OHDSI symposium 2020

I. I., Nonneman, MSc[1]; W., Karman, Ph.D.[1]; K. Antonio, Ph.D.[2,3]; V. Tozzi, MSc[4]; J. Krijgsman van Spangenberg, MSc[1]

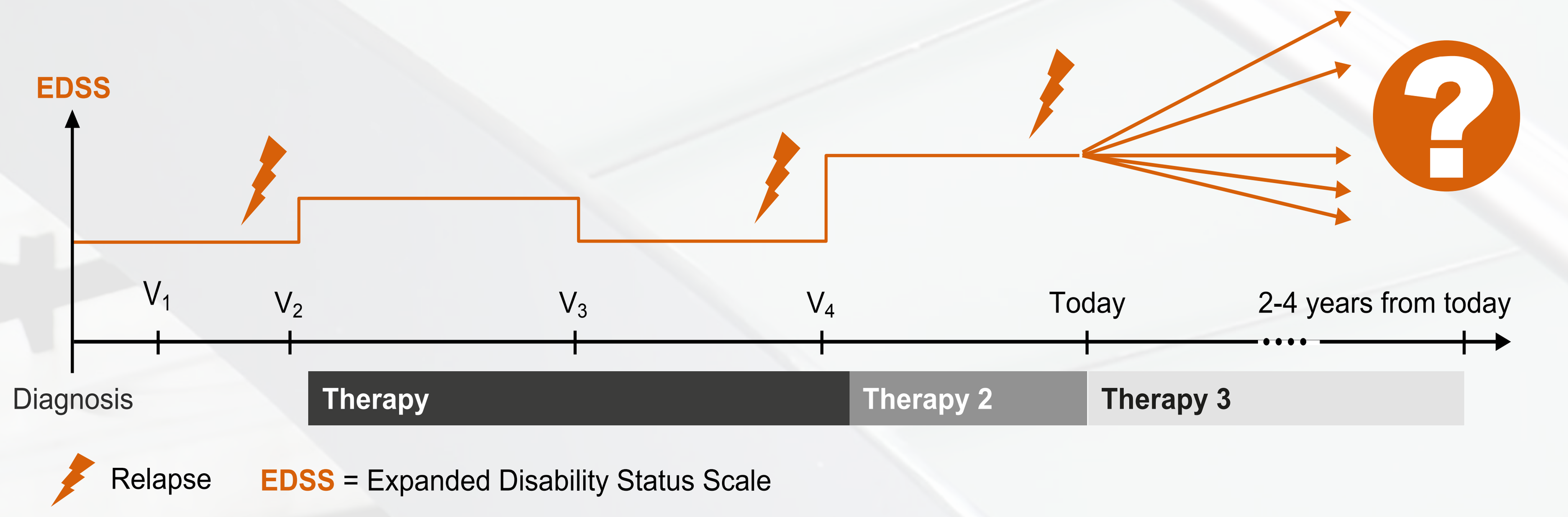
II. 1 PwC, Amsterdam, The Netherlands;
2 Faculty of Economics and Business, University of Amsterdam, The Netherlands;
3 Faculty of Economics and Business, KU Leuven, Belgium;
4 PricewaterhouseCoopers AG, Zürich, Switzerland

Personalised prediction of the time to the first relapse for Relapsing Remitting Multiple Sclerosis

Research goal

- Research shows that **personalised treatment** for relapsing remitting multiple sclerosis (RRMS) patients **can advance treatment effectiveness**
- This study contributes to the personalisation of RRMS treatment by predicting of the time to the first relapse (TTF) of RRMS patients based on individual characteristics and treatment choice.
- We extend the study of *Stühler et al. (2020)*, who present a framework for personalised prediction of treatment response.
- Their model has been implemented in the treatment-decision tool PRHEND which is nowadays successfully used by neurologists in Germany.

