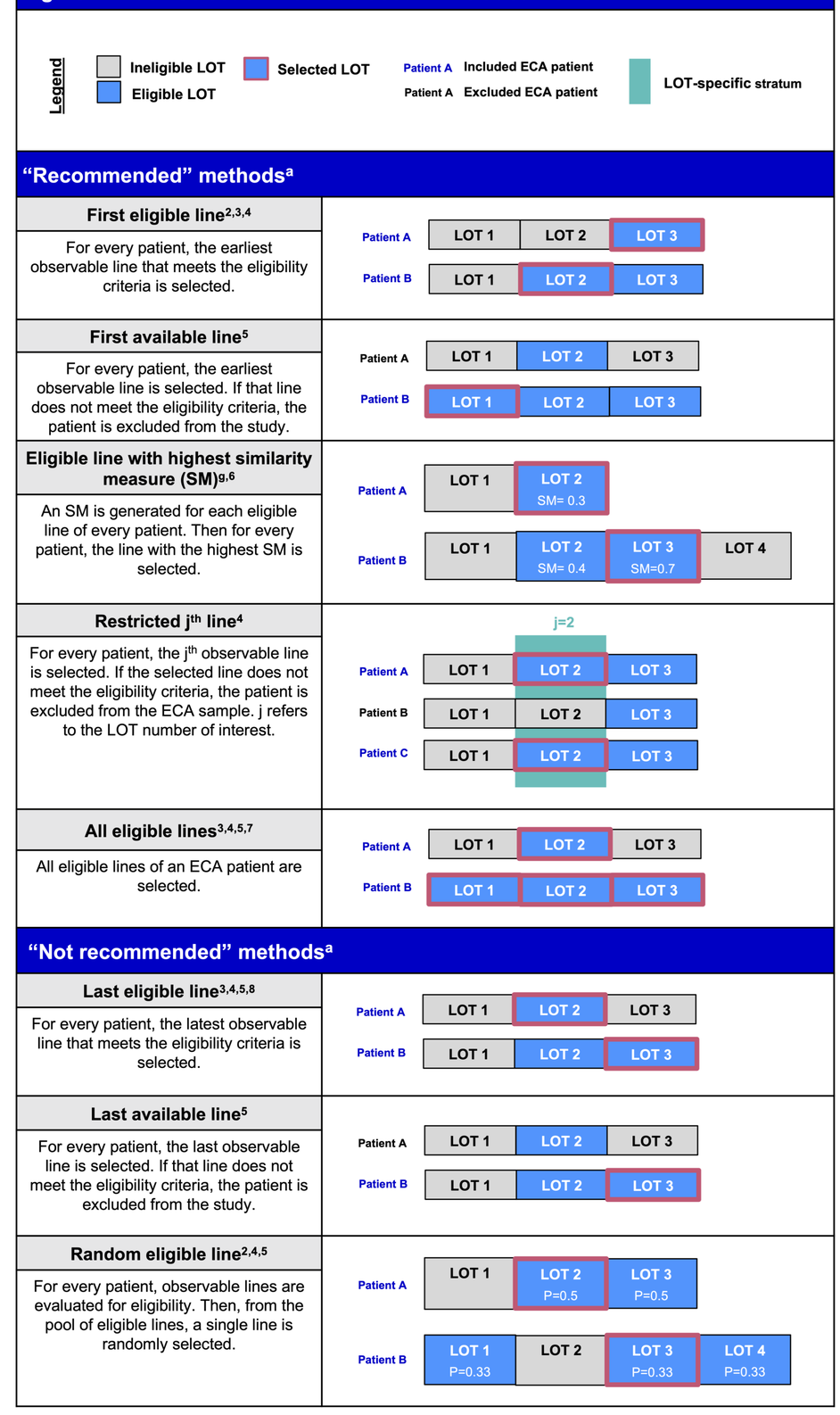
Results

- **Twenty-two** (68.6%) **ECAs** met the inclusion criteria. 9 ECAs were excluded due to treatment history criteria, and 1 ECA did not have individual-level data.
- Most ECAs (n=16; 72.7%) were published **in 2020 or later** (Table 1).
- Most frequent **RWD sources** were **chart review** (n=8; 36.4%) and **EMR** (n=7; 31.8%) (Table 1).
- All ECAs reported **time-to-event outcomes** (Table 1).
- **‘First eligible line’** (n=6; 27.3%) and **‘all eligible lines’** (n=6; 27.3%) were the most commonly used methods (Table 1, Figure 1).
- **Most** ECAs (n=15; 68.2%) used index line selection methods that were classified as **“recommended”**. Though **some** used **“not recommended”** methods (n=4; 18.2%) or were **unclear** about the method used (n=3; 13.6%) (Table 1, Figure 1).
- ECAs using “**recommended**” methods were more likely to **acknowledge** the index line selection **challenge** (66.7% vs 28.6%), and **cite** methodology **references** (53.3% vs. 14.3%) (Table 1, Figure 1).

Table 1. Characteristics of ECAs and index line selection methods by appraisal (N=22)

	All ECAs	ECAs by appraisal[a]	
		Recommended	Not Recommended/ Unclear
	N=22	N=15	N=7
Publication year, n (%)			
2019 or earlier	6 (27.3)	4 (26.7)	2 (28.6)
2020 or later	16 (72.7)	11 (73.3)	5 (71.4)
Disease, n (%)			
MM	7 (31.8)	7 (46.7)	0 (0.0)
NHL[b]	10 (45.5)	6 (40.0)	4 (57.1)
Leukemia[c]	5 (22.7)	2 (13.3)	3 (42.9)
Number of prior LOTs at index, n (%)			
2L+	7 (31.8)	4 (26.7)	3 (42.9)
3L+	7 (31.8)	5 (33.3)	2 (28.6)
4L+	2 (9.1)	2 (13.3)	0 (0.0)
Other[d]	6 (27.3)	4 (26.7)	2 (28.6)
RWD source, n (%)			
Chart-review	8 (36.4)	6 (40.0)	2 (28.6)
EMR (including combined with claims or clinical site databases)	7 (31.8)	5 (33.3)	2 (28.6)
Combined cohort and clinical trial data	3 (13.6)	1 (6.7)	2 (28.6)
