Toxicology Assessment Report
For
Skin Sensitization

ClC1=NC=NC(Cl)=C1

4,6-dichloropyrimidine

REACH*Across*™ 3.1.4 estimates a 96% probability of Skin Sensitization hazard for ClC1=NC=NC(Cl)=C1.

The below resources will aid in completing your IUCLID submission:

1. ECHA - How to use and report (Q)SARs
2. REACH*Across*™ Documentation http://ulreachacross.com
3. REACH*Across*™ QMRF http://ulreachacross.com/Documents/QMRF.xml
4. REACH*Across*™ QPRF http://ulreachacross.com/Documents/QPRF.txt

The below information is supplied to aid in completing a(n) Skin Sensitization submission in IUCLID:

ADMINISTRATIVE DATA

Type of information:

(Q)SAR

Reliability:

2 (reliable with restrictions)

Rationale for reliability:

Results derived from a valid (Q)SAR model and falling into its applicability domain, with adequate and reliable documentation / justification.

Justification for type of information

Software:

http://ulreachacross.com/

Model (incl. version number):

REACH*Across*™ v3.1.4

SMILES or other identifiers used as input for the model:

ClC1=NC=NC(Cl)=C1

Scientific validity of the (Q)SAR model:

- Defined endpoint: Skin Sensitization
- Unambiguous algorithm: REACH*Across*™ provides an unambiguous algorithm definition at http://ulreachacross.com/Documents/reachacross-1.0.0-wp.pdf
- Defined domain of applicability: REACH*Across*™ 3.1.4 defines a probabilistic domain of applicability. Substances predicted with sufficiently high or low probability are included in the domain of applicability.
- Appropriate measures of goodness-of-fit and robustness and predictivity: REACH*Across*™ 3.1.4 uses leave one out cross validation on the ECHA C&L database. These results are reported at http://ulreachacross.com/Documents/reachacross-1.0.0-wp.pdf
- Mechanistic interpretation: N/A

Applicability domain:

- Descriptor domain: Pubchem2D Fingerprints and Similarity Network Features
- Structural and mechanistic domains: Defined by model predictions

- Similarity with analogues in the training set: NA
- Other considerations (as appropriate): NA

Adequacy of the result:

REACH*Across*™ 3.1.4 provides a cross validated prediction of human health hazards. This algorithm is built from billions of chemical similarity comparisons in the UL integrated database and tested on ECHA C&L data. A robust summary is given at http://ulreachacross.com/Documents/reachacross-1.0.0-wp.pdf

Attached Justification:

- Attach the given QMRF (available at http://ulreachacross.com/Documents/QMRF.xml)
- Attach the given QPRF (available at http://ulreachacross.com/Documents/QPRF.txt)

DATA SOURCE

Reference

Title:

REACHAcross-3.1.4 Platform

Reference Type:

Other: other software

Author:

UL Verification Services, Inc.

Year:

2017

Bibliographic source:

http://ulreachacross.com/

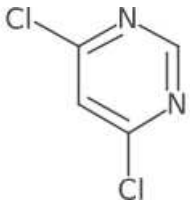
MATERIALS AND METHODS

Principles of method if other than guideline:

REACH*Across*™ v3.1.4 is a hybrid QSAR and read across model that uses chemical similarities paired with supervised learning to make probabilistic estimates on hazard.
Documentation including validation reports and methods information are available at http://ulreachacross.com/Documents/reachacross-1.0.0-wp.pdf
The platform is available at http://ulreachacross.com/

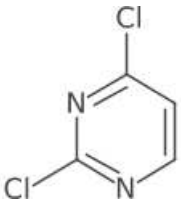
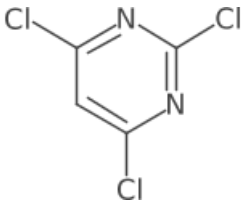
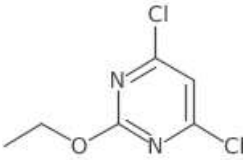
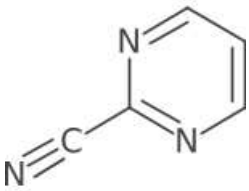
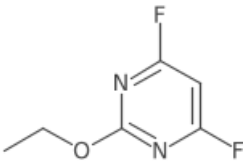
RESULTS AND DISCUSSION

