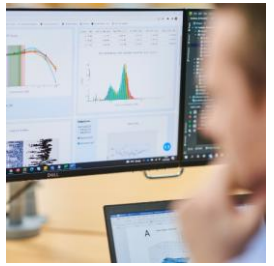
SARA-ICE: A Self-contained Model for Predicting a Human Relevant Point-of-Departure for Skin Sensitization

Georgia Reynolds

EUROTOX 2025

SERS
Safety, Environmental & Regulatory Science

Unilever

SARA-ICE, a Bayesian probabilistic model for skin allergy risk assessment

Database

The core dataset underpinning the model uses data in the NICEATM ICE database.

434 chemicals

1,407 in vivo studies

2,575 in vitro studies

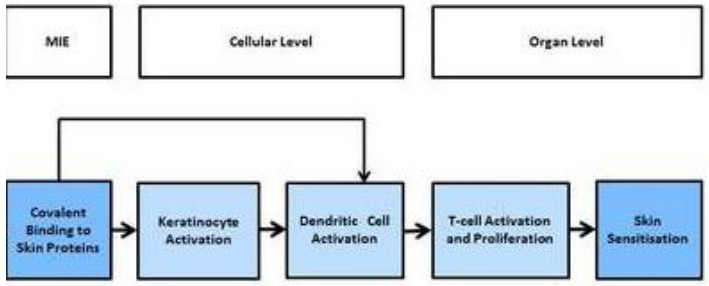
Integrated Chemical Environment

Input Assay Types

OECD TG NAM Assays aligned to key events in the skin allergy AOP.

- DPRA, kDPRA (KE1)
- KeratinoSens (KE2)
- U-Sens, hCLAT (KE3)

- Human (HMT/HRIPT) & LLNA studies may also be used.



Model Outputs

SARA-ICE calculates a continuous measure of sensitiser potency. *ED_{01} (1% sensitising dose in human patch test).*

- A PoD (SARA-ICE DA)

Or

- GHS Classification (SARA-ICE Extended)

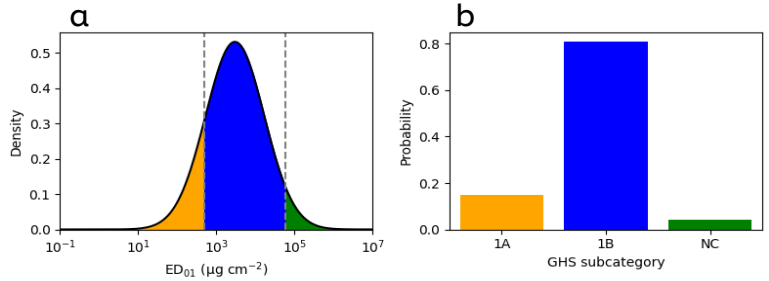


Figure (a) Example estimate of ED_{01} distribution with overlay of GHS subcategories 1A, 1B and NC defined thresholds, (b) probability of each GHS subcategory from ED_{01} distribution

Unilever

OECD Defined Approaches (DAs) for Skin Sensitisation (GL 497)

