

---

# Enhancing single arm trials using external data as control

Camille Marian\*<sup>1</sup>

<sup>1</sup>Laboratoire Servier – Laboratoire Servier – France

## Résumé

**Background:** The gold standard for clinical trials today is the randomized controlled trial (RCT). However, this design is not always feasible due to ethical or practical constraints. On the other hand, single-arm trials are insufficient for accurately estimating treatment effects. This work aims to evaluate various methods for incorporating a synthetic control arm as a solution to emulate an RCT using real-world data.

**Methods:** The use of the Propensity Score Matching (PSM) algorithm to create equivalent populations has been approved by the FDA and will serve as the reference method. Another propensity score-based method, Inverse Probability of Treatment Weighting (IPTW), will also be evaluated. Additionally, a new method incorporating all available data and machine learning algorithms will be investigated: Double Machine Learning (DML) with random forest and extreme gradient boosting.

**Results:** Simulations showed that the standard methods (PSM and IPTW) demonstrate good operational characteristics but may introduce bias in certain settings. In contrast, DML performs well at controlling bias but is often too complex for most scenarios. This method may be more appropriate in cases with a large number of prognostic covariates and complex relationships between them.

**Limitations and Prospects:** The simulations primarily explored scenarios that favored classical methods; more complex scenarios should be investigated. Moreover, machine learning methods require more tuning than is computationally feasible in repetitive settings. Further research is needed to fully evaluate the potential of these methods.

**Keywords:** Randomized controlled clinical trial, Synthetic control arm, Single arm trial, Propensity score, Doubly debiased, Natural history data.

---

\*Intervenant