A prediction model of heterogeneous treatment effects using randomized and observational data

Konstantina Chalkou, Salanti Georgia – University of Bern

Data from:
Fabio Pellegrini – Biogen International GmbH,
Suvitha Subramaniam, Benkert Pascal - University of Basel
One size to fit all

Network meta-analysis is often used in HTA to estimate relative effects between competing treatments.

Synthesis of a network of 5 studies (3910 patients)

Compare Dimethyl Fumarate, Glatiramer Acetate, and Natalizumab in patients with relapsing-remitting MS

Outcome: Relapse at 2 years (binary)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>OR (95% CI)</th>
<th>vs</th>
<th>vs</th>
<th>vs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimethyl fumarate</td>
<td>1.24 [0.84, 1.83]</td>
<td>-</td>
<td>0.71 [0.54, 0.94]</td>
<td></td>
</tr>
<tr>
<td>1.17 [0.84, 1.66]</td>
<td>Glatiramer acetate</td>
<td>-</td>
<td>0.63 [0.45, 0.89]</td>
<td></td>
</tr>
<tr>
<td>2.3 [1.4, 3.7]</td>
<td>1.95 [1.16, 3.28]</td>
<td>Natalizumab</td>
<td>0.31 [0.20, 0.46]</td>
<td></td>
</tr>
<tr>
<td>0.71 [0.54, 0.93]</td>
<td>0.60 [0.44, 0.83]</td>
<td>0.31 [0.20, 0.47]</td>
<td>Placebo</td>
<td></td>
</tr>
</tbody>
</table>
One size does not fit all

Treatment choice is (or should be) personalised
Not all patients have the same response to the same treatment

**Heterogeneous Treatment Effects**
Effect modification operates in many treatments and setting
So, the optimal treatment depends on patients characteristics
Patient A

Predicted probabilities to relapse in two years
Dimethyl Fumarate - 25% / Glatiramer Acetate - 28% / Natalizumab - 27% / Placebo - 50%

Ranking of predicted probabilities to relapse in two years
1. The lowest probability to relapse is under treatment:
Dimethyl Fumarate with 24.8% probability to relapse.
2. Second best choice based on the probability to relapse:
Natalizumab with 26.6% probability to relapse.
Patient B

Predicted probabilities to relapse in two years
Dimethyl Fumarate - 71% / Glatiramer Acetate - 75% / Natalizumab - 58% / Placebo - 84%

Ranking of predicted probabilities to relapse in two years
1. The lowest probability to relapse is under treatment:
   Natalizumab with 58% probability to relapse.
2. Second best choice based on the probability to relapse:
   Dimethyl Fumarate with 71.4% probability to relapse.
One size does not fit all

Treatment choice is (or should be) personalised
Not all patients have the same response to the same treatment
Heterogeneous Treatment Effects
Effect modification operates in many treatments and setting
So, the optimal treatment depends on patients characteristics

At the population level, recommendations should be specific to subgroups and cost-effectiveness will depend on the distribution of effect modifiers within each country
Treatments

- Dimethyl Fumarate
- Glatiramer acetate
- Natalizumab
- Placebo

Predicted Outcome

A

B

C

D

Relative treatment effects

Patient characteristics

Risk score
the probability of the outcome at baseline

Network meta-regression methods

Prognostic modelling methods

IPD from RCTs

IPD from Observational studies or registries
Data

RCTs

3 randomized phase III clinical trials 2990 observations in total
Data

Observational data
Swiss MS Cohort (SMSC)

Patients with confirmed RRMS and at least two-year follow-up period from the baseline visit date

935 patients, each one with 1, 2, or 3 treatment cycles (i.e. repeated measures)

1752 follow-up cycles
Treatments

- Dimethyl Fumarate
- Glatiramer acetate
- Natalizumab
- Placebo

IPD from Observational Data SMSC

Predicted Outcome A
Predicted Outcome B
Predicted Outcome C
Predicted Outcome D

HTE

Risk score

Prognostic model
\[ h(y_i) = \beta_0 + \sum_{j=1}^{n} \beta_j \times PF_{ij} \]

#STAGE1

Re-calibration and re-estimation of the developed risk model and calculation of the new risk score in the RCTs dataset

#STAGE2

Prediction model with IPD Network meta-regression using only the risk score

#STAGE3
**Treatments**

- Dimethyl Fumarate
- Glatiramer acetate
- Natalizumab
- Placebo

**Predicted Outcome**

- A
- B
- C
- D

**IPD from Observational Data SMSC**

**HTE**

**IPD from RCTs**

**Predicted Outcome**

**IPD from**

**Prognostic model**

\[ h(y_i) = \beta_0 + \sum_{j=1}^{n} \beta_j \times PF_{ij} \]

