Leveraging RWD to advance clinical development: Statistical Considerations Related to External Control Arms

Optimizing the Use of Real-World Data through Rigorous Statistical Techniques

MSR132

View all Parexel's posters at ISPOR Europe 2024

Tuli De[1], David W. Warne[2], Angelina Jiang[1] ,Mukesh Kumar Jha[3],, JD Kerr[1], Kateryna Chepynoga, Jackie Vanderpuye-Orgle[1]

[1]Parexel, Durham, North Carolina, USA [2]Parexel, London, UK
[3]Parexel, Hyderabad, India

Background

- External Control Arms (ECA) are becoming increasingly important in clinical trials where randomized controlled trials (RCTs) are difficult or unethical, such as in rare diseases, oncology, or when historical data is needed to supplement a trial.
- The use of ECAs allows researchers to leverage real-world data (RWD) to construct a comparative group outside of the traditional trial design.
- Regulatory agencies, such as the FDA and EMA, have acknowledged the utility of ECAs, but they require highly rigorous statistical methods to ensure that the control arm is comparable to the clinical trial population and to avoid bias. Statistical techniques **like Propensity Score Matching (PSM), Inverse Probability of Treatment Weighting (IPTW), Estimated Propensity Score (ePS), and Bayesian Borrowing** have been proposed and used to mitigate bias and confounding issues. However, these statistical methods need to account for differences between external and trial populations.
- The objective of this study is to review FDA and EMA guidelines regarding statistical methods for ECAs by conducting a targeted literature review to identify key studies that applied ECAs using various statistical techniques,
- The aim is to provide recommendations related to the applicability and fitment of a few optimal methods that have been used in regulatory submissions.

Methods

- **Targeted Literature Review:**
  - Conducted a systematic search of academic databases (PubMed, Google Scholar) and regulatory publications from FDA [1] and EMA [2], [3] related to the use of ECAs in clinical trial, focused on guidelines for real-world evidence (RWE) and statistical standards for ECA.
  - Selection Criteria: Studies were included if they implemented ECA and provided data on regulatory approval or feedback from FDA/EMA.
- **Key Statistical Techniques Examined:**
  - **Propensity Score Matching (PSM):** A method that matches patients in the external control arm to the clinical trial participants based on key covariates. PSM ensures that the baseline. PSM ensures that the baseline characteristics are balanced between the two groups.
  - **Inverse Probability of Treatment Weighting (IPTW):** Uses the probability of receiving a treatment (propensity score) to weight participants and create a pseudo-population that balances covariates across groups.
  - **Estimated Propensity Score (ePS):** Expands on traditional PSM by incorporating additional covariates and empirically weighting them to refine matching and address heterogeneous data sources.
  - **Bayesian Borrowing:** Combines external and internal trial data using Bayesian models. This method dynamically adjusts the degree of borrowing based on the consistency of the external control data with the trial data.
- Recent advancements in Targeted Learning or Causal Machine Learning (CML) offers a promising frontier for improving robustness of ECAs. These methods are designed to handle complex RWD allowing for better control over confounding bias. Hence using these advanced models more precise and reliable estimates of treatment effects in the ECA can be achieved when dealing with high-dimensional or heterogenous data.

Results and Discussion

- The literature review identified key studies that employed different statistical methods in the construction of external control arms:
- The TLR indicated that the following statistical methods have been used in ECAs-

Table 1: FDA and EMA Guidelines on External Control Arms

Regulatory Body	Guidelines	Key Statistical Methods	Applications	Documentation Required
FDA (U.S.)	Real-World Evidence (RWE) Framework (2018)	- Propensity Score Matching (PSM) - Inverse Probability of Treatment Weighting (IPTW) - Bayesian Hierarchical Models - Synthetic Control Arm	Oncology, rare diseases, when RCTs are infeasible	Extensive documentation, sensitivity analyses, clear justification for matching methods
EMA (EU)	Registry-Based Studies Guidelines (2021), RWE for Regulatory Decision Making	- PSM, IPTW - Cohort Matching - Matching-Adjusted Indirect Comparisons (MAIC)	Rare diseases, oncology, post-marketing studies	Detailed validation of data sources, pre-specified statistical analysis plans, sensitivity analyses
Similarities	Clear guidance for statistical methods and documentation should implemented	Propensity Score Matching (PSM) IPTW	Rare Disease, oncology	Sensitivity analysis, clear analysis plan/methods