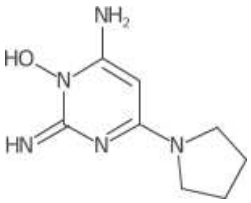
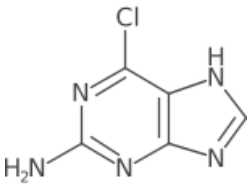
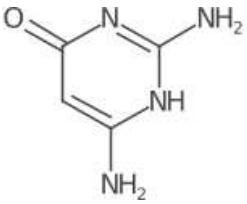
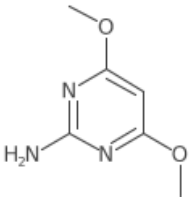
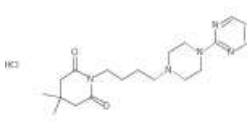
Applicant's summary and conclusion

Target: 

4,6-dichloropyrimidine

ClC1=NC=NC(Cl)=C1

Hazard Analogues	
Structure	Name
 93% Similar	2,4-dichloropyrimidine
 92% Similar	2,4,6-trichloropyrimidine
 84% Similar	4,6-dichloro-2-ethoxypyrimidine
 78% Similar	pyrimidine-2-carbonitrile
 72% Similar	2-ethoxy-4,6-difluoropyrimidine

Non-Hazard Analogues	
Structure	Name
 68% Similar	3-hydroxy-2-imino-6-pyrrolidin-1-ylpyrimidin-4-amine
 65% Similar	6-chloro-7H-purin-2-amine
 62% Similar	2,6-diamino-1H-pyrimidin-4-one
 62% Similar	4,6-dimethoxypyrimidin-2-amine
 58% Similar	4,4-dimethyl-1-[4-(4-pyrimidin-2-ylpiperazin-1-yl)butyl]piperidine-2,6-dione;hydrochloride

Interpretation of results:

Skin Sensitization Hazard

Potency:

Potency Category	Probability
H317: Skin Sens. 1	2%
H317: Skin Sens. 1b	43%
(Skin Sensitization,3)	50%

Conclusions:

REACH*Across*™ v3.1.4 estimates a 96% probability that ClC1=NC=NC(Cl)=C1 is a skin sensitizer

Hazard	Probability
Skin Sensitization Hazard	96%
Not Skin Sensitization Hazard	4%