- **#STAGE1**
  - Prognostic model
  - Risk score

- **#STAGE2**
  - Re-calibration and re-estimation of the developed risk model and calculation of the new risk score in the RCTs dataset

- **#STAGE3**
  - Prediction model with IPD Network meta-regression using only the risk score
Stage 1: Development of the prognostic model

Step 1 - Selection of prognostic factors

- Age
- Sex
- EDSS
- Prior MS treatment
- Months since last relapse
- Disease duration
- Nr. of prior relapses
- Number of Gd enhanced lesions

8 previously identified prognostic factors (at least 2 times included in pre-existing prognostic models)
Stage 1: Development of the prognostic model

Step 2 – Development of the model

Logistic mixed effects model in a Bayesian framework

\[ Y_{ij} \sim \text{Bernoulli}(R_{ij}) \]

\[ \logit(R_{ij}) = \beta_0 + u_{0i} + \sum_{k=1}^{P} (\beta_k + u_{ki}) \times PF_{k,j} \]

**Notation**

- **i**: individuals, where \( i = 1, 2, ..., N \)
- **j**: time point, where \( j = 1, 2, 3 \)
- **PF_{k,j}**: kth prognostic factor at jth time point, where \( k = 1, 2, ..., P \)
- **\beta_0**: fixed effect intercept
- **u_{0i}**: random effect intercept
- **\beta_k**: fixed effect slopes of kth prognostic factor
- **u_{ki}**: the individual-level random slopes of kth prognostic factor
Stage 1: Development of the prognostic model

Step 3 – Sample size efficiency

EPV = 13.7

Recommended more than 10

Our sample size is efficient for
- agreement between apparent and adjusted model performance
- precise estimation of risk

- Riley RD. et al., 2018

Our sample size is not efficient for
- avoiding optimism

Addressed via the shrinkage in the next step
Stage 1: Development of the prognostic model

Step 4 – Shrinkage of coefficients
Bayesian shrinkage methods use a prior on the regression coefficients to address the problem of overfitting in prognostic models.
We used Laplace prior distribution on the regression coefficients to shrink the coefficients.

Step 5 – Handling of Missing data
We used Multilevel Joint Modelling Multiple Imputation approach.
We imputed 10 datasets.
## Stage 1: Development of the prognostic model

### Step 6 - Estimated ORs

<table>
<thead>
<tr>
<th>Prognostic factors</th>
<th>Estimations (ORs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.15</td>
</tr>
<tr>
<td>Age</td>
<td>0.97</td>
</tr>
<tr>
<td>Disease Duration</td>
<td>1.38</td>
</tr>
<tr>
<td>Edss</td>
<td>1.12</td>
</tr>
<tr>
<td>Gd enhanced lesions</td>
<td>1.00</td>
</tr>
<tr>
<td>Number of previous Relapses (1 vs 0)</td>
<td>0.92</td>
</tr>
<tr>
<td>Number of previous Relapses (more than 2 vs 0)</td>
<td>1.12</td>
</tr>
<tr>
<td>Months Since Relapse</td>
<td>0.61</td>
</tr>
<tr>
<td>Treatment Naive</td>
<td>1.15</td>
</tr>
<tr>
<td>Gender</td>
<td>0.28</td>
</tr>
<tr>
<td>Treatment During Cycle (Yes vs No)</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Stage 1: Development of the prognostic model

Step 7 – Internal validation

We used **bootstrap** internal validation approach to correct for optimism in discrimination and calibration ability of the developed model.

Optimism-corrected AUC = 0.67

Optimism corrected calibration-slope = 1.00
Stage 1: Development of the prognostic model

Step 8 – R-shiny app

Prevention of relapses in patients with Relapsing-Remitting Multiple Sclerosis

Risk score
the probability of the outcome at baseline

© The HTx Consortium 2019-2023. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement Nº 825162.
**IPD from Observational Data SMSC**

**Predicted Outcome**

- **A** Dimethyl Fumarate
- **B** Glatiramer acetate
- **C** Natalizumab
- **D** Placebo

**Treatments**

**Risk score**

\[ h(y_i) = \beta_0 + \sum_{j=1}^{n} \beta_j \times PF_{ij} \]

**Prognostic model**

**HTE**

**IPD from RCTs**

**Re-calibration and re-estimation of the developed risk model and calculation of the new risk score in the RCTs dataset**

**Prediction model with IPD Network meta-regression using only the risk score**

#STAGE1

#STAGE2

#STAGE3
Stage 2: Re-calibration and re-estimation of the risk model to RCTs

We update the model to improve predictions for new patients from the new setting (i.e. RCTs)