Table 2: Summary of Literature Review of Statistical Methods for ECAs

Study Title	Statistical Method	Agency	Key Findings
Tafasitamab + LEN in Oncology (2020) [4]	Propensity Score Matching (PSM)	FDA	External control data was accepted, but concerns were raised about covariate imbalance, requiring additional sensitivity analyses.
Rare Disease Application (2021) [5]	Bayesian Borrowing	FDA/EMA	Bayesian model successfully integrated external data with small sample sizes, providing robust efficacy estimates.
Breast Cancer Post-Marketing Study (2019) [6],[7]	Inverse Probability of Treatment Weighting (IPTW)	EMA	IPTW successfully balanced covariates, but extreme weights caused sensitivity issues in later regulatory review.
Orphan Disease Study (2022) [8]	Matching-Adjusted Indirect Comparisons (MAIC)	EMA	MAIC allowed for a more refined matching process, which led to a successful post-marketing submission.
Multiple Myeloma ECA Study (2020) [9]	Estimated Propensity Score (ePS)	FDA	While ePS showed potential for robust matching, it was rejected due to inconsistencies in covariate selection and lack of validation.

Table 3: Comparison of Statistical Methods for ECAs

Method	Strengths	Challenges
Propensity Score Matching (PSM)	- Widely accepted by regulators - Balances observed covariates	- Cannot address unmeasured confounders - Dependent on covariate selection
Inverse Probability of Treatment Weighting (IPTW)	- Effective for balancing large datasets - Reduces selection bias	- Sensitive to model misspecification - Can produce extreme weights
Estimated Propensity Score (ePS)	- Allows for flexible matching of covariates - Suitable for heterogeneous data sources	- Lack of standardization - Difficult to validate in regulatory settings
Bayesian Borrowing	- Dynamic borrowing of information - Robust for small sample sizes	- Computationally intensive - Requires high expertise and regulatory validation
Matching-Adjusted Indirect Comparisons (MAIC)	- Adjusts for population differences between trial and external controls	- Data harmonization can be challenging

- **Key takeaways from Statistical Considerations of ECA:**
  - Among the methods reviewed, Propensity Score Matching (PSM) remains the most commonly used and accepted method due to its simplicity, transparency, and ability to balance baseline covariates. However, Bayesian Borrowing is emerging as a robust alternative, particularly for trials involving small sample sizes or complex RWD.
  - Studies that utilized PSM and IPTW showed the highest rates of approval from both the FDA and EMA. However, advanced methods like ePS and Bayesian Borrowing are gaining traction, especially in areas with more complex datasets and when dealing with RWE.
  - The choice of statistical method should align with the nature of the trial and the quality of the available external data. PSM is recommended for straightforward applications, while Bayesian Borrowing provides flexibility and precision for smaller or more complex trials, though it demands thorough validation and justification.
  - Targeted Learning [10] enhance traditional statistical methods like PSM by allowing flexible modelling of covariates and interactions. These methods have the potential to improve balance between groups in ECAs, thus increasing the validity of the estimated treatment effect.

Conclusions and recommendations

- ***Large, Homogeneous Data****:* Use PSM for balancing covariates, ensuring transparency and ease of interpretation. IPTW can be another reliable alternative here. Hence for standard submissions, these methods work well.
- ***Heterogeneous, Small Populations:*** Use Bayesian Borrowing, which offers greater flexibility but requires extensive validation. This method is recommended to use in regulatory submission when dealing with complex data sources
- ***Higher Precision:*** (ePS) is recommended for its ability to handle a broader range of covariates, making it suitable for more complex datasets.
- ***Emerging Methods:*** Targeted Learning could leverage large, complex datasets while accounting for causal structures makes them particularly critical in setting where traditional methods may struggle.
- ***Direction for future work*****:** Conducting ECA analysis based on a simulated dataset to evaluate the performance of different statistical methods discussed in this poster. Then compare the robustness of these methods under various scenarios and offer recommendations in terms of bias reduction, precisions, and interpretability.

REFERENCES

[1] FDA Real-World Evidence Framework (2018).
[2] EMA Guidelines on Registry-Based Studies (2021).
[3] Methodological Considerations in Using RWD for External Control Arms.
[4] Grzegorz S.et al. Clin Cancer Res 2022 Mar; 28(18): 4003–4017
[5] Moreno et al. Int J Environ Res Public Health 2021 18 (3) 1022
[6] Makito et al. Cancer Manag Res 2022, 14, 623-635
[7] EMA Post marketing guidelines
[8] EMA Orphan Medicine Database
[9] FDA BLA Approval Database
[10] Lauren et al. arXiv:2210.05802v3