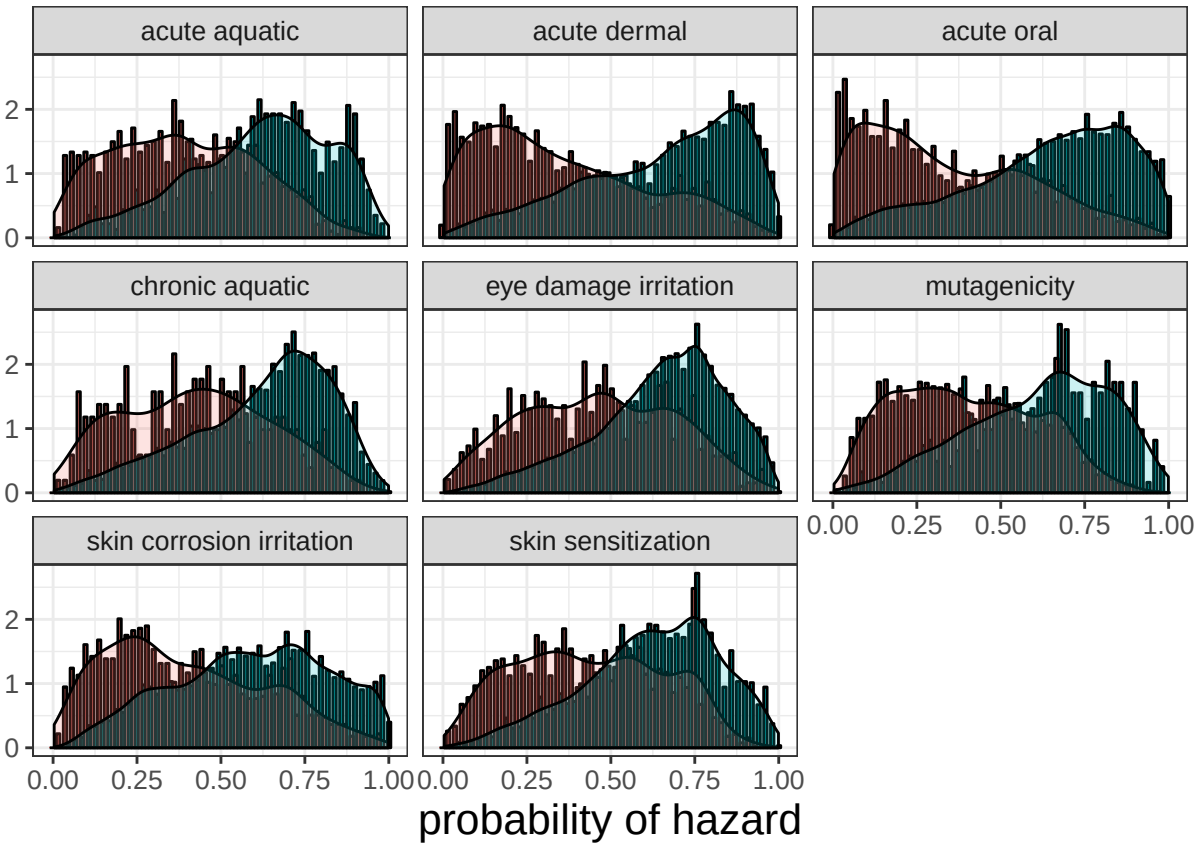
ReachAcross 3.1.4 Binary Evaluation

REACH*Across* 3.1.4 predicts chemical hazard potencies for 8 endpoints:

1. acute dermal (negative, h313, h312, h311, h310)
2. mutagenicity (negative, h340, h341)
3. acute oral (negative, h303, h302, h301, h300)
4. acute aquatic (negative, h402, h401, h400)
5. chronic aquatic (negative, h413, h412, h411)
6. skin sensitization (negative, h317 type 1, h317 type 1b)
7. skin corrosion irritation (negative, h316, h315, h314)
8. eye damage irritation (h320, h319, h318)

