Poster SA6

Comparing Established and Novel Index Date Selection Methods: A Simulation Study in Two Clinical Contexts

Neisha Opper[1], Shivani Aggarwal[1], David Goldfarb[2], Hoa Le[3]

[1]Landmark Science, Inc., Los Angeles, CA, USA, [2]Landmark Science, Inc. New York, NY, USA, [3]University of North Carolina at Chapel Hill, NC, USA

Scan here for e-poster

Background & Rationale

- Appropriate index date selection is fundamental to target trial emulation and other comparative effectiveness research, especially in areas such as hematologic and solid cancers where patients commonly have multiple eligible lines of therapy (LoTs).
- Several established index date selection strategies exist, but prior work has shown that each have different strengths and limitations depending on context.
- A recent study introduced and evaluated a novel methodology, stratified random LoT (SRL), for large B-cell lymphoma (LBCL) using real-world data.
- Simulation studies provide an opportunity to directly compare methods under controlled conditions where the true treatment effect is known.
- Prior simulation studies have been generic rather than disease-specific, leaving open questions about how disease dynamics should inform method selection.

Objective: To compare the performance of established and novel index date selection methods using simulated data emulating two disease contexts characterized by multi-line treatment patterns.

Methods

Data Source

- Synthetic cohorts were generated to emulate trial-like intervention patients, true controls, and external controls in **relapsed/refractory (r/r) LBCL** and EGFR-positive **advanced non-small cell lung cancer (aNSCLC)**.
- Disease settings were selected to **reflect distinct clinical contexts with differing LoT patterns and progression dynamics**.
- Overall survival was simulated under scenarios where prognosis deteriorated with increasing LoT.
- True controls were generated by re-simulating intervention patients under control parameters.

Index Date Selection Methods

We evaluated four LoT-based index date selection methods in a simulation study (Figure 1).

- First eligible line (**FEL**): selects the earliest LoT meeting eligibility criteria.
- Random eligible line (**REL**): randomly selects one eligible LoT per patient.
- All eligible lines (**AEL**): includes all eligible LoTs per patient; requires accounting for within-patient correlation.
- Stratified random line (**SRL**): randomly selects one eligible LoT per patient, stratified to match the LoT distribution of the trial population.

Analyses

- Treatment effects were estimated using Cox proportional hazards models under naïve, stabilized inverse probability of treatment weighting (sIPTW), and standardized mortality ratio (SMR) approaches.
- Propensity score models were estimated based on the disease-specific covariates plus line of therapy.

Performance Evaluation

- Index methods were compared based on their ability to recover the true treatment effect across repeated simulation iterations (n=5000).
- Reported performance metrics include bias and coverage.
- The impact of sample size, effect size, and overlap in prior LoTs was evaluated.

Key Findings

- Method choice meaningfully impacted performance, with differences in both bias and coverage across index date selection strategies.
- All eligible lines (AEL) demonstrated minimal bias and moderate-to-high coverage (82–86%); weighting improved coverage for AEL in LBCL but reduced it in NSCLC (Figure 3a & 3b).
- SRL performed comparably to AEL in LBCL, while avoiding reliance on multiple observations per patient (Figure 3b). Handling multiple observations per patient requires additional analytic complexity to address within-patient correlation and can make interpretation less intuitive.
- SRL generally outperformed FEL, with lower bias and/or higher coverage across analyses (e.g., LBCL sIPTW: 96% vs 68%; NSCLC sIPTW: 81% vs 64%) (Figure 4a & 4b).
- The comparative performance of SRL improved in settings where LoT was a stronger prognostic factor or LoT distributions were less comparable between arms, highlighting the importance of aligning index date selection with underlying disease dynamics and empirical features of comparison groups (Figure 5a & 5b).
- Random eligible line (REL) performed poorly, with substantial bias and near-zero coverage that was not resolved by weighting (Figures 3-5).

Limitations

- Simulations reflect clinically motivated but simplified disease processes. Future work should incorporate greater clinical realism (e.g., treatment–covariate interactions, alternative survival distributions) and extend to indications with different disease dynamics (e.g., indolent diseases such as follicular lymphoma).
- The simulated trial arm was limited to a first eligible line, reflecting common comparisons between external controls and single arm trials, but limiting generalizability to settings with multiple entry points, treatment switching, or discontinuation.
- Direct performance comparisons do not take into account that indexing strategies correspond to different estimands (e.g., AEL reflects a per-line effect, while single-line methods estimate a patient-level effect). Real-world method selection should also consider target estimand for appropriate causal inference.
- Weighting results should be interpreted with caution because baseline covariate distributions were the same across all simulated arms; in real-world application, weighting is likely to reduce residual confounding due to the absence of this design construct.

Why is this Research Important?

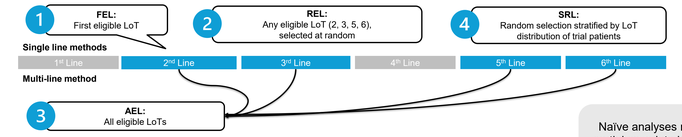
- Misspecified index dates can bias effectiveness estimates and reduce study validity, particularly in disease areas characterized by multi-line treatment patterns.
- The use of RWE to support clinical and regulatory decision-making has grown more common, particularly in areas like oncology and hematology, increasing the demand for rigorous design.
- These findings provide practical guidance on method selection in clinically relevant settings, helping inform better design choices for real-world studies.

