# Addressing treatment implementation bias in the construction of high dimensional propensity score

**N.H. Thurin[1], J. Jové[1], P. Blin[1], M. Rouyer[1], S. Lamarque[1], P. Bosco-Levy[1], N. Moore[1], C.Droz-Perroteau[1]**

[1] Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France

## Background

- Formally, covariate for propensity score (PS) construction are assessed before treatment onset.
- **However differences occurring in patient journeys after the decision to treat and before the treatment onset may bias the score** (*e.g.* pre-chemotherapy assessment)

## Objectives

**To address treatment implementation bias** induced by differential patient journeys between the decision to treat and the treatment initiation.

## Methods

### Study design

- **Comparative effectiveness study** to compare 1st-line advanced cancer treatments:
  - **intravenous agent**
  - **oral treatment**

### Data source

- Extraction from the **French nationwide healthcare database (SNDS)**
- From 01/01/2009 to 12/31/2016

### Study population

- **1 213** patients initiating an **intravenous** agent,
- **2 442** patients initiating an **oral** agent.

### Construction of the high dimensional Propensity Scores (hdPS)

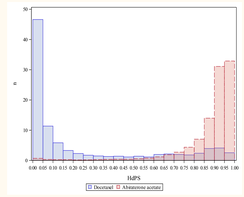
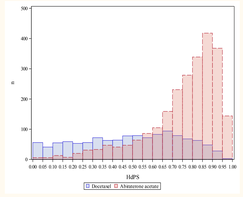
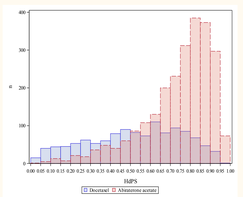
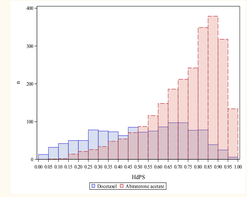
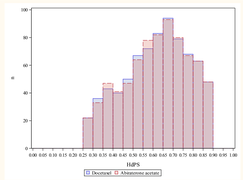
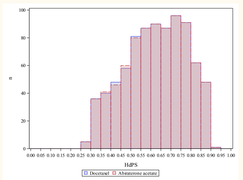
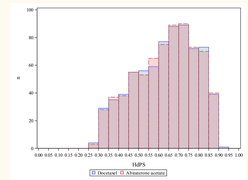
- Different **hdPS models** to estimate the probability for a patient to be treated by an intravenous *versus* oral agent
  - 100 variables empirically selected from 5 dimensions
    - Long term disease registration
    - Hospital discharge diagnoses
    - Dispensed drugs
    - Performed laboratory tests
    - Performed medical procedures
  - Extra forced variables judged clinically pertinent by experts
  - Different covariates assessment period length
- **1:1 matching** on hdPS, and potentially other forced variables

### hdPS performance assessment

- hdPS distributions
- **C-statistics value** (the closer to 0.5 the better)
- Number of **matched-patient pairs**
- **Number of variables with a standardized difference (SD) > 10% between the comparator groups**
- Number of variables associated with the outcome with SD>10%

## Results

***Table 1. Comparison of successive hdPS models***

| | | Model 1 | *Model 1.1* | Model 2 | Model 3 |
|---|---|---|---|---|---|
| **Forced variables** | Age at initiation treatment | ✔ | ✔ | ✔ | ✔ |
| | Disease stage before index date | ✔ | ✔ | ✔ | ✔ |
| | Charlson comorbidity index | ✔ | ✔ | ✔ | ✔ |
| | Total healthcare costs | ✔ | ✔ | | |
| | Treatment of bone metastases | | | | ✔ |
| | Cancer specific procedure 1 | | | | ✔ |
| | Dispensing of antineoplastic agents | | | | ✔ |
| **Covariate assessment period length** | | **[-1 year; index date]** | **[-1 year; -1 month]** | **[-1 year; -1 month]** | **[-2 years; -1 month]** |
| **hdPS distribution** (Intravenous / Oral) | **Before matching** |  |  |  |  |
| | **After trimming and matching 1:1** | Not applicable |  |  |  |
| **Number of potential matched-patients pairs** | | 273 | 765 | 830 | 716 |
| **Matching strategy** | | hdPS (± 0.01) | hdPS (± 0.01) | hdPS (± 0.01)<br>Age ± 1 an | hdPS (± 0.01)<br>Previous stage of cancer<br>Diagnosis date ± 1 year |
| **C-statistic value** | | 0.713 | 0.614 | 0.586 | 0.603 |
| **Number of variables with SD >10%** | | Not applicable | 52 | 48 | 17 |
| | **and statistically associated with the outcome** | Not applicable | 14 | 12 | 5 |

- Setting adjustment from **model 1** to **model 2** resulted in the improvement of the number of matched-patient pairs (830 vs. 273) and of the C-statistic value (0.586 vs 0.713); however 12 variables with SD>10% remained statistically associated with the outcome
- Intermediate model (**model 1.1**) showed that this improvement resulted mainly from the **exclusion of the month preceding the index date** from the covariate assessment period. This effect could be explained by a differential healthcare pathway between the time treatment is decided and the time the treatment is actually initiated (e.g. pre-therapeutic imaging, dosage).
- Setting adjustment from **model 2** to **model 3** induced a slight diminution of the number of matched patient-pairs (716 vs. 830) but led to a reduced number of variables with SD>10% associated with the outcome (5 vs. 12). The **model 3** was used for adjustment in survival analyses.

## Conclusion

- ✓ **Pre-exposure window should be routinely assessed and potentially excluded when constructing propensity score, especially when patients have followed distinct care pathways.**
- ✓ **Ideally, covariate assessment period should stop at the time treatments are decided upon and not yet started.**

***The Authors declare that there is no conflict of interest.***