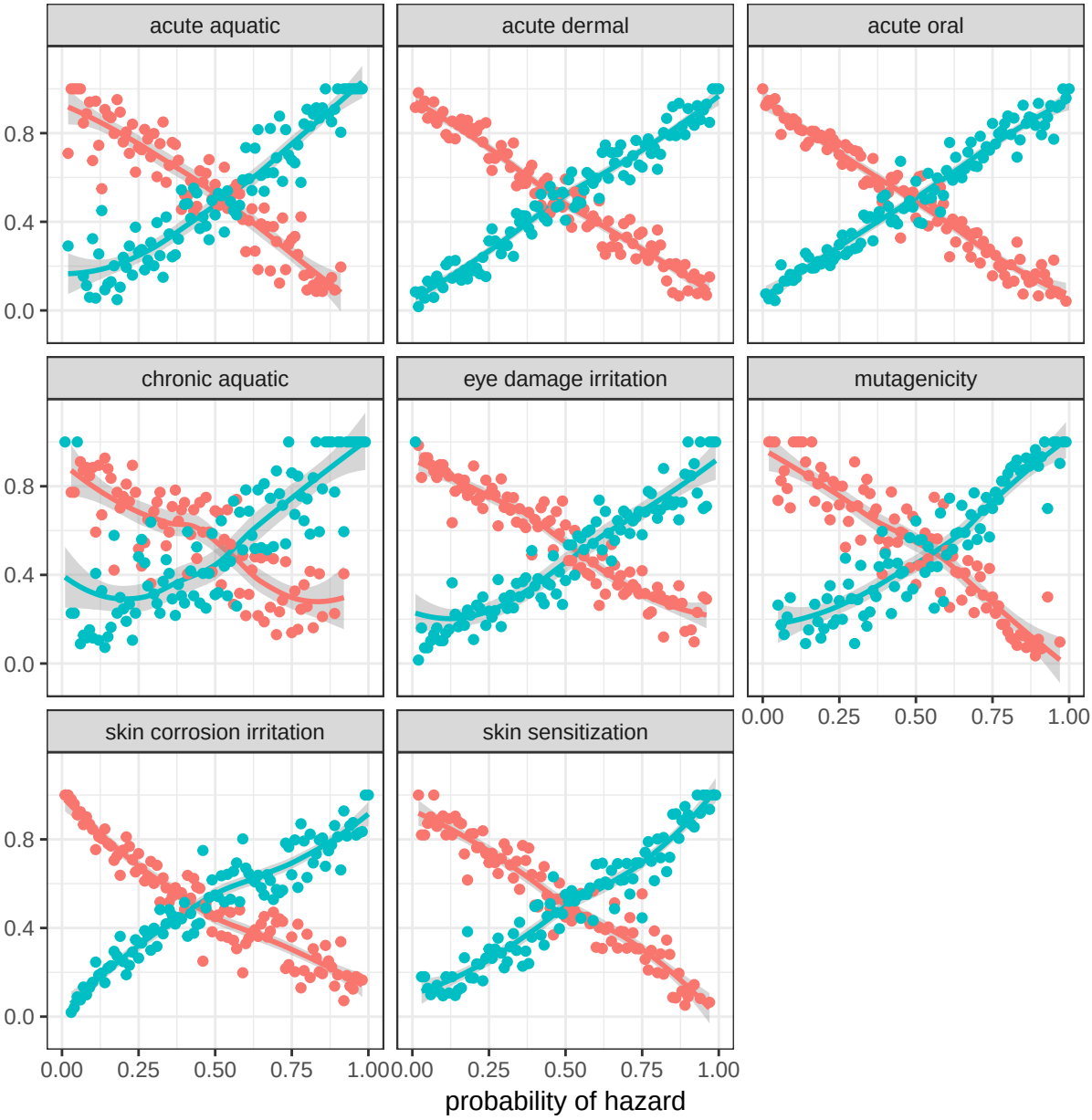
The below figures evaluate REACHAcross performance in predicting binary hazards.

1 Negative / Positive Distribution over Hazard Probability



The above plots are density plots with **blue positives** and **orange negatives**. Ideally the orange density would be on the far left and blue density on the far right.