Registry	2 (9.1)	2 (13.3)	0 (0.0)
Administrative hospital data	1 (4.5)	0 (0.0)	1 (14.3)
Prospective cohort study	1 (4.5)	1 (6.7)	0 (0.0)
Clinical trial stage, n (%)			
Phase 1 single-arm	1 (4.5)	1 (6.7)	0 (0.0)
Phase 1/2 or 1b/2 single-arm	9 (40.9)	7 (46.7)	2 (28.6)
Phase 2 single-arm	9 (40.9)	5 (33.3)	4 (57.1)
Phase 3	3 (13.6)	2 (13.3)	1 (14.3)
Reported outcomes, n (%)			
Time-to-event	22 (100.0)	15 (100.0)	7 (100.0)
OS	22 (100.0)	15 (100.0)	7 (100.0)
PFS	13 (59.1)	10 (66.7)	3 (42.9)
TTNT	8 (36.4)	5 (33.3)	3 (42.9)
DOR	3 (13.6)	3 (20.0)	0 (0.0)
Response	12 (54.5)	8 (53.3)	4 (57.1)
Studies that explicitly acknowledge index selection issue[e], n (%)	12 (54.5)	10 (66.7)	2 (28.6)
Index selection method used in the main analysis, n (%)			
Single-line methods	13 (59.1)	9 (60.0)	4 (57.1)
First eligible line	6 (27.3)	6 (40.0)	0 (0.0)
First available line	1 (4.5)	1 (6.7)	0 (0.0)
Last eligible line	2 (9.1)	0 (0.0)	2 (28.6)
Last available line	1 (4.5)	0 (0.0)	1 (14.3)
Eligible line with highest PS/similarity measure	2 (9.1)	2 (13.3)	0 (0.0)
Random eligible line	1 (4.5)	0 (0.0)	1 (14.3)
Multiple-line methods	6 (27.3)	6 (40.0)	0 (0.0)
All eligible lines	6 (27.3)	6 (40.0)	0 (0.0)
Unclear method	3 (13.6)	0 (0.0)	3 (42.9)
Authors provided justification for index selection method in the main analysis, n (%)			
Yes	13 (59.1)	9 (60.0)	4 (57.1)
No	9 (40.9)	6 (40.0)	3 (42.9)
Methodology paper cited, n (%)			
Hernan et Robins 2016[2]	7 (31.8)	6 (40.0)	1 (14.3)
Backenroth 2021[5]	2 (9.1)	2 (13.3)	0 (0.0)
Hampson 2024[6]	1 (4.5)	1 (6.7)	0 (0.0)
No relevant citations	15 (68.2)	9 (60.0)	6 (85.7)
Index selection method used in sensitivity analyses, n (%)			
First eligible line	3 (13.6)	1 (6.7)	2 (28.6)
Restricted line	2 (9.1)	1 (6.7)	1 (14.3)
Other[f]	1 (4.5)	1 (6.7)	0 (0.0)
No index selection sensitivities	16 (72.7)	12 (80.0)	4 (57.1)