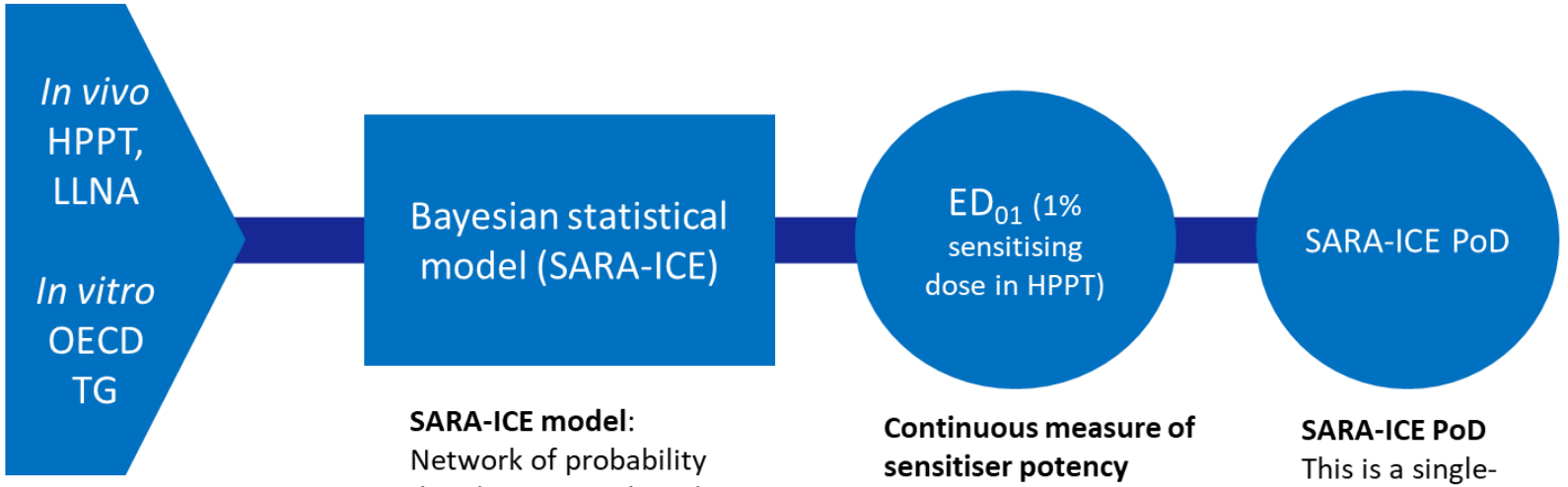
- In 2021, OECD Test Guideline 497 was adopted, meeting regulatory requirements for:
  - DAs that discriminate between sensitizers and non-sensitizers
  - DAs that discriminate strong from weak/moderate sensitizers (i.e., GHS potency categories)

- In 2021, the US and UK began a joint led feasibility study project under OECD for **evaluating a defined approach** that can provide a **point of departure** for quantitative risk assessment
- In 2024, the project began drafting an update to OECD TG 497 to incorporate DAs for PoD determination

- **In June 2025, an update to OECD GL 497, including SARA-ICE DA, was published.**

Unilever

SARA-ICE DA (OECD GL 497)



SARA-ICE database:
434 chemicals
1,407 *in vivo* studies
2,575 *in vitro* studies

SARA-ICE model:
Network of probability distributions to describe associations between all data types

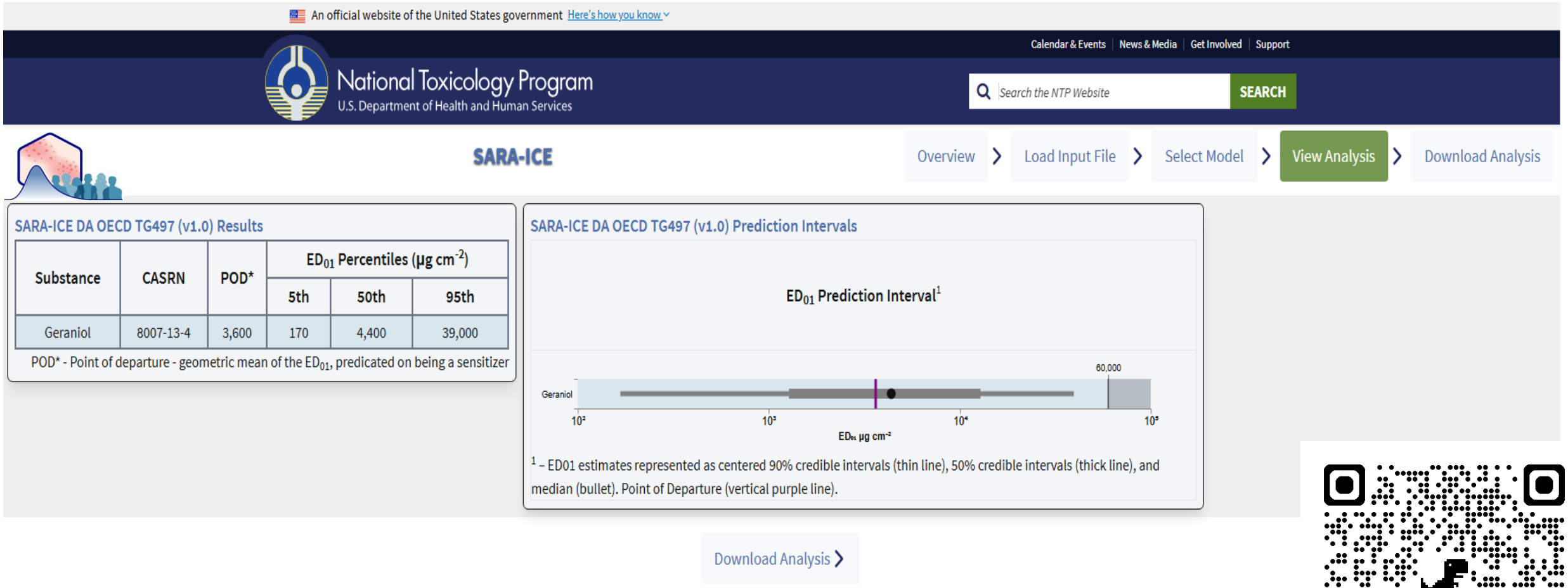
Continuous measure of sensitiser potency
Probability distribution of a random variable defined as the dermal dose required to induce sensitisation in 1% of a HPPT-eligible population.