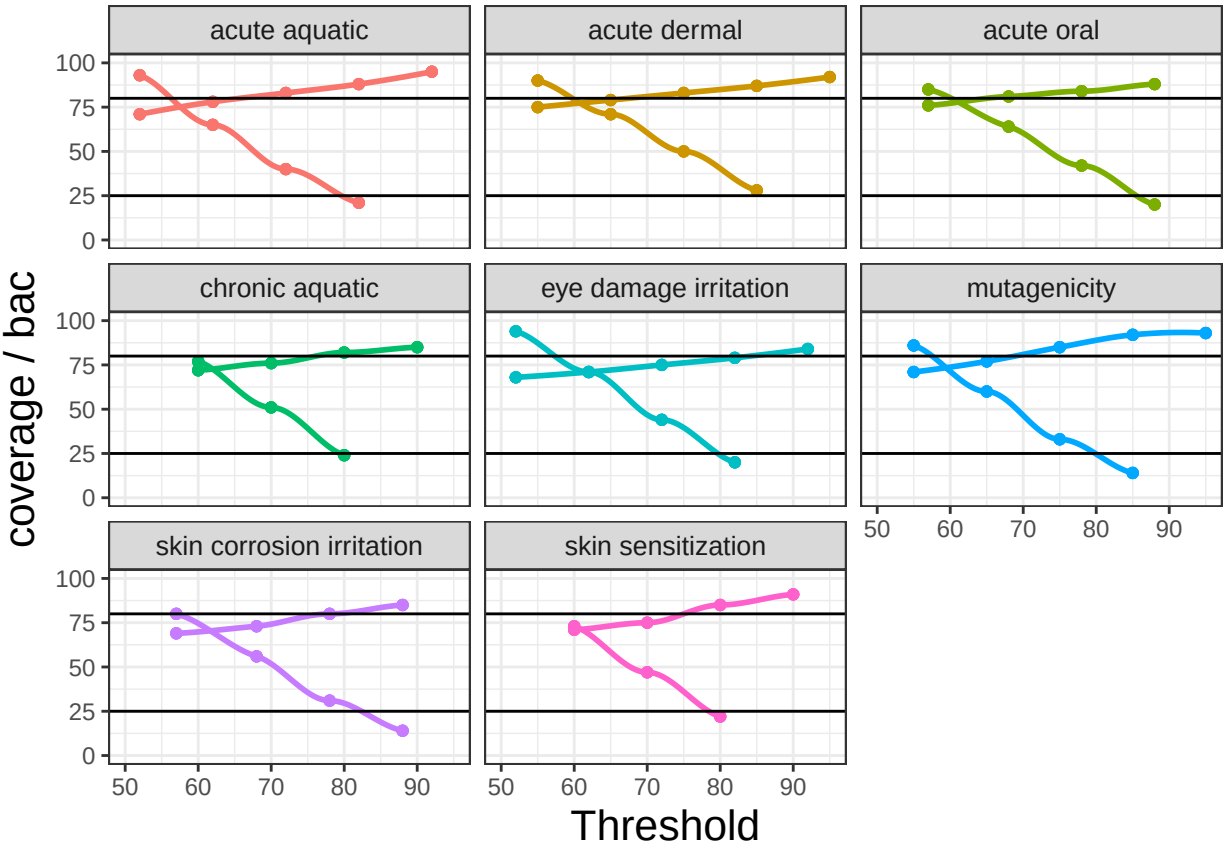
2 Prevalence Adjusted Positive / Negative Ratio



The above plots are **adjusted ratios** of **negatives (orange)** to **positives (blue)** at each probability of hazard. Ratios are adjusted to acount for prevalence. For many of these endpoints there are far more positives than negatives. To account for this difference in prevalence negatives and positives are weighted to create an equal **adjusted prevalence**.

3 Accuracy vs Coverage

In general there is a tradeoff between coverage and accuracy. In the below graphs balanced accuracy increases above 80% (top black line) for every endpoint and coverage drops to 0% for every endpoint as threshold for prediction increases.



4 Sensitivities and Specificities

The terms in the below table are defined below:

1. **negative threshold (negT)**: a probability value below which a negative call is made
2. **positive threshold (posT)**: a probability value above which a positive call is made
3. **Sensitivity (Se)**: Percent of covered hazardous compounds correctly predicted.
4. **Specificity (Sp)**: Percent of covered non-hazardous compounds correctly predicted.
5. **Balanced Accuracy (BAC)**: a measure of accuracy (1/2 sensitivity * 1/2 specificity).
6. **coverage (cover)**: Percent of labeled compounds that falls above/below the pos/neg threshold.

Thresholds were selected to:

1. optimize balanced accuracy
2. optimize coverage
3. maintain equal distance between 0.5 and the negative and positive thresholds

endpoint	pos	neg	total	negT	posT	Se	Sp	BAC	cover
skin sensitization	2865	1886	4751	41%	59%	79%	63%	71%	75%
eye damage irritation	14778	944	15722	41%	59%	83%	56%	70%	79%
acute oral	10225	1932	12157	41%	59%	77%	77%	77%	82%
mutagenicity	600	2795	3395	41%	59%	75%	71%	73%	76%
skin corrosion irritation	13757	1348	15105	41%	59%	69%	71%	70%	76%
acute dermal	4322	1978	6300	41%	59%	77%	75%	76%	82%
acute aquatic	1122	921	2043	41%	59%	76%	74%	75%	77%
chronic aquatic	2552	251	2803	41%	59%	78%	66%	72%	79%

Balanced accuracies are over 70% for every endpoint. Coverage is over 70% for every endpoint. With the exception of skin corrosion/irritation sensitivities are above 70% for every endpoint.