Variable of interest



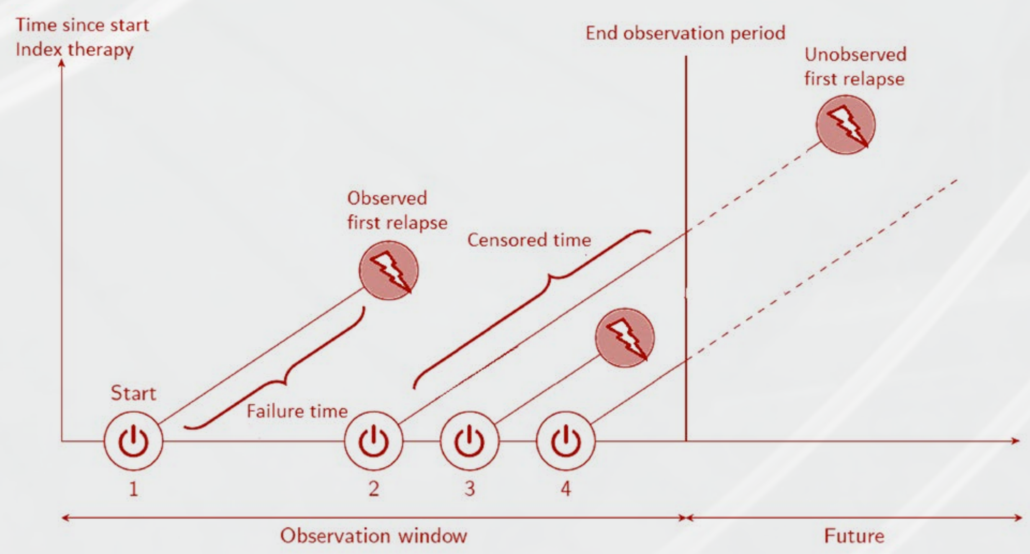
Indicator of treatment response
Which treatment can increase the **time to the first relapse (TTF)** after a patient switches therapies?

Time to event data

- Since 2008, **NeuroTransData GbmH**, a Germany-wide network ofphysicians has maintained a database that currently documents more than 20.000 MS patients.

Data characteristics:

- Following privacy measures, a set of simulated observations is created to perform the analysis.
- The database contains **demographic data**, such as patient's age and gender, as well as **clinical data**, such as patient's quality of life, diagnosis and treatments.
- The final data set includes 3000 observations and around 12 relevant features.
- The data is **highly right-censored** (70%) and includes **ties**.



Methodology

To propose the most effective treatment per patient, the TTF is used as therapy effectiveness indicator. To predict the TTF based on demographical and clinical covariates, we firstly model relapse time using 3 different models:

1. **The Cox Proportional Hazards model (Cox PH model)**
   Semiparametric,
   MLE with Efron's partial likelihood
   $$\lambda_c(t_j, x_i) = \lambda_0(t_j) e^{x_i^T \beta}$$
2. **The Negative Binomial Time to event model (NBT model)**
   Assumes a Negbin. distribution for the # of relapses
   MLE
   $$\lambda_T(t, x) = \frac{\theta e^{x^T \beta} \mu_0'(t, \alpha)}{\theta + e^{x^T \beta} \mu_0(t, \alpha)}$$
3. **Gradient Boosting Machine (GBM)**
   A loss function based on Efron's partial likelihood

Statistical validation

Validation technique:

- Training set (90%), validation set (10%)

Model comparison using two statistical measures:

- Harrell's concordance statistic
- Brier Score

To handle the high percentage of right-censoring in the data, **Inverse Probability of Censoring Weights** is applied to the statistical measures to obtain a better reflection of the population. The censoring distribution is modelled by a marginal Kaplan-Meier estimate, under the assumption of random censoring.

Results and future prospects

Results

- Based on model results, clinical information and demographics prove to be important predictors for the TTF
- GBM performs statistically the best
- Models overall good in predicting survival probabilities on short timeline (1 year), predictions worsen in long run
- GBM displays most diversity in therapy ranking

Future prospects

- Combining Cox PH model and GBM, to open the black box and create explainable models

pwc

Contact: iris.nonneman@pwc.com