“Reducing risk of reproducing research wrong: creation of a Patient-Level Prediction model library”

Ross D. Williams
Department of Medical Informatics, Erasmus MC
r.williams@erasusmc.nl
BACKGROUND

Lots of prediction models are made, few are validated, even fewer are implemented in clinical practice

This is because of technical difficulties in validating (e.g. different coding systems ICPC, ICD10 etc.)

Infeasibility of application (e.g. excessive numbers of covariates)

Lack of trust in modelling

A new toy is more fun than an old one...
How can we remove these barriers?

OMOP CDM removes many of the technical barriers to replication but doesn’t necessarily help with increasing trust and usability.

This is where a library containing evidence that can be explored interactively adds value.

No need to trawl through rds files on github.
The library will contain the model, the cohorts and the performance as well as various metadata about developer, database and package versions.

At each iteration of maturity a DOI will be created preserving the research at that moment.

The interactivity will allow any party to explore the results of the study themselves.
## How will the library work?

<table>
<thead>
<tr>
<th>T</th>
<th>O</th>
<th>TAR</th>
<th>AUC</th>
<th>Dev</th>
<th># of External Validations</th>
<th>Model</th>
<th>T Size</th>
<th>O Count</th>
<th>O Incidence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP/OP/ER visits of patients presenting with Covid flu or flu-like symptoms AND no symptoms or pneumonia in prior 60d</td>
<td>Hospitalizations with pneumonia</td>
<td>0 to 30 days</td>
<td>0.85</td>
<td>OptumDoD</td>
<td>11</td>
<td>Lasso Logistic Regression - Data Driven</td>
<td>37500</td>
<td>1678</td>
<td>4.47</td>
</tr>
<tr>
<td>GP/OP/ER visits of patients presenting with Covid flu or flu-like symptoms AND no symptoms or pneumonia in prior 60d</td>
<td>Hospitalizations with pneumonia</td>
<td>0 to 30 days</td>
<td>0.84</td>
<td>OptumDoD</td>
<td>13</td>
<td>COVER-H</td>
<td>2082277</td>
<td>105030</td>
<td>5.04</td>
</tr>
<tr>
<td>GP/OP/ER visits of patients presenting with Covid flu or flu-like symptoms AND no symptoms or pneumonia in prior 60d</td>
<td>Hospitalizations with pneumonia or ARDS or sepsis or AKI requiring intensive services or resulting in death in 30d</td>
<td>0 to 30 days</td>
<td>0.84</td>
<td>OptumDoD</td>
<td>12</td>
<td>COVER-I</td>
<td>2082277</td>
<td>29905</td>
<td>1.44</td>
</tr>
<tr>
<td>GP/OP/ER visits of patients presenting with Covid flu or flu-like symptoms AND no symptoms or pneumonia in prior 60d</td>
<td>Persons who die</td>
<td>0 to 30 days</td>
<td>0.90</td>
<td>OptumDoD</td>
<td>13</td>
<td>COVER-F</td>
<td>2082277</td>
<td>11407</td>
<td>0.55</td>
</tr>
<tr>
<td>GP/OP/ER visits of patients presenting with Covid flu or flu-like symptoms AND no symptoms or pneumonia in prior 60d</td>
<td>Hospitalizations with pneumonia or ARDS or sepsis or AKI</td>
<td>0 to 30 days</td>
<td>0.87</td>
<td>OptumDoD</td>
<td>13</td>
<td>Lasso Logistic Regression - Data Driven</td>
<td>37500</td>
<td>2617</td>
<td>6.98</td>
</tr>
</tbody>
</table>

[https://rdwilliams.shinyapps.io/predictionlibrary/](https://rdwilliams.shinyapps.io/predictionlibrary/)
WHAT DOES THIS ADD?

Centralised, version controlled location for storing studies

Ability to explore results with a variety of metrics and graphics

Ability to download study package directly and run against local data

This can then be uploaded to increase number of external validations
A database of these models will be created that will be searchable by author, name, model type, setting etc

This is again a focus on increasing the accessibility and usability

This feature will be helpful for clinicians/regulators looking for relevant models as well as prediction researchers looking for a set of problems to apply new methods too.
THANKS FOR LISTENING

Questions?

Email: r.williams@erasmusmc.nl