SARA-ICE PoD
This is a single-valued quantity at a physically possible dose that may serve as a suitable value for input into exposure-based risk assessment.

Assay	HPPT	LLNA	DPRA	kDPRA	KeratinoSens	h-CLAT	U-Sens
Inputs into SARA-ICE	Dermal dose # tested # sensitised	**EC_3** or max. cx. tested if no response	**% depletion** of cys and lys peptide	**Log Kmax**	**$EC_{1.5}$** or max. cx. tested **IC50** or max. cx. tested	**CD86 EC_{150}, CD50 EC_{200}** or max. cx. tested **CV_{75}** or max. cx. tested	**CD86 EC_{150}** or max. cx. tested **CV_{75}** or max. cx. tested

Unilever

SARA-ICE DA User Interface



The SARA-ICE Model User Interface (for OECD 496 DA and Extended versions) is available on the NTP Website.

SARA-ICE NAM PoD: Application in Risk Assessment

Acceptable Margin of Exposure (acceptable MoE) = a value above which a risk assessor may usually conclude low risk for their safety assessment.

We have developed an approach to define NAM PoD equivalent acceptable MoEs so that SARA-ICE PoDs can be applied in current risk assessment methodologies such as QRA.

Published in the **ICCS Best Practise Guidance 2025**.

Table 11. Side by Side Comparison of the Acceptable MoE from a Traditional NESIL-Based Risk Assessment to that for PoD_{NAM} from SARA-ICE

Acceptable MoE for traditional, NESIL-based risk assessment	Median of the distribution for the acceptable MoE for SARA-ICE PoDs
3	3
10	10
30	30
100	100
300	360
1,000	1,700
Source: Reynolds et al. 2025	

Box 8. Example Safety Assessment Using PoD_{NAM} and NAM Acceptable MoE:

Scenario: 0.02% of 'Substance X' in a deodorant

Exposure: CEL = 5 µg/cm^2

PoD_{NAM}: = 4800 µg/cm^2 (derived using SARA-ICE DA)

$MoE = \frac{4800}{5} = 960$

Acceptable MoE = 360 (translated from a traditional Acceptable MoE of 300 for products applied to the underarm (Appendix B. Historically applied Safety Assessment Factors (SAFs)to a NAM Acceptable MoE using Reynolds et al., 2025 as in Table 11)

***Safety Assessment Conclusion*:** No appreciable risk; MoE of 960 > Acceptable MoE of 360

Summary

- SARA-ICE DA fulfils a gap in the current OECD TG 497 on defining a PoD for risk assessment
- SARA-ICE allows flexible use of a range of OECD TG NAMs (as well as historical LLNA/Human data)
- SARA-ICE NAM PoDs can be applied in current skin sensitisation risk assessment methods by applying calculated acceptable MoEs, replacing traditional SAFs