LANDMARK SCIENCE

Presented at ISPOR 2026 | May 17–20, 2026 | Philadelphia, Pennsylvania.

Results

Figure 1. Schematic of the Index Date Selection Methods Evaluated



Naïve analyses reflect partial covariate balance because distributions were not arm-specific.

Table 1. Disease-Agnostic Simulation Parameters in Primary Analyses

Number of patients sampled per arm	750
Eligibility proportion of external controls	75%
Mean OS model	$\mu = \text{Base}_{\text{arm}} + k_{\text{arm}} \cdot \sum_{i=1}^{x} \text{centered } x_i - \text{LoT deterioration}$
Clinical deterioration by LoT	Modelled as an additive penalty per line to represent worsening health per LoT; additional penalties are applied as interaction terms for patients with certain characteristics (e.g., metastases).

Disease-agnostic parameters were modelled after Hatswell AJ, Deighton K, Snider JT, Brookhart MA, Faghmous I, Patel AR. Approaches to selecting "time zero" in external control arms with multiple potential entry points: a simulation study of 8 approaches. Medical Decision Making. 2022 Oct;42(7):893-905.
NOTE: Time to progression (TTP) was used to construct patient timelines (e.g., progression vs. death depending on whether survival time was greater than TTP at the end of each line).

Table 2. Disease-Specific Simulation Parameters in Primary Analyses

	aNSCLC	r/r LBCL
Distribution of baseline patient characteristics per arm	Age: Continuous (Mean: 66, SD:9) Sex: Male/Female (Female: 56%) ECOG Score: 0-2 (0: 40%, 1: 35%, 2: 25%) Smoking Status: Yes/No (Yes: 7%) Adenocarcinoma: Yes/No (Yes: 93%) Brain metastases: Yes/No (Yes: 25%) Liver metastases: Yes/No (Yes: 5%)	Age: Continuous (Median: 62; Range: 20-89) Sex: Male/Female (Female: 37%) ECOG Score: 0-2 (0: 30%, 1: 46%, 2: 24%) Extranodal disease: Yes/No (Yes: 60%) Bulky disease: Yes/No (Yes: 20%) Prior HSCT: Yes/No (Yes: 18%) Refractory to last LoT: Yes/No (Yes: 93%) Chemorefractory; Yes/No (Yes: 75%)
Starting line of therapy	Intervention Patients & True Controls: Binomial (probability = 0.67, size=6) External controls: Binomial (probability = 0.33, size=6)	Intervention Patients & True Controls: Binomial (probability = 0.42, size=6) External controls: Binomial (probability = 0.25, size=6)
Clinical deterioration by LoT	Moderate prognostic value; ~1 month decrement per LoT for an average patient)	Substantial prognostic value; ~ 2 months decrement per LoT for an average patient)
Administrative censoring	Intervention Patients/True Controls: 36 months External controls: 25 months	All arms: 25 months

Disease-specific parameters were based on Patel J, Meng J, Le H, Tanaka Y, Phani S, Salas M, Wu C, Sternberg D, Esker S, Anderson JP, Crowley A. Real-world treatment patterns and clinical outcomes among patients with metastatic or unresectable EGFR-mutated non-small cell lung cancer previously treated with osimertinib and platinum-based chemotherapy. Advances in Therapy. 2024 Aug;41(8):3299-315 for aNSCLC and Van Le H, Van Naarden Braun K, Nowakowski GS, Sermer D, Radford J, Townsend W, Ghesquieres H, Menne T, Porpaczy E, Fox CP, Schuetzbauer C. Use of a real-world synthetic control arm for direct comparison of lisocabtagene maraleucel and conventional therapy in relapsed/refractory large B-cell lymphoma. Leukemia & lymphoma. 2023 Feb 23;64(3):573-85 for r/r LBCL.
NOTE: LoT distribution was offset by 1 for LBCL to reflect relapsed/refractory status.

