Paper AS08

Index Date Selection in Externally Controlled Arms: Methods, Impact, and Practical Considerations

Lorraine Fang, Lixiao Chen, Kelvin Lin, Kevin Yang, Peishan (Jenny) Han,

Marc De Benedetti, Lihua Yue

Bristol Myers Squibb, Lawrence Township, NJ, USA

ABSTRACT

Establishing an external control arm (ECA) for single-arm trials requires defining the index date (time zero), the point at which a patient would be eligible in a contemporary study. In real-world data (RWD), patients often have multiple qualifying lines of therapy (LOT), creating complexity in index-line selection. We evaluated several approaches, including First Eligible Line, Last Eligible Line, Second-to-Last Eligible Line, All Eligible Lines, Restricted to 3rd Line, Stratified Random Line and Random Eligible Line. Using real study data, we examined the impact of these methods on treatment-effect estimates for a target therapy versus standard of care (SOC). To our knowledge, this is the first study to compare index-date assignment strategies in RWD using actual study data rather than simulations, providing important insights for the design and interpretation of ECAs.

INTRODUCTION

Externally controlled trials (ECTs) are increasingly used to evaluate treatment effects when randomized controls are infeasible or unethical, such as single-arm trials with strong efficacy signals or rare diseases. RWD offers an opportunity to construct comparator cohorts retrospectively. Index date selection is a critical design component in externally controlled arms (ECAs) using real-world data (RWD). Unlike clinical trial participants who have a clearly defined, protocol-driven index date (e.g., treatment infusion, first dose starting), RWD patients may meet eligibility criteria at multiple points in time across several lines of therapy (LOTs). This paper summarizes seven index date selection methodologies, evaluates their impact using real patient-level data from a large multi-source Large B-Cell Lymphoma dataset, and highlights methodological considerations, strengths, and risks associated with each method. Our empirical findings demonstrate that although overall treatment effect direction between Breyanzi (a CAR-T therapy developed by BMS) and ECA is consistent across methods, small but meaningful variations in effect estimates, sample size, balance, and potential biases must be considered when selecting an approach for regulatory-grade ECAs.

Index date selection therefore determines:

- which LOT defines the patient's pseudo-baseline,
- how prognostic factors are measured,
- the patient's eligibility status at that time,

- and ultimately, the estimated treatment effect.

This paper summarizes the methodological landscape of index date selection and reports findings from a real-data example using seven approaches applied to relapsed/refractory LBCL patients.

BACKGROUND & MOTIVATION

Clinical trial subjects are indexed at the start of treatment (e.g., CAR-T infusion). For RWD subjects, however, the researcher must choose retrospectively among several potential LOTs at which patients satisfy eligibility. Choosing different timepoints can:

- introduce survivor bias,
- inflate or deflate outcomes,
- alter covariate distributions,
- affect propensity score balance,
- and influence sample size and statistical power.

The dataset used in this analysis combines global, multi-source LBCL RWD from Cota, Flatiron, GRN, IQVIA US and EU registries, with identical inclusion/exclusion criteria applied as in the TRANSCEND trial. Of 1450 patients screened, 606 met high-level comparator criteria, forming the Initial Comparator Cohort (ICC). Among them, 443 patients had ≥1 eligible LOT.

INDEXING METHODS

Seven indexing methods were used to assign pseudo-baselines to the 443 eligible RWD patients.

1. Stratified Random Line (SRL):
   RW patients are randomly sampled to match the clinical LOT distribution, without duplication. This ensures balance on the highly prognostic factor, number of prior LOTs.
   • Strengths: Excellent balance; aligns with the clinical trial LOT distribution.
   • Weaknesses: Requires a large RWD source; reduces sample size.

2. First Eligible Line (FEL):
   Index at the earliest LOT at which a patient meets all eligibility criteria.
   • Strengths: Simple and reproducible.
   • Weaknesses: Prone to bias toward better outcomes due to earlier indexing.

3. Restricted to 3rd Line (3L):
   Index at LOT = 3 for both trial and RW subjects.
   • Strengths: Fully standardizes LOT, eliminating LOT imbalance.
   • Weaknesses: Significantly reduces sample size; may lower statistical power.

4. All Eligible Lines (AEL):
   Reuse subjects at every eligible LOT (i.e., multiple rows per subject).

• Strengths: Maximizes sample size; produces stable effect estimates.
• Weaknesses: Within-patient correlation is not fully addressed.