OECD GL 497

SARA-ICE User Interface

ICCS BPG

Unilever

Thank You

SERS

Safety, Environmental & Regulatory Science

georgia.reynolds@unilever.com

seac.unilever.com

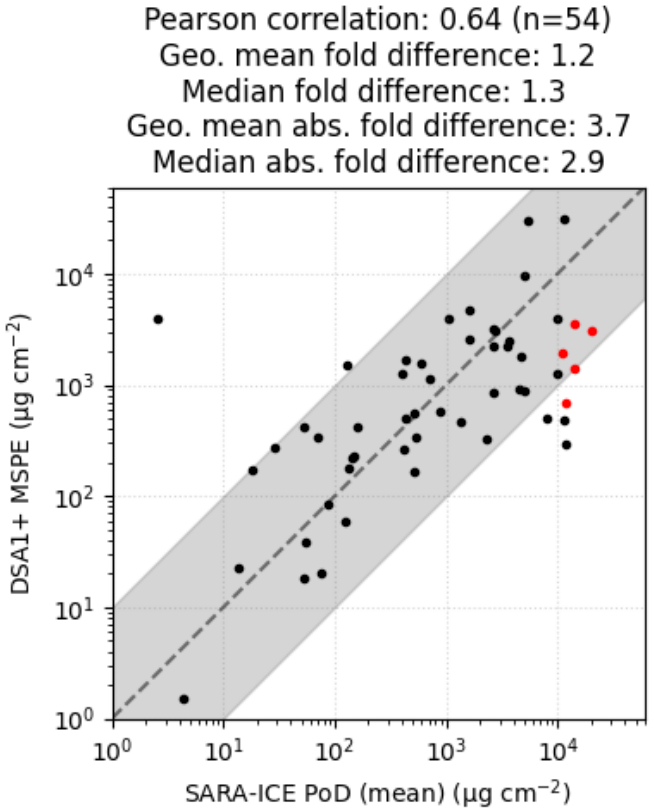
The NICEATM Group

Back up slides

Unilever

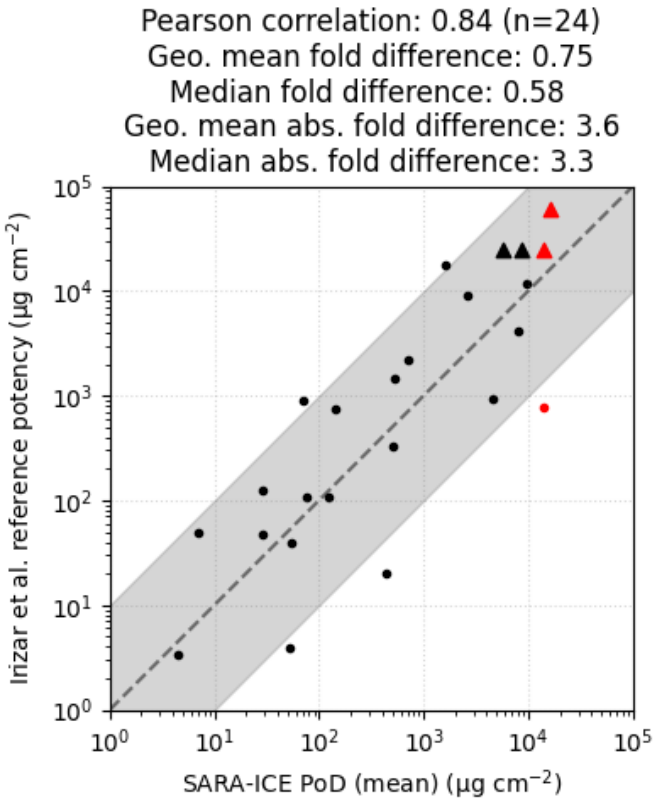
Evaluation of the SARA-ICE PoD

SARA-ICE vs reference DSA1+



SARA-ICE mean PoDs (from NAM data) relatively unbiased relative to reference DSA1+.
PoDs on average around 3-fold away.

