External Validation of STopTox as an Alternative to Animal Testing for Toxicological Assessments

Ricardo Scheufen Tieghi[#a], Cleber C. Melo-Filho[#a], Holli-Joi Martin[a], José Teófilo Moreira-Filho[b], Tripp LaPratt[c], Dave Allen[c], Judy Strickland[c], Alexander Tropsha*[a,d], Nicole Kleinstreuer*[b], Eugene N. Muratov*[a].

[#] These authors contributed equally.

[a] *UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA.*

[b] *National Toxicology Program Interagency Center for Evaluation of Alternative Toxicological Methods (NICEATM), Research Triangle Park, NC, 27711, USA.*

[c] *Inotiv-RTP, Morrisville, 27560, NC, United States*

[d] *Predictive, LLC, Raleigh, 27614, NC, United States*

*To whom correspondence and materials requests should be addressed: 100K Beard Hall, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA; Telephone: (919) 966-2955; FAX: (919) 966-0204; E-mail: nicole.kleinstreuer@nih.gov, alex_tropsha@unc.edu, murik@email.unc.edu.

Competing Interests

AT and ENM are co-founders of Predictive, LLC, which develops novel alternative methodologies and software for toxicity prediction. All the other authors declare no conflicts.

Data Collection Protocol

Integrated Chemical Environment (ICE) Overview:

1. ICE contains highly curated and verified datasets for various specific toxicity endpoints.
2. The datasets can be downloaded by accessing the link:
    a. (https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION, accessed June 2024)
3. Datasets of interest can be directly downloaded as xlsx files from their "datasets tab." All original and complete endpoint-specific datasets from ICE are available in the folder: "ICE Datasets."
    a. File descriptions:
        i. Oral Toxicity – 9110 chemicals for 4 endpoints
        ii. Acute Dermal Toxicity – 275 chemicals for 3 endpoints
        iii. Acute Inhalation Toxicity – 1781 chemicals for 3 endpoints.
        iv. Eye Irritation Toxicity – 454 chemicals for 7 endpoints
        v. Skin irritation/Corrosion – 564 chemicals for 16 endpoints
        vi. Skin Sensitization – 1956 chemicals for 26 endpoints.
4. All datasets were filtered to include only compounds tested *in vivo* using the same OECD Test Guidelines the original STopTox was designed to predict. Specifically,
    a. Eye Irritation/Corrosion
        i. OECD TG 405
    b. Skin Irritation/Corrosion
        i. OECD TG 404
    c. Skin Sensitization
        i. OECD TG 429
    d. Acute Oral
        i. OECD TG 401, 420, 423, 425
    e. Acute Dermal
        i. OECD TG 402
    f. Acute Inhalation
        i. OECD TG 403

Data Collection Protocol:

The search was executed independently for the endpoints available in the STOptox web platform. The endpoints of interest were three acute systemic (acute oral, acute dermal, and acute inhalation toxicity) and three topical (eye irritation/corrosion, skin irritation/corrosion, and skin sensitization).

1. First, we navigated to ICE's "Search" tool.
2. The scope of the search used was "Intersection."
3. For "Chemical Input," all 26 available chemical quick lists were selected.
4. For "Datasets," all *in vivo* assays were included for each endpoint. Using skin sensitization as an example:
    a. Navigate to the sensitization tab and filter for "Murine Local Lymph Node Assay (LLNA)."
5. Select "Run."
    a. In this example, 886 records were found for 322 unique substances.
6. The file was exported, and the CASRN of the chemicals in ICE were cross-referenced with those in STopTox's training set.
    a. Duplicate chemicals or compounds missing SMILES were removed.
    b. The remaining compounds were analyzed, and their hazard category predictions were performed using the STopTox web platform (https://stoptox.mml.unc.edu/, accessed February 2025, version 1.0). The predictions were then compared to the original *in vivo* outcome for external validation.
7. Steps 1-6 were repeated for all endpoints available in the STOptox Web platform.

Endpoint-specific "dataset" filters used:

1. Eye Irritation/Corrosion:
    a. "Rabbit Draize Skin Irritation/Corrosion Test" AND "in-vivo" AND "rabbit," AND "GHS Classification."
2. Skin Irritation/Corrosion:
    a. "Rabbit Draize Skin Irritation/Corrosion Test" AND "in vivo" AND "GHS"
3. Skin Sensitization:

    a. “Murine Local Lymph Node Assay (LLNA)” AND “Mouse” AND (“GHS” OR “Call”)
4. Acute Oral:
    a. “in vivo” AND “Rat Acute Oral Toxicity” AND “Rat” AND (“GHS Classification” OR “LD50”
5. Acute Dermal:
6. “Rat Acute Dermal Toxicity” AND “in vivo” AND (“GHS Classification” OR “LD50”) AND “rat”
7. Acute Inhalation:
    a. “Rat Acute Inhalation Toxicity” AND “in vivo” AND “LC50” AND “GHS Classification”

Model Benchmarking:

1. OECD QSAR Toolbox (https://qsartoolbox.org/, accessed February 2025, Version 4.7.1) was downloaded.
2. PredSkin software: (http://predskin.labmol.com.br/, accessed February 2025, Version 3.0) was used through their online web platform.
3. SMILES lists were created using the same compounds outside the STopTox training set, which are available in GitHub.
4. The SMILES list for each endpoint was loaded into the standalone platform and predicted using the QSAR Toolbox and PredSkin models.
5. Models information:
    a. Eye Irritation:
        i. Test organisms (species): Rabbit
        ii. Endpoint: Primary Irritation Index
        iii. Type of method: In vivo
        iv. Prediction approach: QSAR
        v. SAR/QSAR: Severe Skin Irritation in Rabbit - Danish QSAR DB CASE Ultra model."
    b. Skin Irritation/Corrosion:
        i. Test organisms (species): Rabbit

ii. Endpoint: Primary Irritation Index
iii. Type of method: In vivo
iv. Prediction approach: QSAR
v. SAR/QSAR: Severe Skin Irritation in Rabbit - Danish QSAR DB battery model

c. Skin Sensitization
i. Endpoint: EC3 <OR> S M W N <OR> Skin sensitisation
ii. Type of method: In vivo
iii. Assay: GPMT <OR> LLNA
iv. Prediction approach:
v. SAR/QSAR: Skin sensitization for DASS

d. Acute oral:
i. Test organisms (species): Rat
ii. Endpoint: LD50
iii. Duration: 48 h
iv. Prediction approach: QSAR
v. SAR/QSAR: Acute toxicity in Rat, Oral - Danish QSAR DB ACDLabs model

6. The predictions were compared to the original calls assigned and evaluated using the CCR, SE, SP, PPV, and NPV formulas.