ScreenMCM: A Machine Learning-Based Product Screening Tool to Accelerate Medical Countermeasure Development

Rahul Kumar Goyal, M.S
Center for Translational Medicine
Department of Practices Sciences and Health Outcomes Research
School of Pharmacy
University of Maryland Baltimore, MD
ML is a New Tool for the Advancement of Precision Medicine

Data

Statistical methods

Mechanistic/Semi-mechanistic methods

Machine Learning

Data Analysis

Application

Jarugula et al. The Journal of Clinical Pharmacology; 2021
MCMs treat ARS and are Approved by US-FDA after Animal Testing

Acute Radiation Syndrome (ARS)

Necrosis of Bone Marrow Cells (Myelosuppression)

Mortality

Cytopenia

Product in development

Tested in Animals

FDA Approved Product

*MCM – medical countermeasure

Failure Rate

75%
**ScreenMCM Accelerates Product Screening**

<table>
<thead>
<tr>
<th>Current status</th>
<th>Trial Duration</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60 days</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 10 20 30 40 50 60</td>
<td></td>
</tr>
<tr>
<td>With tool</td>
<td>10 days</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 10 20 30 40 50 60</td>
<td></td>
</tr>
</tbody>
</table>

**FDA Modernization Act 2.0:** This bill authorizes the use of certain alternatives to animal testing, including cell-based assays and computer models, to obtain an exemption from the US-FDA to investigate the safety and effectiveness of a drug.

ScreenMCM Was Built by Pooling Existing Data

- 3 studies
- 501 non-human primates
- 12 biomarkers
- Daily data for 60 days
- 5 radiation doses

BIG DATA

New Product

Marker 1

Marker 2

Marker 3

Marker 4

✓ Effective (OR)

X Ineffective
Study Designs and Data

3 studies – 60 days each
- **S1** (N=105) (LD₅₀)
  - V1, 24h
  - L1, 24h
  - L2, 48h

- **S2** (N=288) (LD₇₀)
  - V1, 48h
  - V2, 48h + Azithromycin
  - L1, 48h
  - L2, 72h
  - L3, 96h
  - L4, 120h
  - L5, 48h + Azithromycin

- **S3** (N=108) (LD₅₀, LD₇₀)
  - V1, 48h, LD₅₀
  - V2, 48h, LD₇₀
  - L1, 48h, LD₅₀
  - L2, 48h, LD₇₀

**12 Biomarkers**
- RBC, HGB, HCT, RETI, PLAT
- WBC, ANC, ALC, MONO, LGUNSCE, BASO, EOS

**Blood sampling times**
- 1-30 days – Daily
- 31-60 days – Every 3 days

**# Animals**
- 501
- ~15k rows of data

**Longitudinal measurements**

---

1. Final Study Report, Study14-045 Lovelace Biomedical Research Institute, Doyle-Eisele M, et al.
2. Final Study Report, Study TSK-0144 CiToxLabs, Ascah A, et. al.
3. Final Study Report Study 1017-3493, CitTox Labs, Pouliot M, et. al.
Workflow to Predict Mortality due to ARS Using Supervised Machine Learning

01 Data Preparation
- Data from a time frame
- Predictors: Raw levels and/or aggregation of time windows
- Outcome: 60-day mortality

02 Train-test split
- 60% Training
- 40% Testing

03 Model Fitting
- Elastic-net regression
- XGBoost
- Support vector machines
- Random forest

04 Qualification
- ROC-AUC
- Accuracy
- Calibration Plot
- Spiegelhalter Z-test

05 Prediction performance
Time Frame Selection and Data Preparation

- **Depletion starts** – Day 4
- **Recovery starts** – Day 13
- **95% Deaths in Days 12-24**

**Time window aggregation metrics**
- Area under curve
- Slope
- Maximum
- Minimum
- Mean
- Auto-correlation
- Change from baseline
- Daily biomarker levels

**Final dataset**
- 500 animals x 136 predictors

**Data Preparation**
- Recovery starts – Day 13
- Depletion starts – Day 4
- 4-10 days

Line is shown for mean ± standard error of the mean

- Study Day (Days)
- % Decedents
- Platelets x10^3 cells/uL
- Days Post Irradiation