SARA-ICE PoDs vs Irizar et al. benchmarks



SARA-ICE mean PoDs (from NAM data) relatively unbiased relative to Irizar et al. reference potency.
PoDs on average around 3.5-fold away.

Red points: compounds predicted to be non-sensitising at a hazard probability threshold of 0.77

SARA-ICE DA (Extended Version)

Input

In vivo HPPT, LLNA

In vitro OECD TG

SARA-ICE database:
434 chemicals
1,407 *in vivo* studies
2,575 *in vitro* studies

Bayesian statistical model (SARA-ICE)

SARA-ICE model:
Network of probability distributions to describe associations between all data types

ED_{01} (1% sensitising dose in HPPT)

Continuous measure of sensitiser potency
Probability distribution of a random variable defined as the dermal dose required to induce sensitisation in 1% of a HPPT-eligible population.

GHS classification thresholds:
Threshold 1A/1B: 500 µg cm^{-2}
Thresholds 1B/NC: 60,000 µg cm^{-2}

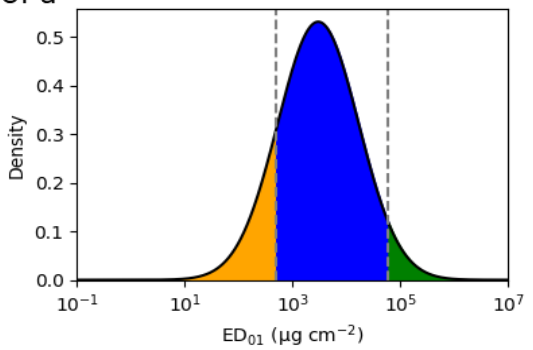
GHS classification probabilities

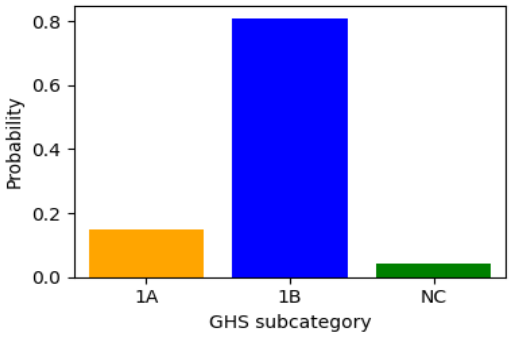
Categorical measure of sensitiser potency
Probability that chemical potency should be categorised as GHS 1A, 1B or NC.

Decision model:
Call 1 if $P(1) > \theta_{bin}$
Call NC if $P(NC) > \theta_{bin}$
Call 1A if $P(1A \mid 1) > \theta_{sub}$
Call 1B if $P(1B \mid 1) > \theta_{sub}$

GHS classification decision model

GHS classification
GHS call if probability passes thresholds chosen within the decision model