Figure 1. Index Selection Methods used in included ECAs



Limitations

- This study relied on the search and screening of hemato-oncology ECAs conducted by Hermans et al. Relevant ECAs may have been missed as a result, and findings may not generalize to the broader oncology field.
- Appraisal of index line selection methods was based on methodology papers identified through a targeted literature search. Competing references may have been missed.
- Formal comparisons between studies were not conducted due to the heterogeneity of the ECAs.

Conclusions

- Around a third of the identified ECAs used index line selection methods not recommended by the literature or were unclear about their approach.
- Despite the proliferation of ECAs in hemato-oncology, insufficient attention is being given to index line selection methodology.
- Development of guidance on index line selection methods by regulatory and HTA bodies is welcome to ensure unbiased appraisal of RW-derived ECA evidence in application packages.

Electronic Poster

Please scan this quick response (QR) code with your smartphone app to view this poster. If you do not have a smartphone, access the poster via the internet at:
https://scientificpubs.congressposter.com/p/q3td0m12cmirdr4o

Abbreviations: ALL: acute lymphocytic leukemia; AML: acute myeloid leukemia; CLL: Chronic lymphocytic leukemia; DLBCL: diffuse large B-cell lymphoma; DOR: duration of response; ECA: external control arm; EMR: electronic medical records; IMiD: immunomodulatory drugs; LOT: line of therapy; MM: multiple myeloma; MF: mycosis fungoides; NHL: non-Hodkins lymphoma; OS: overall survival; PFS: progression-free survival; PI: proteosome inhibitor; PS: propensity score; R/R: relapse/refractory; RWD: real-world data; SM: similarity measure; SS: Sezary syndrome; TCE: triple class exposure; TTNT: time-to-next-treatment.

Notes:
[a]Categorization of indexing methods as " recommended" and “not recommended" is based on appraisals found in methodological papers discussing the indexing methods[2-8], and not by the authors of the present analysis.
[b]Including DLBCL, Follicular lymphoma, Advanced MF/SS, Mantle cell lymphoma
[c]Including ALL, CLL, AML
[d]Including 4L+ or refractory to PI and IMiD and TCE, 2L-4L, R/R
[e]Considered acknowledged when authors state that individuals in the RWD source can be eligible for inclusion in the ECA at different time points.
[f]Subset of first eligible line restricted to a predefined calendar-period
[g]Similarity measure is defined as the conditional probability that the patient was prescribed the intervention as their LOT given the values of their covariates at the time of starting this LOT.

References: 1. Hermans et al. 2024. JAMA Oncol 10: 1426–1436. **2.** Hernán et al. 2016. Am J Epidemiol 183: 758–764. **3.** Hatswell et al. 2022. Med Decis Making 42: 893-905. **4.** Van Le et al. 2024. Epidemiol Methods 13. **5.** Backenroth et al. 2021. Pharm Stat 20: 783–792. **6.** Hampson et al. 2024. Stat Biopharm Res 16: 1–10. **7.** Sun et al. 2024. J Biopharm Stat. **8.** Suissa. 2021. Epidemiology 32: 94–100.

Acknowledgments: This study was sponsored by Pfizer.

Disclosures: Michael Wallington and Yong Chen report employment and stock ownership and/or stock options in Pfizer. Iman Fakih and Ann-Sophie Demers report employment at STATLOG.

Contact: Michael Wallington | michael.wallington@pfizer.com

Presented at the 2025 ISPOR EU Meeting | November 9-12, 2025 | Glasgow, Scotland