01 Data Preparation
Elastic-net regression algorithm performed similar to other algorithms based on ROC-AUC & Accuracy

**Grid-search CV Results**

- **Accuracy**
- **ROC-AUC**

**Perfect prediction**

**Random guessing**

**Calibration Plot**

- Spiegelhalter Z-test p-value = 0.2

**Predictor Importance**

- CFB – change from baseline
- ACFn – auto-correlation factor with lag of n days
- MIN – minimum
- MAX - maximum

**Plots**

- Mean CV Performance (95% CI)
- Over all Importance

**Predicted Mortality % vs Observed Mortality %**

- Calibration Plot

**Legend**

- E_NET: Elastic-net regression
- RF – random forest
- SVM – support vector machine
- XGB - XGBoost

- MEAN CV Performance (95% CI)

- Overall Importance
Final Elastic-net Regression Model Provides Greater than 70% Accuracy and ROC-AUC on the Test (Unseen) Dataset

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value(^1)</th>
<th>95% CI(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.71</td>
<td>(0.66, 0.77)</td>
</tr>
<tr>
<td>ROC-AUC</td>
<td>0.75</td>
<td>(0.67, 0.81)</td>
</tr>
<tr>
<td>Balanced Accuracy</td>
<td>0.63</td>
<td>(0.57, 0.7)</td>
</tr>
<tr>
<td>F1 Score</td>
<td>0.80</td>
<td>(0.77, 0.84)</td>
</tr>
<tr>
<td>MCC(^3)</td>
<td>0.31</td>
<td>(0.17, 0.46)</td>
</tr>
<tr>
<td>NPV(^4)</td>
<td>0.63</td>
<td>(0.51, 0.77)</td>
</tr>
<tr>
<td>PPV(^5)</td>
<td>0.73</td>
<td>(0.7, 0.78)</td>
</tr>
</tbody>
</table>

\(^1\)Value refers to test performance on the test dataset
\(^2\)95% CI was obtained using 2000 bootstraps on the test dataset
\(^3\)Matthew’s Correlation Coefficient
\(^4\)Negative Predictive Value
\(^5\)Positive Predictive Value
### Application of ScreenMCM

<table>
<thead>
<tr>
<th>Products in pipeline</th>
<th>Hypothetical treatment effect on platelets</th>
<th>Placebo survival rate</th>
<th>ScreenMCM</th>
<th>Treatment survival rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product 1</td>
<td>No effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product 2</td>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product 3</td>
<td>50%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product 4</td>
<td>✓ 100%</td>
<td>✓ 23%</td>
<td>✓ 23%</td>
<td>✓ 40%</td>
</tr>
</tbody>
</table>
Conclusions

- Increased probability of trial success
- Shorten product development time
- Reduced animal testing
- Continuous process improvement
Machine Learning is Being Expanded Across Therapeutic Areas to Achieve the Goal of Precision Medicine

**Individualized treatment planning for lung cancer**

**Input**
- Inflammatory cytokines
- Patient characteristics
- Treatment plan

**Output**
- Radiation toxicity (pneumonitis)

**Biomarker identification for immunotherapy**

**Input**
- CD8 cell penetration & treatment response

**Output**
- CD8 cell penetration & treatment response
Acknowledgments & Conflicts of Interest

• Partner Therapeutics, Inc. is the project sponsor.

• Data used in this analysis was generated in three NHP studies supporting Leukine®’s FDA-approval as a MCM to treat Acute Radiation Syndrome which were funded by the Office of the Assistant Secretary for Preparedness and Response (ASPR), Biomedical Advanced Research and Development Authority (BARDA), under Contract number HHSO1002013000051.

• Thank you to the collaborators for problem formulation, funding, guidance, and critique
  • Mathangi Gopalakrishnan, University of Maryland Baltimore
  • Joga Gobburu, University of Maryland Baltimore
  • John L. McManus, Partner Therapeutics
Thank You