Figure 2a. Overall Survival Curves by Index Date Selection Method for aNSCLC

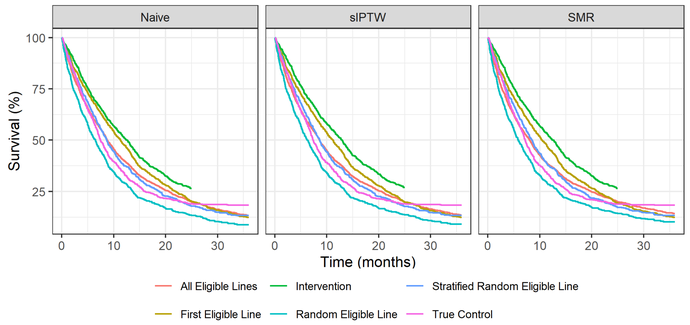


Figure 2b. Overall Survival Curves by Index Date Selection Method for r/r LBCL

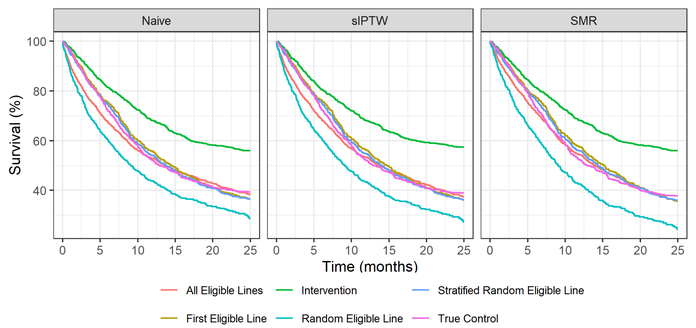


Figure 3a. Effect Estimates by Index Date Selection Method for aNSCLC

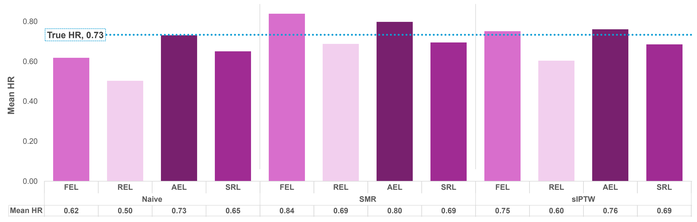


Figure 3b. Effect Estimates by Index Date Selection Method for r/r LBCL

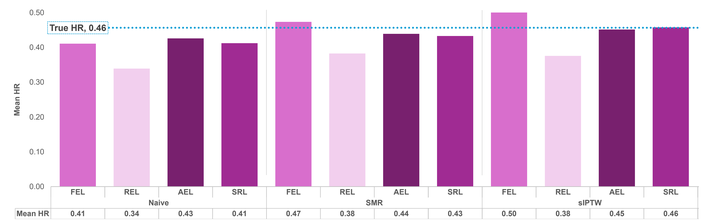


Figure 4a. Coverage and Bias by Index Date Selection Method for aNSCLC in Naïve & Weighted Analyses

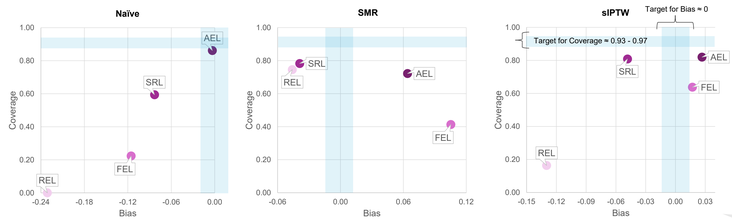
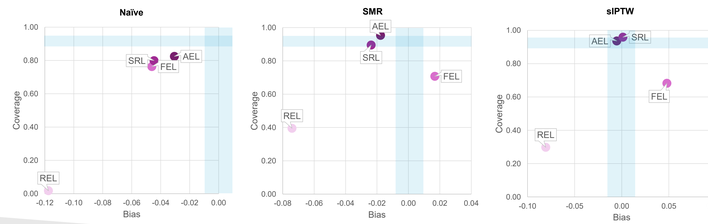


Figure 4b. Coverage and Bias by Index Date Selection Method for r/r LBCL in Naïve & Weighted Analyses



Coverage diagnostics are essential; low bias alone may mask poor inferential performance.

Figure 5a. Indexing Methods Ranked by Bias & Coverage Across Scenarios for aNSCLC

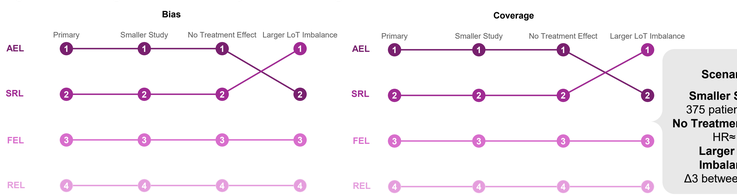
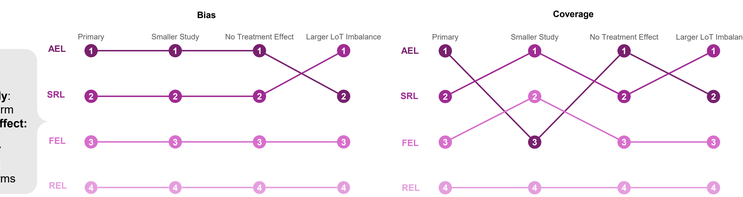


Figure 5b. Indexing Methods Ranked by Bias & Coverage Across Scenarios for r/r LBCL



References: References are available upon request to the corresponding author at Shivani@landmarkscience.com; **E-poster and supplementary materials**: Copies of this poster and supplementary materials obtained through the QR code or link are for personal use only and may not be reproduced without permission;
Funding: This study was funded by Landmark Science, Inc.; **Disclosures:** NO, SA, DG are employees of or are contracted to Landmark Science, Inc.; HL is affiliated with University of North Carolina at Chapel Hill;
Acknowledgements: The authors thank Anthony Hatswell and Sudhir Venkatesan for their contributions to this project.