Abstract No.
MSR82

Dealing With Non-Comparative Data: Proposed Alternatives and Considerations to Conduct Indirect Treatment Comparisons (ITCs)

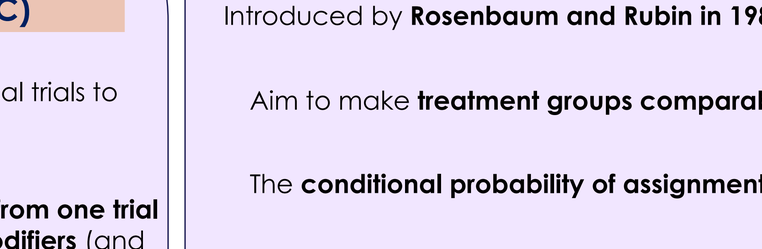
Le Nouveau P[1], Gauthier A[2]
[1]Amaris Consulting, Paris, France, [2]Amaris Consulting, Barcelona, Spain

BACKGROUND

- **Indirect treatment comparisons (ITCs)** have been developed to allow generating **relative treatment effects vs all relevant comparators**, beyond the comparators selected in the pivotal trial.
- Standard ITC methods require **homogeneity across trials** in terms of **treatment effect modifiers**, as well as **connectivity of network of evidence.**
- However, in some cases **data** available to conduct the ITC is **non-comparative**, **limiting the use of standard methods**:

Single-arm clinical trial

Difference in trial design



Difference in populations

METHODS & OBJECTIVES

- Recommendations from the NICE Technical Support Documents (TSDs) were reviewed and previous experiences encountered within Amaris were considered. [1,2]

➢ To suggest **alternatives** when dealing **with non-comparative data**.

➢ To provide **guidance** on key **considerations at feasibility assessment** phase to define relevant ITC approach.

Results – Methods to perform ITCs with non-comparative data

- Different **factors** will help identifying the **most appropriate approach** to conduct the ITC, and should be considered during **feasibility assessment phase**:

Connectivity: anchored vs unanchored vs

Individual patient data (IPD) availability vs

Number of comparators vs

IPD available for pivotal trial only

Matching-Adjusted Indirect Comparison (MAIC)

Simulated Treatment Comparison (STC)

Objective

Comparing indirectly two treatments not directly compared in clinical trials to **overcome heterogeneity or non-connectivity issues.**

Principle

Estimating relative treatment effect between two treatments **using IPD from one trial and aggregated data for the second one** while **adjusting for effect modifiers** (and prognostic factors depending on anchored status of the ITC)

- **MAIC**: **similar to a propensity score** weighting method
- **STC**: based on **outcome model fitted to IPD** as function of covariates

Why conduct MAIC and/or STC?

- **Unanchored** ITC
- High level of **heterogeneity** between studies

Key assumptions

- All **effect modifiers** (and **prognostic factors** for unanchored ITCs) **are known and observed** in both trials so that they can be adjusted for.
- Assumes that the **trial with IPD includes the competitor's trial population.**
- **Results** obtained **only valid in the population of comparator's trial.**

Additional assumptions for STC

- Use of **identity link function**, potential bias if done on another scale.
- For **time-to-event outcomes**: assumes a **parametric distribution.**
- Assume that **parameters of the regression** estimated on trial with IPD available are **applicable to competitor's population.**

What type of data to inform on comparator in MAIC/STC?

Both clinical trials, as well as observational studies, as long as enough data is available

What if comparators' are connected in a network?

Results from MAIC/STC related to one comparator of this network can be used **to connect pivotal trial to the global network**

IPD available for both trials

Propensity score approaches in brief

Introduced by **Rosenbaum and Rubin in 1983[3]** to address the issue of **selection bias** in observational research.

Objective

Aim to make **treatment groups comparable** or to **control for** the effect of prognostic factors on outcomes.

Principle

The **conditional probability of assignment to a particular treatment** given a vector of observed covariates.

Key steps

Estimate **propensity score** across all patients via a **logistic** regression
Ensure **sufficient overlap** between groups
Select **adjustment method**, main options: **matching**, **weighting**, stratification, doubly robust

AI to generate propensity scores: our experience

Objective: To compare methods (logistic regression vs machine learning [ML]) to estimate propensity scores when constructing an external control arm using oncology data from Project Data Sphere®

Methods compared

1. Logistic regression (standard approach)
2. Classification and Regression Trees (CART)
3. Naïve Bayes classifier
4. Support Vector Machines

Project steps

Choose data with **sufficient overlap** in **baseline** demographic and clinical **characteristics**
Using data from the **control arm of two data sources** in order to compare **performance of methods**

Targeted literature review → **Identification of data sources** → **Comparison of survival and baseline characteristics** → **Methods to evaluate the propensity score** → **Comparison of algorithms** → **Dissemination of findings**

To identify the **most appropriate machine learning algorithms** to apply and what **prognostic factors** should be considered

In terms of **covariate balance** and **homogeneity of survival distributions**

AI to generate propensity scores: our findings

Data source

➢ Use of **historical clinical trial** data generally **allows for more clinically relevant variables**, not available in larger claims and/or electronic health records databases and **more overlap** between original populations.

Sample size & variables

➢ Given limited number of variables and small sample size, **ML algorithms** generally **did not perform better than logistic regression** to create an external control arm.
➢ **Having all relevant prognostic variables** available in both data sets **is key** to a successful model!

Predictive power

➢ More patients were excluded when using ML algorithms due to their high predictive power.

DISCUSSION

- **Towards personalized medicine**: new drug development focused on **rare diseases and target populations often justifies non-RCT designs** – especially in disease areas such as **oncology** or **hematology**.
- **Alternatives** can be considered **to conduct ITCs**, should they be **conducted in line with recommendations**, built on a **strong rationale** and validated clinically.

Key considerations suggested to build a strong comparison

Prognostic factors and effect modifiers

➢ **Strong rationale** required on list of factors.
➢ Could be based on literature, clinical expertise, IPD analyses, etc.

Choice and comparability of trials

➢ **Strong justification of choice of trials** required (e.g. external arm)
➢ How was it identified? Additional trials? Comparability?

Methodological guidelines and references

➢ Will **ensure robustness**.
➢ Should more complex approach be used, individual steps could still be **related to specific guidelines**

Involvement of multiple stakeholders

➢ Will **help justifying approach** and will **ensure expectations are met**.
➢ Could be clinical/ methodological experts, country affiliates, etc.

Extensive set of sensitivity analyses

➢ Will help **showing robustness** and understanding **impact of choices** made.
➢ Could include: use of other trials, alternatives approaches, etc.

Validity of results

➢ Will help **giving credibility** to the approach.
➢ Could be based on clinical experts, comparison with trials, literature, etc.

References:
1. Faria et al. 2015, NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data.
2. Philippo et al. 2016, NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE.
3. Rubin et al. 1983, The central role of the propensity score in observational studies for causal effects.

Abbreviations:
AI: Artificial intelligence, CART: Classification and Regression Trees (CART), IPD: Individual patient data, ITC: Indirect treatment comparison, MAIC: Matching-adjusted indirect comparison, ML: Machine learning, NICE: National Institute for Health and Care Excellence, RCT: Randomized controlled trial, STC: Simulated treatment comparison, TSD: Technical Support Document

ISPOR Europe, Vienna, AT; November 6 – 9, 2022