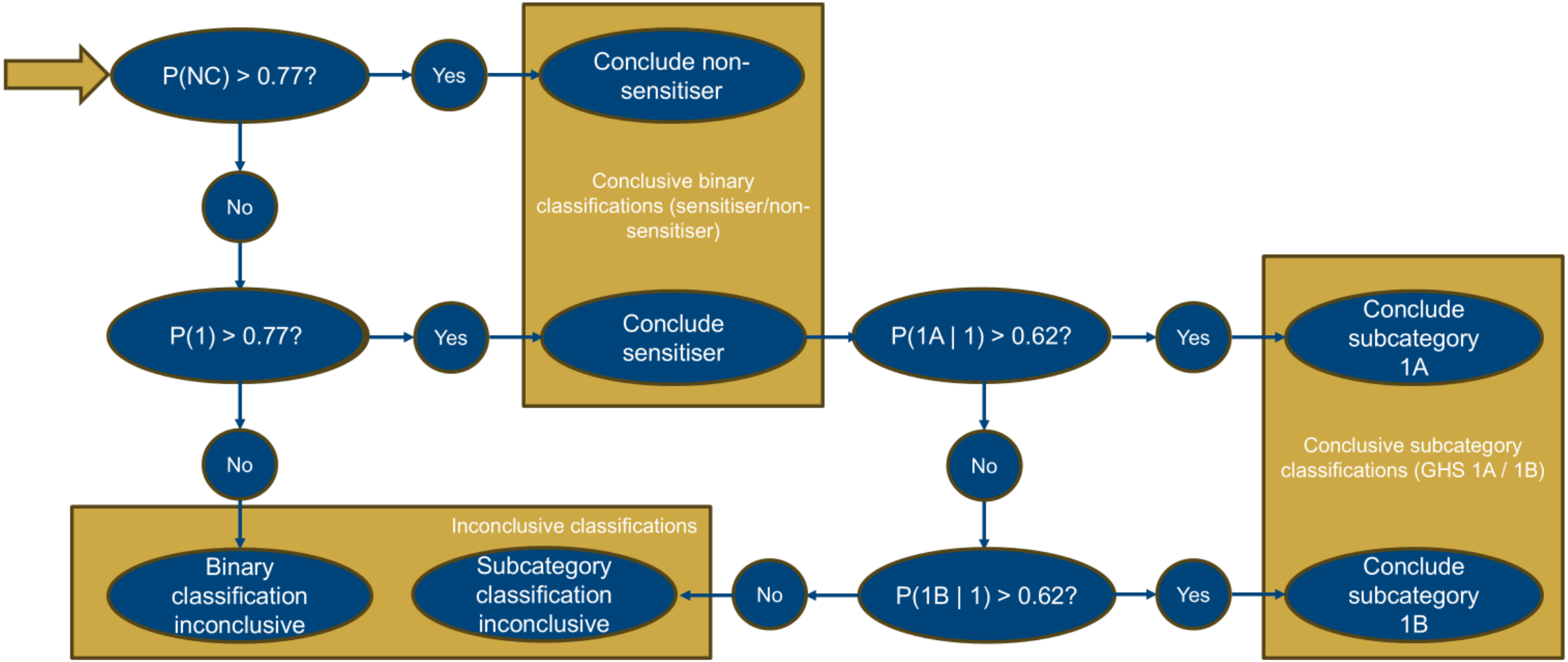


θ_{bin} = selected probability threshold for making a binary classification (1/NC)

θ_{sub} = selected threshold for making a sub-classification of 1A of 1B, contingent on class 1 being true

Unilever

GHS Classification Decision Model (SARA-ICE Extended)



SARA-ICE NAM vs OECD DASS benchmarks

The SARA-ICE decision model has been evaluated against OECD benchmark classifications.

Binary classifications

Human, $\Theta_{bin} = 0.77$	SARA-ICE 1	SARA-ICE NC	Inconclusive	Total
Reference 1	37	5	13	55
Reference NC	0	5	6	11
Total	37	10	19	66

Sensitivity: 88%
Specificity: 100%
Balanced accuracy: 94%
Inconclusive rate on reference class 1: 24%
Inconclusive rate on reference class NC: 55%

LLNA, $\Theta_{bin} = 0.77$	SARA-ICE 1	SARA-ICE NC	Inconclusive	Total
Reference 1	89	9	37	135
Reference NC	2	19	12	33
Total	91	28	49	168

Sensitivity: 91%
Specificity: 90%
Balanced accuracy: 91%
Inconclusive rate on reference class 1: 27%
Inconclusive rate on reference class NC: 36%

Subcategory classifications

Human, $\Theta_{bin} = 0.77$, $\Theta_{sub}=0.62$	SARA 1A	SARA 1B	SARA NC	Inconclusive	Total
Reference 1A	14	2	0	5	21
Reference 1B	3	7	5	16	31
Reference NC	0	0	5	6	11
Total	17	9	10	27	63

Sensitivity 1A: 88%, Specificity 1A: 85%, Balanced accuracy 1A: 86%
Sensitivity 1B: 47%, Specificity 1B: 90%, Balanced accuracy 1B: 69%
Sensitivity NC: 100%, Specificity NC: 84%, Balanced accuracy NC: 92%
Average balanced accuracy: 82%
Inconclusive rate on reference class 1A: 24%
Inconclusive rate on reference class 1B: 52%
Inconclusive rate on reference class NC: 55%