5. Random Eligible Line (REL):
   Randomly select among a patient's eligible LOTs.
   • Strengths: Preserves the natural LOT distribution.
   • Weaknesses: Cannot guarantee LOT balance.

6. Last Eligible Line (LEL):
   Index at the most recent LOT at which the patient remains eligible.
   • Weaknesses: Risks survivor bias and reflects a more advanced disease state.

7. Second-Last Eligible Line (SEL):
   Index at the second most recent eligible LOT.
   • Weaknesses: Similar limitations to LEL; behavior depends on the LOT pattern.

COHORT DEFINITIONS

Inclusion/exclusion criteria similar to TRANSCEND trial were applied to comparator cohort selection. These I/E criteria include:

- Age ≥18 at LBCL diagnosis,
- Prior anthracycline + rituximab exposure,
- Completion of ≥2 LOTs and starting a subsequent LOT,
- Histology consistent with DLBCL NOS, FL3B, HGL, or PMBCL.

The ICC (N=606) was refined to:

- 443 subjects with ≥1 eligible LOT,
- 257 subjects used under SRL,
- 381 subjects under FEL,
- 653 index events (443 unique patients) under AEL,
- 118 subjects under 3L alignment.

STATISTICAL METHODS

Propensity score (PS) stabilized inverse probability of treatment weighting (sIPTW) was used to balance prognostic factors between the RWD comparator cohort and TRANSCEND trial arm. Twelve prognostic factors included:

- Age
- Gender
- Time since diagnosis
- Number of prior LOTs
- LOTs per year since diagnosis
- Refractory status to last therapy
- Best previous response
- Prior HSCT

- Chemosensitivity
- Bulky disease presence
- Extranodal disease
- Disease stage

RWD missing data were imputed 25 times. Primary endpoints examined were: ORR (overall response rate), CRR (complete response rate), OS (overall survival), PFS (progression-free survival)

RESULTS

Across all indexing methods, the treatment effect consistently favored Breyanzi, with adjusted HRs and RRs showing large, statistically significant differences between arms.

Key Findings

- All seven methods produced consistent directionality of treatment effect.
- Median OS for ECAs ranged from 5.7 to 8.0 months across methods.
- Median PFS ranged narrowly from 2.1 to 2.5 months.
- SRL and AEL showed the strongest covariate balance.
- 3L and LEL methods produced the most extreme OS values due to LOT constraints.

Numerical Highlights

- Adjusted ORR for RWD ECAs: 34–40% vs ~74% for TRANSCEND.
- HR for OS ranged from 0.45 to 0.53, consistently favoring treatment.
- HR for PFS ranged from 0.55 to 0.60, similar across indexing methods.

DISCUSSION

While treatment effect magnitude was stable due to the strong efficacy of Breyanzi, subtle differences across indexing methods were observed:

Pros & Cons Observed:

Method	Pros	Cons
SRL	Excellent balance; avoids LOT bias	Requires large data; reduces N
FEL	Simple	Biased toward better outcomes
3L	Fully standardized LOT	Lowest N; reduced power
AEL	Maximizes effective sample size	Within-patient correlation unresolved
REL	Simpler version of SRL	Cannot ensure balance
LEL	—	Survivor bias risk
SEL	—	Similar to LEL limitations

Index date selection is not a trivial technical decision—it can meaningfully shape cohort characteristics, covariate balance, and downstream causal inference. For regulatory use, methods that reduce bias and maintain balance (e.g., SRL, AEL) are generally advantageous.

SUMMARY

This analysis provides the first real patient-level empirical comparison of seven index date selection strategies applied to a multi-source real-world LBCL dataset. Although all indexing methods produced consistent treatment-effect directionality, differences in baseline balance, sample size, and susceptibility to bias clarify that no single indexing strategy is universally optimal. Selection should be guided by:

- study design objectives,
- available sample size,
- LOT distribution,
- regulatory expectations,
- and sensitivity to survivor/selection biases.

Careful methodological justification for index selection is essential for creating credible RWD-based comparator cohorts.

REFERENCES

1. Van Le, Hoa, De Benedetti, Marc, Yue, Lihua, Fang, Lorraine, et al. "Effect of designations of index date in externally controlled trials: an empirical example." *Epidemiologic Methods*, 2024.

2. Link to the publication: - https://doi.org/10.1515/em-2023-0041?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle

CONTACT INFORMATION

Author: Lorraine Fang **Email:** Lorraine.Fang@bms.com