<table>
<thead>
<tr>
<th>Methods</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Update</td>
<td>0.57</td>
</tr>
<tr>
<td>Update only the intercept (Re-calibration)</td>
<td>0.57</td>
</tr>
<tr>
<td>Update intercept and coefficients (Re-calibration)</td>
<td>0.57</td>
</tr>
<tr>
<td>Model revision (Re-calibration &amp; selective re-estimation)</td>
<td>0.61</td>
</tr>
</tbody>
</table>
Risk of relapse in two years in RCTs

AUC = 0.61
Dimethyl Fumarate

Glatiramer acetate

Natalizumab

Placebo

**Predicted Outcome A**

**Predicted Outcome B**

**Predicted Outcome C**

**Predicted Outcome D**

**Treatments**

**IPD from Observational Data SMSC**

**IPD from RCTs**

**Risk score**

**Prognostic model**

\[ h(y_i) = \beta_0 + \sum_{j=1}^{n} \beta_j \times PF_{ij} \]

**Prediction model with IPD Network meta-regression using only the risk score**

**Re-calibration and re-estimation of the developed risk model and calculation of the new risk score in the RCTs dataset**

**#STAGE1**

**#STAGE2**

**#STAGE3**
Stage 3: IPD Network Meta-regression

\[ Y_{ijk} \sim \text{Bernoulli}(p_{ijk}) \]

\[
\text{logit}(p_{ijk}) = \left\{ \begin{array}{ll}
    u_j + B \times (\logit R_{ij} - \logit R_{j}) & \text{if } k = b_j \\
    u_j + D_{bk} + B \times (\logit R_{ij} - \logit R_{j}) + G_{bk} \times (\logit R_{ij} - \logit R_{j}), & \text{if } k \neq b_j
\end{array} \right.
\]

**Notation**

- \( i \): Individuals
- \( j \): study
- \( k \): treatment
- \( b_j \): baseline treatment in study \( j \)
- \( B \): Individual level covariate regression term for Risk / the impact of Risk as prognostic factor
- \( D_{bk} \): the treatment effect of treatment \( k \) versus placebo / fixed effect
- \( G_{bk} \): The interaction of treatment and risk. Different for each treatment vs study’s control / the impact of Risk as effect modifier

Saramago et al., 2012
## Stage 3: IPD Network Meta-regression

### Results: Estimation of model parameters

OR for relapse for one unit increase in logit-risk in untreated patients (placebo) - \((\text{exp(B)}) = 2.7\) (2.1, 3.9)

<table>
<thead>
<tr>
<th>Drug</th>
<th>OR for relapse versus placebo at the study mean risk ((\text{exp(D)})) &amp; 95% Cr. Intervals</th>
<th>OR versus placebo for one unit of increase in the logit risk ((\text{exp(G)})) &amp; 95% Cr. Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natalizumab</td>
<td>0.28 (0.21, 0.37)</td>
<td>0.60 (0.31, 1.15)</td>
</tr>
<tr>
<td>Glatiramer Acetate</td>
<td>0.53 (0.34, 0.78)</td>
<td>0.73 (0.32, 2.10)</td>
</tr>
<tr>
<td>Dimethyl Fumarate</td>
<td>0.43 (0.3, 0.57)</td>
<td>0.89 (0.50, 1.87)</td>
</tr>
</tbody>
</table>
Stage 3: IPD Network Meta-regression

Results: Estimation of model parameters

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean</th>
<th>Less than 25% Risk</th>
<th>More than 75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natalizumab</td>
<td>46%</td>
<td>33%</td>
<td>57%</td>
</tr>
<tr>
<td>Glatiramer Acetate</td>
<td>61%</td>
<td>43%</td>
<td>75%</td>
</tr>
<tr>
<td>Dimethyl Fumarate</td>
<td>57%</td>
<td>34%</td>
<td>75%</td>
</tr>
</tbody>
</table>

Best treatment
Natalizumab
1% Absolute benefit compared to
Dimethyl Fumarate

Best treatment
Natalizumab-28% Absolute benefit compared to
Dimethyl Fumarate
R-shiny apps

https://cinema.ispm.unibe.ch/shinies/koms/
We developed a new framework to combine observational data and RCTs via prognostic research and network meta-regression. This model allows personalized predictions under several treatment options. Modern estimation and selection methods such as shrinkage are not available in network meta-regression. Our multi-stage model enables their use. The models needs IPD data from (some) RCTs.

**Extensions**

We will use measures relevant to clinical usefulness to validate the model. We will include RCTs that have only aggregated data. We will include cost-effectiveness analysis.