LLNA, $\Theta_{bin} = 0.77$, $\Theta_{sub}=0.62$	SARA 1A	SARA 1B	SARA NC	Inconclusive	Total
Reference 1A	27	3	0	8	38
Reference 1B	12	22	8	43	85
Reference NC	0	1	19	13	33
Total	39	26	27	64	156

Sensitivity 1A: 90%, Specificity 1A: 81%, Balanced accuracy 1A: 85%
Sensitivity 1B: 52%, Specificity 1B: 92%, Balanced accuracy 1B: 72%
Sensitivity NC: 95%, Specificity NC: 89%, Balanced accuracy NC: 92%
Average balanced accuracy: 83%
Inconclusive rate on reference class 1A: 21%
Inconclusive rate on reference class 1B: 51%
Inconclusive rate on reference class NC: 39%

Estimates of the ED_{01} use NAM data only (1xDPRA, 1xKeratinoSens, 1xh-CLAT, 1xkDPRA).
Sensitivity, specificity and accuracy is computed for **conclusive** classifications only.

SARA-ICE DA (Extended Version)

Calendar & Events | News & Media | Get Involved | Support

National Toxicology Program
U.S. Department of Health and Human Services

Search the NTP Website SEARCH

SARA-ICE

Overview > Load Input File > Select Model > View Analysis > Download Analysis

GHS Thresholds

Users can select Probability Thresholds for GHS Call for their own individual use requirements. Sliders are restricted to minimum allowable thresholds for predicting hazard or sub-category as defined in the evaluation found in Reinke et al., 2025. Evaluation of categorization performance was carried out using thresholds of 0.77 (hazard) and 0.62 (sub-category), as described in Supplementary Data 2: Performance of SARA-ICE against OECD benchmark GHS dataset.

GHS Hazard Probability Threshold
Category 1 vs Not Categorized (NC)
Minimum: 0.67
0.77

GHS Sub-Cat Probability Threshold
1A vs 1B (assuming Category 1)
Minimum: 0.5
0.62

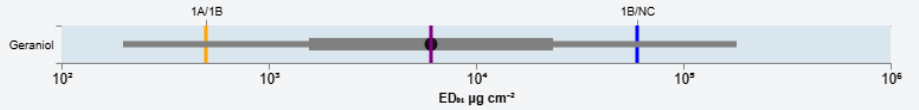
SARA-ICE Extended Model (v1.0) Results

Substance	CASRN	Mean[1]	SPUR[2]	ED_{01} Percentiles (µg cm^{-2})			SARA-ICE Probability GHS Subcategory				GHS Call
				5th	50th	95th	1A	1B	1	NC	
Geraniol	8007-13-4	6,100	30	200	6,000	>60,000	0.11	0.76	0.87	0.13	1B

Mean[1] – Geometric Mean of the ED_{01} distribution
SPUR[2] – (SARA-ICE Prediction Uncertainty Ratio) fold-difference between the median (50th percentile) and the 5th percentile

SARA-ICE Extended Model (v1.0) Prediction Intervals

ED_{01} Prediction Interval[1]



[1] – ED01 estimates represented as centered 90% credible intervals (thin line), 50% credible intervals (thick line), and median (bullet). Geometric Mean (vertical purple line).

Unilever

Application of the SARA-ICE Models

Example Case Study: Geraniol

- Using NAM data only, generate a PoD (SARA-ICE DA) and GHS Classification (SARA-ICE Extended)

SARA-ICE Input Data:

Substance Name	CASRN	MW (g/mol)
Geraniol	8007-13-4	154.25

DPRA	
Depletion Cys (%)	Depletion Lys (%)
12.3	2.6

kDPRA
log Kmax (M-1 s-1)
-3.4

KeratinoSens
EC1.5 (uM)
209.8

hCLAT	
CD54, EC200 (ug/ml)	CD86, EC150 (ug/ml)
>168	123

USENS
CD86, EC150 (ug/ml)
53.6