Collaborative Modeling Project for Predicting Acute Oral Toxicity

Kamel Mansouri1, Agnes Karmaus1, Jeremy Fitzpatrick2, Grace Patlewicz2, Prachi Pradeep2, David Allen1, Warren Casey4, and Nicole Kleinstreuer4

1) ILS, RTP, NC, USA; 2) EPA/NCCT, RTP, NC, USA; 3) ScitoVation LLC, RTP, NC, USA; 4) NIH/NEIHS/DNTP/NICEATM, RTP, NC, USA

Abstract

With an increasing number of chemicals to assess for acute systemic toxicity potential and a lack of sufficiently predictive in vitro approaches, 3D model developers provide an alternative to predict acute oral toxicity and bridge data gaps. NICEATM, in partnership with the Collaborative Acute Toxicity Modeling Initiative (CATMoS), organized an international collaborative project to develop in silico models for predicting acute oral toxicity (1). Instead of groups participating and testing 15 predictive models each using a dataset of 11,682 chemicals, models were developed for five endpoints: LD50 value, EPA hazard categories, GHS hazard categories, very toxic (LD50 < 50 mg/kg), and non-toxic (LD50 ≥ 2000 mg/kg). The resulting consensus predictions leverage the strengths and overcome the limitations of individual modeling approaches. The CATMoS approach to combine the five independent calls (2) and the use of the CATMoS consensus tool for predicting acute oral toxicity (3) demonstrates that consensus predictions can be generated for new chemical structures and are made available as open-source models via the OPERA predictive tool, which provides applicability domain assessments and accuracy estimates (4-5). CATMoS predictions for acute oral toxicity can be used to fill data gaps for chemicals and identify new chemical entities for regulatory assessment.

Project Data

- Endpoints: five endpoints were selected by the ICCVAM ATWG member agencies to serve as endpoints for predictive modeling within the CATMoS project.
- Collected data: 34,508 rat oral LD50 values for 16,297 chemicals selected by the ICCVAM ATWG.

15,688 chemical structures
21,205 LD50 values

Collaborators

A consortium of 35 participants/groups from around the globe representing academia, industry, and government.

Models Performance Evaluation

Single models evaluation. Resulting scores (per model) from the evaluation procedure. Evaluation scores (Sn, Sp, Balanced accuracy (BA), Sensitivity (Sn), Specificity (Sp)) for LD50, EPA, and GHS categories.

Models Evaluation Table

<table>
<thead>
<tr>
<th>Models</th>
<th>LD50 (32 models)</th>
<th>EPA (26 models)</th>
<th>GHS (24 models)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Sn</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Average Sp</td>
<td>0.99</td>
<td>0.97</td>
<td>0.90</td>
</tr>
<tr>
<td>Balanced accuracy</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Consensus models evaluation

- Definitions and algorithms
- Robustness: balance between (Goodness of fit (R2) + Predictivity (RP2) + Robustness (R2))

Conclusions

- CATMoS implementation in OPERA

References


Acknowledgements

This paper was prepared under a cooperative agreement with federal funds from the National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN273201500010C. Disclaimer: The views expressed above do not necessarily represent the official positions of any federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied.