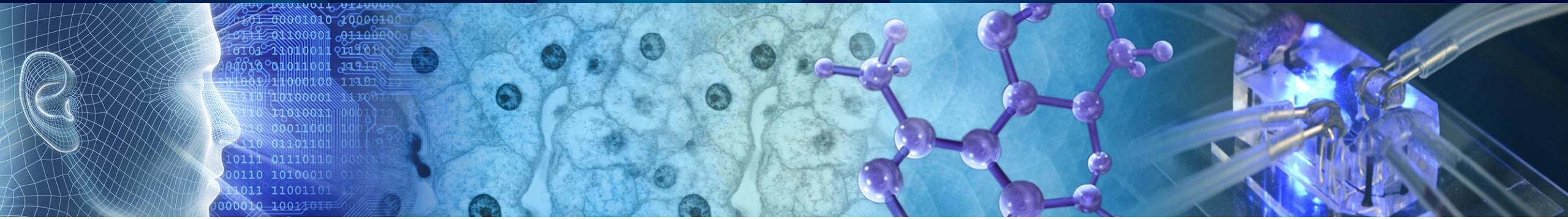
National Institute of Environmental Health Sciences
Division of Translational Toxicology



The SARA-ICE Model for Predicting Skin Sensitizer Potency

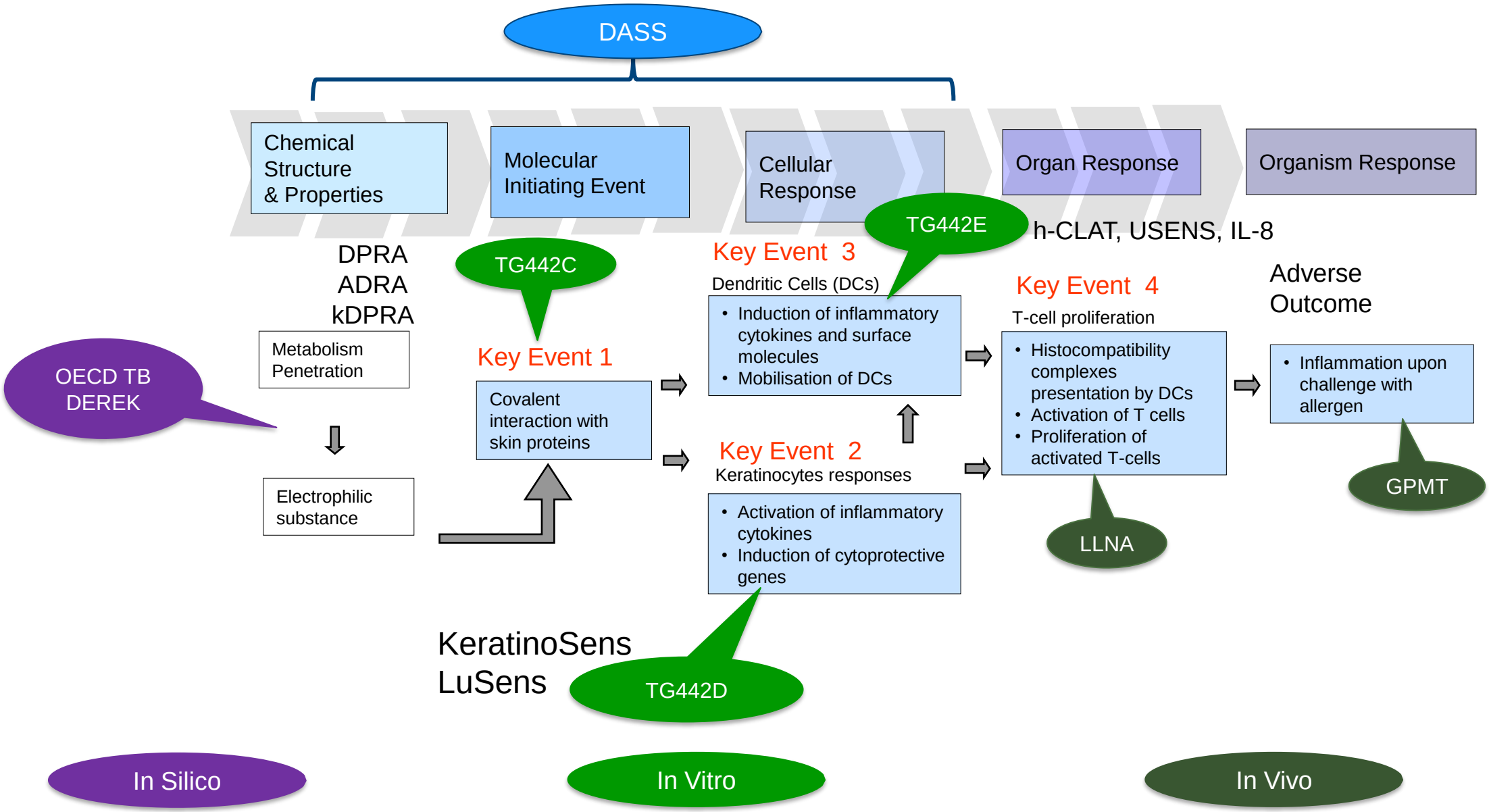
Emily N. Reinke, PhD, DABT,
Inotiv, Contractor supporting the National Toxicology Program Interagency Center for the Evaluation of Alternative Test Methods (NICEATM)

ESTIV Congress 2024
4 June 2024

Disclaimer: Inotiv staff provide technical support for NICEATM, but do not represent NIEHS, NTP, or the official positions of any federal agency.

National Institute of Environmental Health Sciences
Division of Translational Toxicology

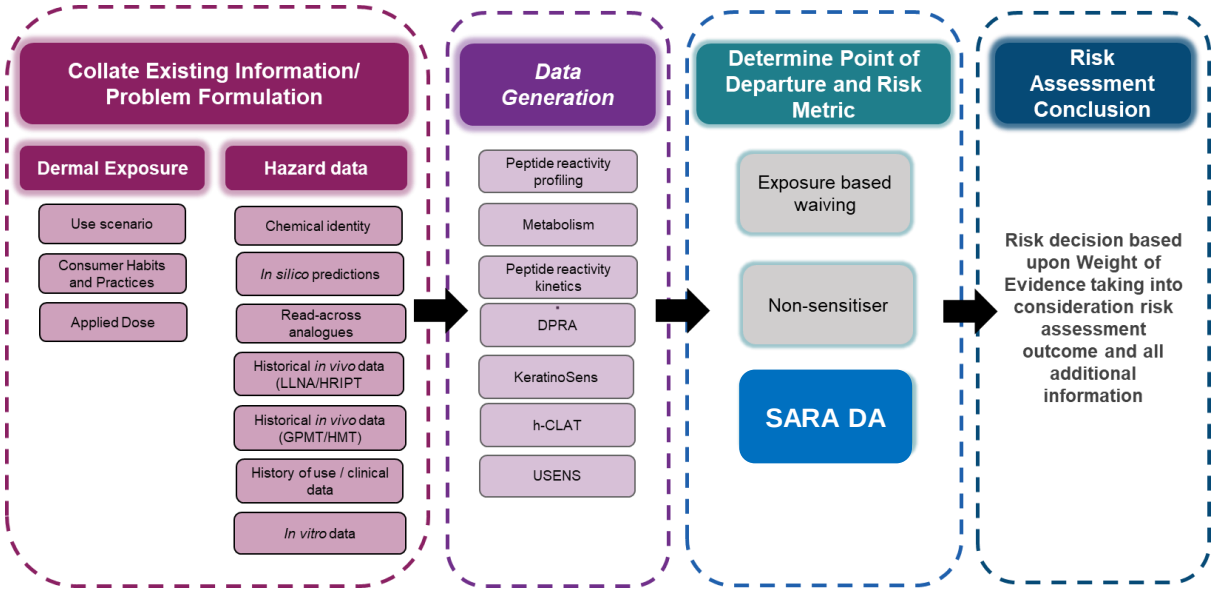
Test Methods Mapped to AOP

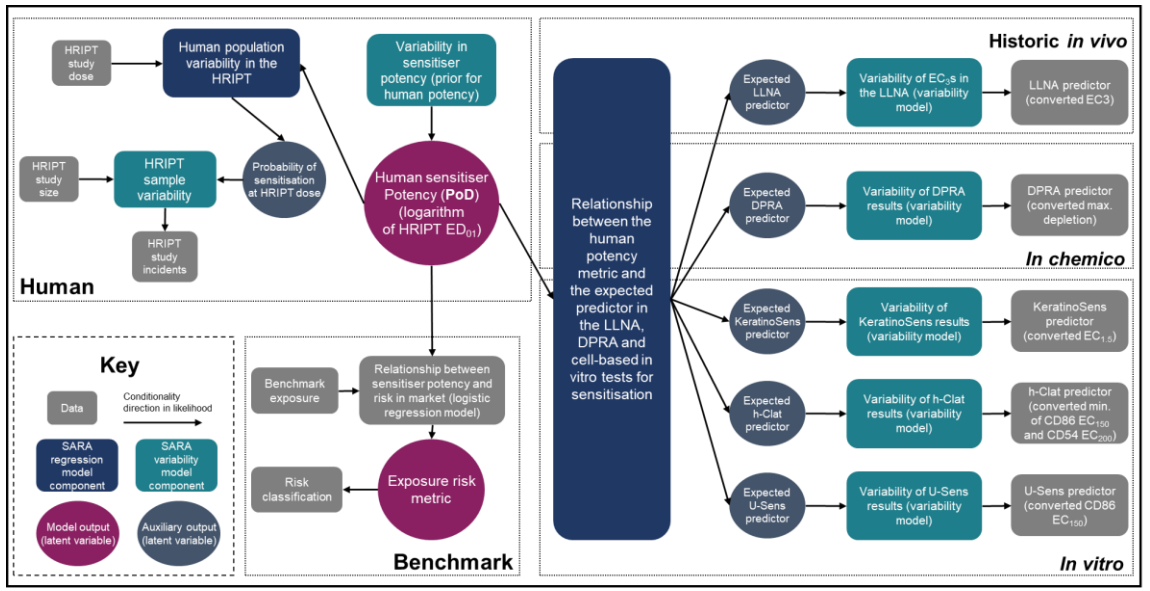


OECD Defined Approaches for Skin Sensitization Guideline Project

- Extensive curation efforts undertaken to build LLNA (168 substances) and human (66 substances) reference databases
- Applicability domain and DA confidence were defined
- The resulting Guideline 497 was adopted in 2021
- It meets regulatory requirements of:
  - DAs that discriminate between sensitizers and non-sensitizers
  - DAs that discriminate strong from weak/moderate sensitizers (i.e., GHS potency categories)
- Future work will cover DAs that address regulatory needs of quantitative risk assessment
  - US and UK leading a project under OECD for evaluating a defined approach that can provide a point of departure for quantitative risk assessment

Skin Allergy Risk Assessment Defined Approach (SARA DA) was developed for application as part of a tiered, WoE NGRA framework





- **Unilever NGRA framework for Skin Allergy** was designed to use a WoE based upon all available information, accommodate range of consumer product exposure scenarios and provide a quantitative point of departure and risk metric → **SARA DA**

The use-case of the SARA DA is to estimate:

1. ED_{01}, the dose at which there is a 1% chance of sensitization in an HPPT-eligible population
2. Probability that a consumer exposure to some chemical is 'low risk', conditional on the available data and the model

Reynolds et al 2022 https://pubmed.ncbi.nlm.nih.gov/35835397/

Unilever Team

Georgia Reynolds
Nicola Gilmour
Joe Reynolds
Gavin Maxwell

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

NICEATM News - 2021 Issue 25: May 27

In this Newsletter:

NICEATM to Collaborate with Unilever on Development of Predictive Model for Skin Sensitization

NICEATM to Collaborate with Unilever on Development of Predictive Model for Skin Sensitization

NICEATM has entered into an agreement with consumer products company Unilever to collaboratively test and further develop their Skin Allergy Risk Assessment (SARA) predictive model. SARA is a computational model that uses a variety of input data to estimate a probability that a chemical will cause an allergic skin reaction in humans. NICEATM will test the SARA model using a variety of chemical data sets, including chemicals of interest to U.S. and international regulatory agencies. NICEATM and Unilever will also work together to expand the SARA model to include data generated by NICEATM. The intent is to make the SARA model openly available for public use along with other NICEATM predictive models. Availability of the SARA model will help further reduce animal use for the endpoint of skin sensitization, and will improve upon existing efforts by providing points of departure for quantitative human risk assessment.

Information about other NICEATM projects to evaluate alternatives to animal use for skin sensitization is available at https://ntp.niehs.nih.gov/go/ACDtest.

Reference: Reynolds et al. Probabilistic prediction of human skin sensitizer potency for use in next generation risk assessment. Comput Toxiol 9:36-49. https://doi.org/10.1016/j.comtox.2018.10.004

NICEATM Team

Nicole Kleinstreuer
Emily Reinke
Dori Germolec
Dave Allen
Tripp LaPratt
Michaela Blaylock
(Judy Strickland)
(Jim Truax)

Modification of SARA DA to create SARA-ICE

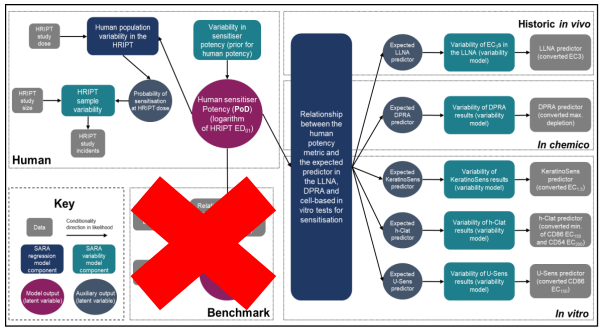
Database
Aim to expand the core dataset underpinning the model using data in the ICE database (relaxing the constraint that chemicals be limited to cosmetic ingredients).

Risk benchmarking
Drop the risk benchmarking component of the model – the current set of benchmarks are limited to use of consumer goods. Use the model for human potency estimation for quantitative risk assessment.

GHS classification
Add functionality to predict GHS classification (estimated as a class probability to communicate uncertainty in classification).

Integrated Chemical Environment

ICE: Integrated Chemical Environment (nih.gov)



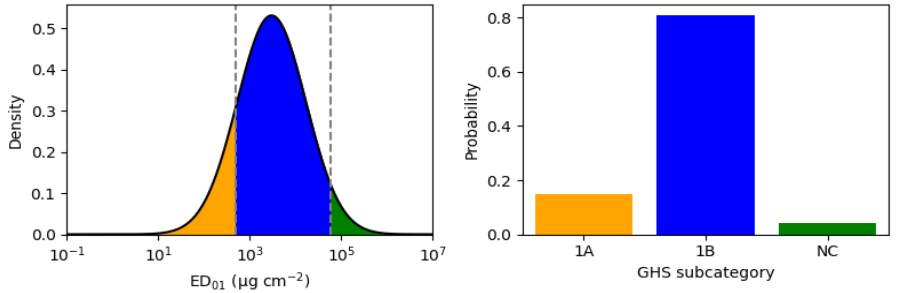
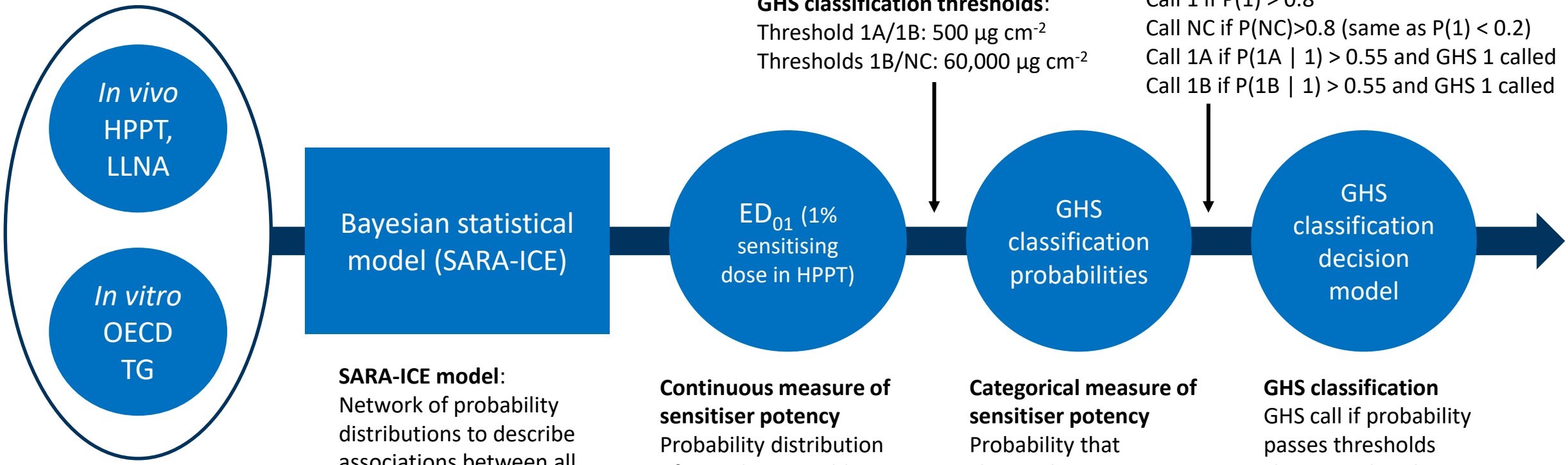


Figure (a) Example estimate of ED_{01} distribution with overlay of GHS subcategories 1A, 1B and NC defined thresholds, (b) probability of each GHS subcategory from ED_{01} distribution

SARA-ICE DA: Skin Allergy Risk Assessment - Integrated Chemical Environment Defined Approach



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.

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

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

The SARA-ICE database

Study type	HPPT	LLNA	DPRA	kDPRA	KeratinoSens	h-CLAT	U-Sens
Inputs into SARA-ICE	Dermal dose, number tested, number sensitised	EC_3 or maximum concentration tested if no response observed	% depletion of cysteine and lysine peptides	Log Kmax	$EC_{1.5}$ or maximum concentration tested IC50 or maximum concentration tested	CD86 EC_{150}, CD50 EC_{200} or maximum concentration tested CV_{75} or maximum concentration tested	CD86 EC_{150} or maximum concentration tested CV_{75} or maximum concentration tested
Number of studies in database	871	536	650	361	972	428	164
Number of unique CASRN with this study type	276	195	251	185	258	211	90

434 distinct CASRN

SARA-ICE NAM vs OECD DASS benchmarks

Binary classifications

Human, θ_{bin} = 0.80	**SARA 1**	**SARA NC**	**Inconclusive**	**Total**
OECD 1	37	4	14	55
OECD NC	0	4	7	11
Total	37	8	21	66
Sensitivity: 90% Specificity: 100% **Balanced accuracy: 95%**				
LLNA, θ_{bin} = 0.80	**SARA 1**	**SARA NC**	**Inconclusive**	**Total**
OECD 1	87	6	42	135
OECD NC	2	19	12	33
Total	89	25	54	168
Sensitivity: 94% Specificity: 90% **Balanced accuracy: 92%**				

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

Estimates of the ED01 use NAM data only (1xDPRA, 1xKeratinoSens, 1xh-CLAT, 1xkDPRA)

Sensitivity, specificity and acccuracy is computed for **conclusive** classifications only.

SARA-ICE NAM vs OECD DASS benchmarks

Subcategory classifications

Human, θ_{bin} = 0.80, θ_{sub}=0.55	**SARA 1A**	**SARA 1B**	**SARA NC**	**Inconclusive**	**Total**
OECD 1A	14	2	0	5	21
OECD 1B	4	9	4	14	31
OECD NC	0	0	4	7	11
Total	18	11	8	26	63
Sensitivity 1A: 88%, Specificity 1A: 81%, Balanced accuracy 1A: 84% Sensitivity 1B: 53%, Specificity 1B: 90%, Balanced accuracy 1B: 71% Sensitivity NC: 100%, Specificity NC: 88%, Balanced accuracy NC: 94% **Average balanced accuracy: 83%**					

LLNA, θ_{bin} = 0.80, θ_{sub}=0.55	**SARA 1A**	**SARA 1B**	**SARA NC**	**Inconclusive**	**Total**
OECD 1A	28	4	0	6	38
OECD 1B	16	22	5	42	85
OECD NC	0	1	19	13	33
Total	44	27	24	61	156
Sensitivity 1A: 88%, Specificity 1A: 75%, Balanced accuracy 1A: 81% Sensitivity 1B: 51%, Specificity 1B: 90%, Balanced accuracy 1B: 71% Sensitivity NC: 95%, Specificity NC: 93%, Balanced accuracy NC: 94% **Average balanced accuracy: 82%**					

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

Estimates of the ED01 use NAM data only (1xDPRA, 1xKeratinoSens, 1xh-CLAT, 1xkDPRA)

Sensitivity, specificity and acccuracy is computed for **conclusive** classifications only.

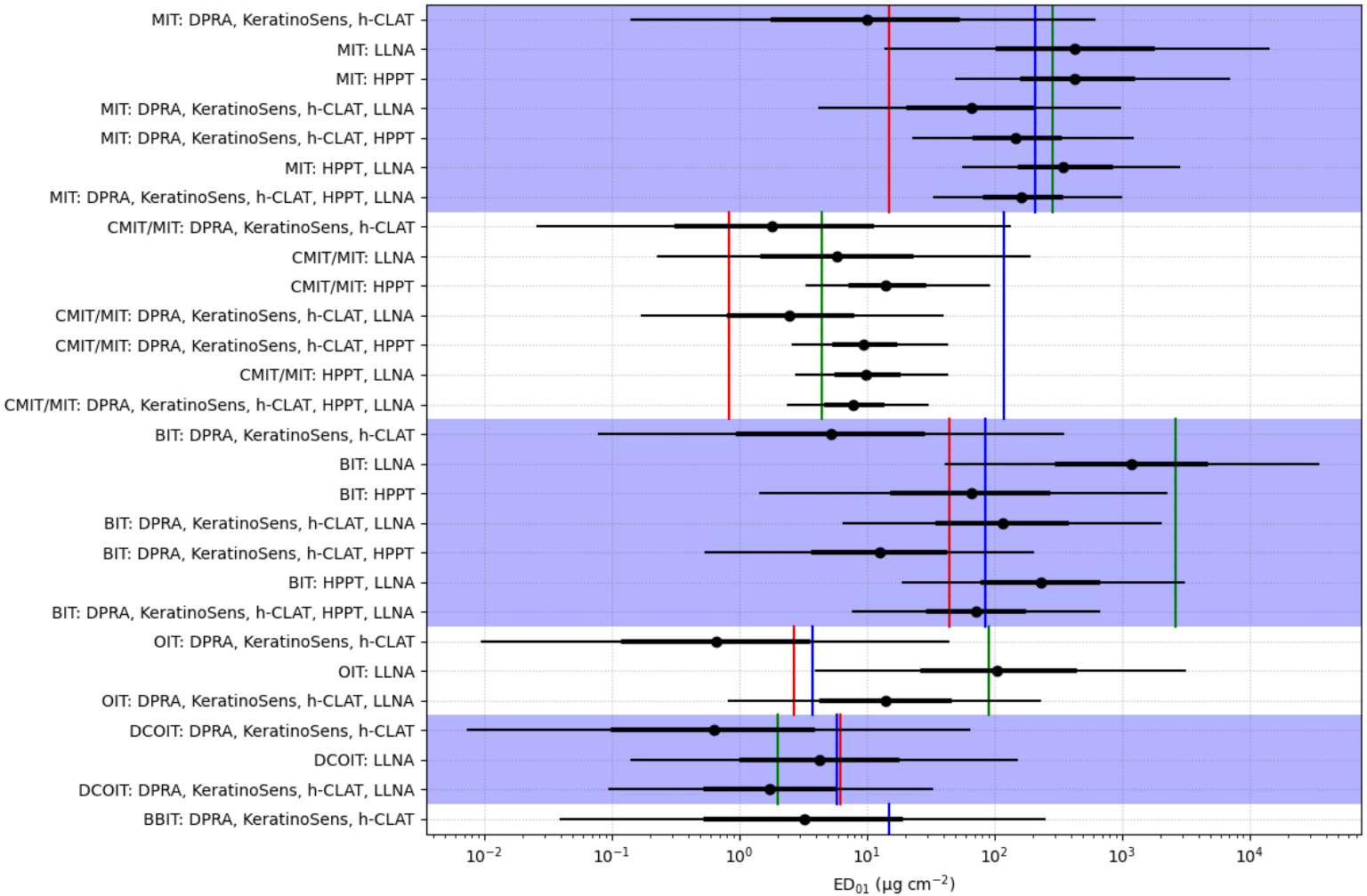
Case Studies

SARA-ICE - Isothiazolinones

Study Type	Chemical						Source
	MIT	CMIT/MIT	BIT	OIT	DCOIT	BBIT	
DPRA	Cysteine depletion: 100% Lysine depletion: 0%	Cysteine depletion: 100% Lysine depletion: 10.6%	Cysteine depletion: 100% Lysine depletion: 0%	Cysteine depletion: 100% Lysine depletion: 1.3%	Cysteine depletion: 100% Lysine depletion: 11.6%	Cysteine depletion: 100% Lysine depletion: 0%	NICEATM IT report, Appendix A, Table 2
KeratinoSens™	$EC_{1.5}$: 9.54 µM IC_{50}: 108 µM	$EC_{1.5}$: 3.41 µM IC_{50}: 19.9 µM	$EC_{1.5}$: 3.14 µM IC_{50}: 57.8 µM	$EC_{1.5}$: 2.19 µM IC_{50}: 12.7 µM	$EC_{1.5}$: 1.32 µM IC_{50}: 4.65 µM	$EC_{1.5}$: 3.84 µM IC_{50}: 53.0 µM	NICEATM IT report, Appendix A, Table 5
h-CLAT	CD54 EC_{200}: 11.6 µg ml^{-1} CD86 EC_{150}: 11.8 µg ml^{-1} CV_{75}: 24.6 µg ml^{-1}	CD54 EC_{200}: 2.63 µg ml^{-1} CD86 EC_{150}: 2.81 µg ml^{-1} CV_{75}: 3.04 µg ml^{-1}	CD54 EC_{200}: 7.63 µg ml^{-1} CD86 EC_{150}: 7.84 µg ml^{-1} CV_{75}: 13.1 µg ml^{-1}	CD54 EC_{200}: 0.95 µg ml^{-1} CD86 EC_{150}: 7.26 µg ml^{-1} CV_{75}: 8.8 µg ml^{-1}	CD54 EC_{200}: 0.92 µg ml^{-1} CD86 EC_{150}: >1.08[1] µg ml^{-1} CV_{75}: 0.9 µg ml^{-1}	CD54 EC_{200}: 3.01 µg ml^{-1} CD86 EC_{150}: 3.15 µg ml^{-1} CV_{75}: 3.3 µg ml^{-1}	NICEATM IT report, Appendix A, Tables 7 & 8
LLNA	EC_3: 0.4% to > 4.5% (4 studies)	EC_3: 0.0049% to 0.048% (9 studies)	EC_3: 1.5% to 32.4% (7 studies)	EC_3: 0.2% to 0.66% (4 studies)	EC_3: 0.0041% to 0.011% (2 studies)		NICEATM IT report, Appendix C
HPPT	DSA: 10 µg/cm^2 to 30 ug/cm^2 N_{tested}: 75 to 210 $N_{sensitised}$: 0 to 1 (6 studies)	DSA: 0.83 µg/cm^2 to 79 ug/cm^2 N_{tested}: 45 to 602 $N_{sensitised}$: 0 to 7 (13 studies)	DSA: 45 µg/cm^2 to 91 ug/cm^2 N_{tested}: 54 to 58 $N_{sensitised}$: 0 to 5 (2 studies)				Strickland et al., 2023; Herzler et al., 2024

Reinke et al., 2024, in draft

SARA-ICE - ED_{01} PoD estimates



ED01 estimates represented as centered 90% credible intervals (thin line), 50% credible intervals (thick line) and median (bullet). Red lines indicate the reference NESIL, blue lines are plotted at the EPA POD and green lines are plotted at the reference LLNA EC3.

NESILs (ECHA; Burnett et al., 2021; Novick et al., 2013; Ladics et al., 2020); EPA POD (EPA DOCKET (https://www.regulations.gov/document/EPA-HQ-OPP-2017-0720-0011); LLNA EC3 (Strickland et al., 2023)

SARA-ICE – MIT example – Probability that an exposure is less than the ED_{01}

Input combination	Exposure (µg cm^{-2})											
	0.01	0.03	0.1	0.3	1	3	10	30	100	300	1000	3000
DPRA	0.93	0.89	0.82	0.75	0.65	0.55	0.43	0.32	0.23	0.16	0.096	0.058
KeratinoSens	0.99	0.99	0.97	0.94	0.88	0.79	0.67	0.54	0.39	0.27	0.16	0.092
h-CLAT	1	1	0.99	0.98	0.94	0.89	0.79	0.67	0.51	0.37	0.23	0.14
DPRA, KeratinoSens	0.98	0.96	0.91	0.83	0.71	0.57	0.41	0.27	0.15	0.084	0.038	0.018
DPRA, h-CLAT	0.99	0.99	0.96	0.92	0.82	0.7	0.53	0.37	0.22	0.12	0.057	0.027
KeratinoSens, h-CLAT	1	1	0.99	0.97	0.93	0.86	0.73	0.58	0.4	0.26	0.14	0.067
DPRA, KeratinoSens, h-CLAT	1	0.99	0.97	0.92	0.82	0.69	0.5	0.33	0.17	0.082	0.032	0.012
DPRAx2, KeratinoSensx2, h-CLATx2	1	1	0.98	0.95	0.88	0.76	0.58	0.39	0.21	0.096	0.035	0.012
LLNAx4	1	1	1	1	1	0.99	0.97	0.91	0.77	0.57	0.34	0.17
DPRAx2, kDPRAx1, KeratinoSensx2, h-CLATx2, U-Sensx1	1	1	1	0.98	0.94	0.84	0.66	0.43	0.22	0.091	0.029	0.0091
DPRAx2, KeratinoSensx2, h-CLATx2, LLNAx4	1	1	1	1	1	0.98	0.91	0.73	0.43	0.19	0.047	0.0095
DPRAx2, kDPRAx1, KeratinoSensx2, h-CLATx2, U-Sensx1, LLNAx4	1	1	1	1	1	1	0.97	0.88	0.61	0.32	0.095	0.019

*Comparison of ED01 estimates (based on different combinations of inputs) and probability that exposures are the less than the ED01. Thresholds of 0.2 (**orange - ≥ 80% likelihood that exposure is greater than ED_{01}**) and 0.8 (**blue - ≥80% likelihood that exposure is less than ED_{01}**).*

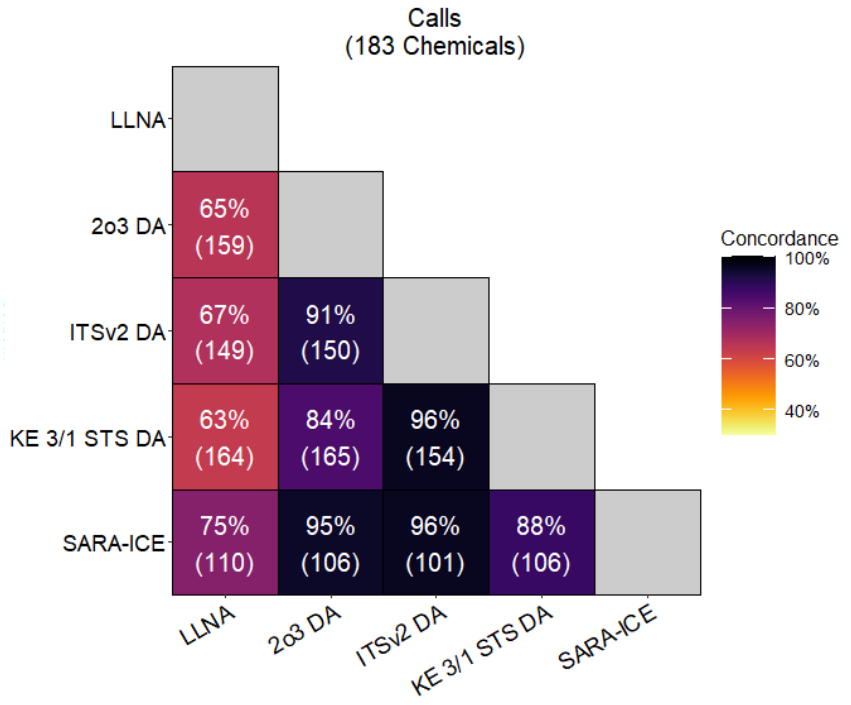
National Institute of Environmental Health Sciences
Division of Translational Toxicology

GHS estimates for MIT for different SARA-ICE data inputs

Chemical	Input combination	ED01 (μg/cm^2)	Pr(GHS 1A)	Pr(GHS 1B)	Pr(NC)	Classification GHS binary	Classification GHS subcategory	Reference GHS Category
MIT	DPRA, KeratinoSens, h-CLAT	9.9	0.94	0.06	0.00	1	1A	1A
MIT	LLNA	4.3e+02	0.53	0.46	0.01	1	Inconclusive*	
MIT	HPPT	4.8e+02	0.54	0.45	0.00	1	Inconclusive*	
MIT	DPRA, KeratinoSens, h-CLAT, LLNA	66	0.89	0.11	0.00	1	1A	
MIT	DPRA, KeratinoSens, h-CLAT, HPPT	1.5e+02	0.84	0.16	0.00	1	1A	
MIT	HPPT, LLNA	3.7e+02	0.61	0.39	0.00	1	1A	
MIT	DPRA, KeratinoSens, h-CLAT, HPPT, LLNA	1.7e+02	0.86	0.14	0.00	1	1A	
CMIT/MIT	DPRA, KeratinoSens, h-CLAT	1.9	0.98	0.02	0.00	1	1A	1A
CMIT/MIT	LLNA	6	0.98	0.02	0.00	1	1A	
CMIT/MIT	HPPT	15	1.00	0.00	0.00	1	1A	
CMIT/MIT	DPRA, KeratinoSens, h-CLAT, LLNA	2.6	1.00	0.00	0.00	1	1A	
CMIT/MIT	DPRA, KeratinoSens, h-CLAT, HPPT	9.8	1.00	0.00	0.00	1	1A	
CMIT/MIT	HPPT, LLNA	10	1.00	0.00	0.00	1	1A	
CMIT/MIT	DPRA, KeratinoSens, h-CLAT, HPPT, LLNA	8.1	1.00	0.00	0.00	1	1A	
BIT	DPRA, KeratinoSens, h-CLAT	5.3	0.96	0.04	0.00	1	1A	1
BIT	LLNA	1.2e+03	0.33	0.64	0.03	1	1B	
BIT	HPPT	63	0.83	0.16	0.00	1	1A	
BIT	DPRA, KeratinoSens, h-CLAT, LLNA	1.2e+02	0.81	0.19	0.00	1	1A	
BIT	DPRA, KeratinoSens, h-CLAT, HPPT	12	0.98	0.02	0.00	1	1A	
BIT	HPPT, LLNA	2.4e+02	0.69	0.31	0.00	1	1A	
BIT	DPRA, KeratinoSens, h-CLAT, HPPT, LLNA	73	0.93	0.07	0.00	1	1A	
OIT	DPRA, KeratinoSens, h-CLAT	0.66	0.99	0.01	0.00	1	1A	1A
OIT	LLNA	1.1e+02	0.78	0.22	0.00	1	1A	
OIT	DPRA, KeratinoSens, h-CLAT, LLNA	14	0.98	0.02	0.00	1	1A	
DCOIT	DPRA, KeratinoSens, h-CLAT	0.64	0.99	0.01	0.00	1	1A	1A
DCOIT	LLNA	4.4	0.98	0.02	0.00	1	1A	
DCOIT	DPRA, KeratinoSens, h-CLAT, LLNA	1.7	1.00	0.00	0.00	1	1A	
BBIT	DPRA, KeratinoSens, h-CLAT	3.2	0.97	0.03	0.00	1	1A	1

Federal Agency Partners Chemical Set

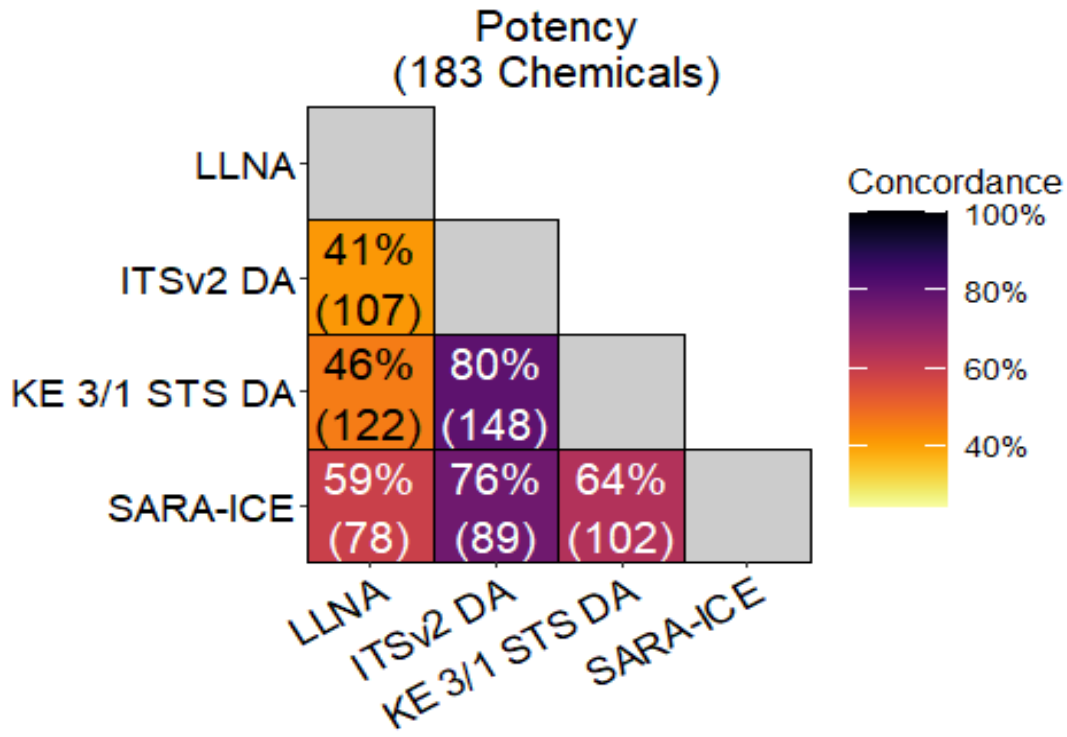
- 183 substances tested in DPRA, KeratinoSens, and h-CLAT
- Evaluated against LLNA and existing DAs with SARA-ICE Output for hazard and GHS category



Defined Approach	Sensitivity	Specificity	Balanced Accuracy	False Positive Rate	False Negative Rate	Number of Chemicals Predicted	Inconclusive (LLNA GHS 1/NC)
2o3	76.00%	47.46%	61.73%	53%	24%	159	9 (6/3)
ITSv2	84.69%	33.33%	59.01%	67%	15%	149	18 (8/10)
KE 3/1 STS	87.38%	22.95%	55.16%	77%	13%	164	0
SARA-ICE	83.33%	53.13%	68.23%	47%	17%	110	63 (33/30)

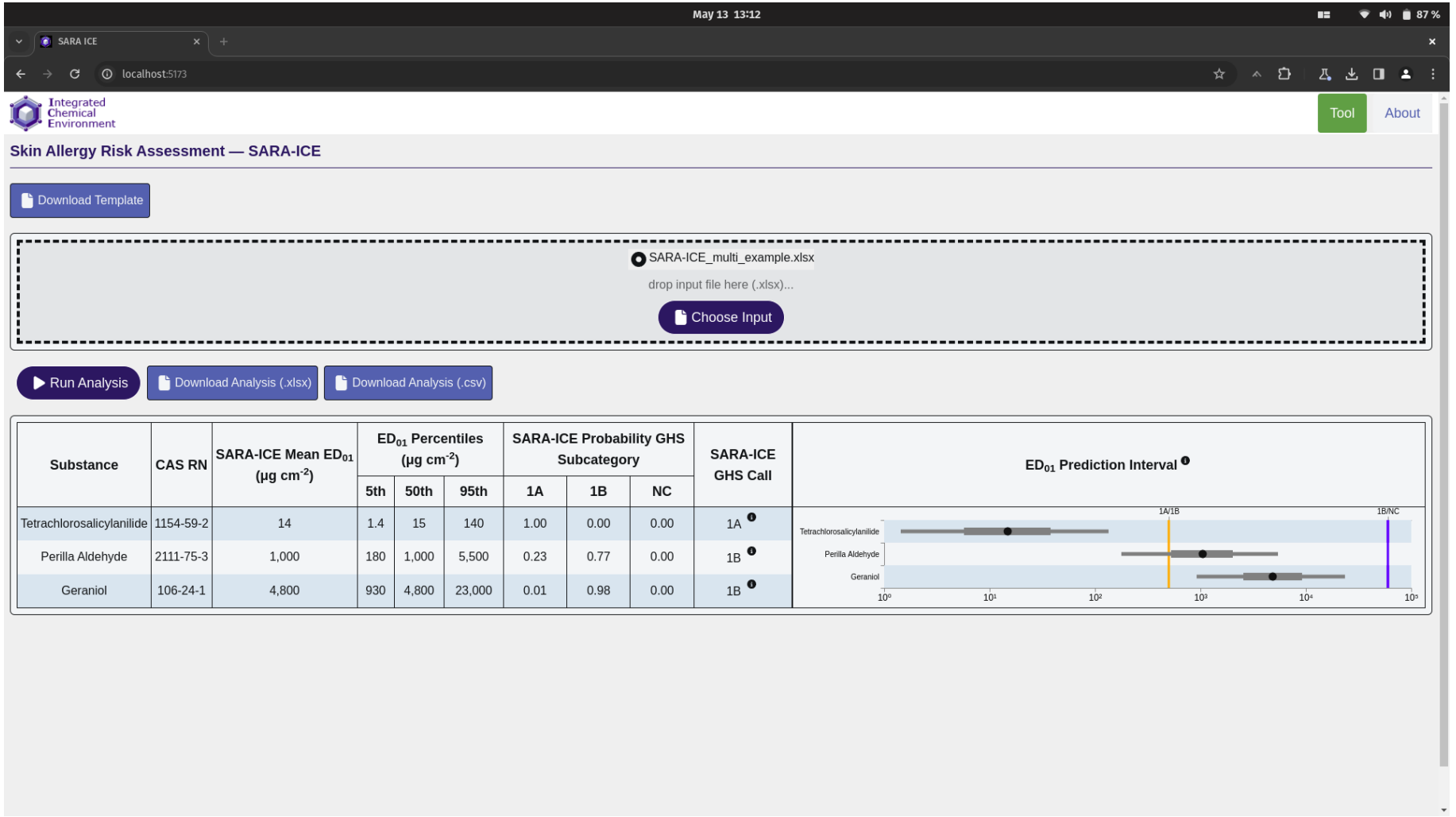
Potency Estimates

- 183 substances tested in DPRA, KeratinoSens, and h-CLAT
- Evaluated against LLNA and existing DAs with SARA-ICE Output for hazard and GHS category



Defined Approach	Accuracy	Underpredicted	Overpredicted	Number of Chemicals Predicted	Inconclusive (LLNA GHS 1A/1B/NC)
ITSv2	41%	26%	33%	102	19 (4/6/9)
KE 3/1 STS	46%	21%	33%	122	0
SARA-ICE	59%	18%	23%	78	51 (7/18/26)

SARA-ICE Container



Substance	CAS RN	SARA-ICE Mean ED_{01} (µg cm^{-2})	ED_{01} Percentiles (µg cm^{-2})			SARA-ICE Probability GHS Subcategory			SARA-ICE GHS Call	ED_{01} Prediction Interval
			5th	50th	95th	1A	1B	NC		
Tetrachlorosalicylanilide	1154-59-2	14	1.4	15	140	1.00	0.00	0.00	1A	
Perilla Aldehyde	2111-75-3	1,000	180	1,000	5,500	0.23	0.77	0.00	1B	
Geraniol	106-24-1	4,800	930	4,800	23,000	0.01	0.98	0.00	1B	

Beta testing of the container is ongoing, volunteers wanted

Conclusions

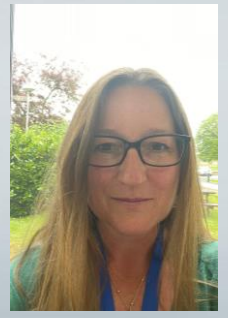
- SARA-ICE DA is being adapted for regulatory use through expanded data and functionality
- SARA-ICE DA shows good concordance with sensitizer binary and GHS sub-category classifications against OECD DASS benchmark data
- Case studies demonstrated benefits of SARA-ICE DA:
  - estimates human potency (ED_{01}) with uncertainty
  - estimates with in vitro and in vivo data inputs
  - estimates with incomplete and repeat datasets
- Evaluation of the SARA-ICE DA, including thresholds for conclusive predictions and performance impact, is ongoing within the OECD DASS expert group
- SARA-ICE is packaged for download for local implementation and **is available** for beta testing upon request via the NICEATM website (https://ntp.niehs.nih.gov/whatwestudy/niceatm) or by contacting me (Emily.Reinke@inotivco.com).

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Acknowledgments

The NICEATM Group

Unilever



Subscribe to NICEATM News email list

Integrated Chemical Environment

emily.reinke@inotivco.com

Model assumptions

HPPT

1. There is a dermal dose at which there is a 1% chance of inducing sensitisation in a randomly selected individual from a HPPT-eligible population.
2. The probability of inducing sensitisation in a HPPT increases with dose.
3. Each individual within a HPPT-eligible population has a personal threshold for sensitisation to any given chemical. This threshold may be greater than the maximum possible dose.
4. The distribution of the base-10 logarithm of personal thresholds has a Gaussian shape. The standard deviation is chemical-specific; different chemicals have different variabilities within the human population with respect to sensitivity to induction of sensitisation.
5. The number of individuals sensitised in a HPPT study follows a logit-normal-binomial compound distribution.

Model assumptions

Non-HPPT data

1. Data from the LLNA, DPRA, kDPRA, KeratinoSens, h-CLAT and U-Sens assays can be transformed to such that it is reasonable to variability in chemical-specific data in terms of a normal distribution (transformations mostly involve logarithms).

2. The same transformations put data on a scale in which it is reasonable to assume linear relationships between the average transformed datapoint on the base-10 logarithm of the ED_{01}.

3. The relationships between the average results can be described by a multivariate Gaussian distribution.

4. Variability in each test is chemical-specific. There is a latent variable for each test and each chemical which defines the variance of the chemical in the particular test.

5. Chemical-specific variance parameters can be estimated using partial pooling. The population of variances for each tested can be learnt and used to regularise chemical-specific estimates when limited data is available.

Computation

The SARA-ICE model is a mathematical model – it's assumptions and equations are expressible with pen and paper.

Learning model parameters requires numerical computation: the model is realised numerically using the programming language Stan. Python is used to process model inputs and outputs.

Computation requires many CPU cycles; however, a production version of the model has been developed to alleviate this limitation.

A standalone, downloadable version of the model is being created by NICEATM.

SARA-ICE.stan

Check on Save

```
178
179 }
180 model {
181
182   // In vitro regression
183   L_Omega ~ lkj_corr_cholesky(2);
184   for (i in 1:9) {
185     residual_scale[i] ~ normal(0, 2);
186   }
187   for (i in 1:N_chem) {
188     residual_raw[i] ~ normal(0, 1);
189   }
190
191   // Chemical distribution
192   theta_raw ~ normal(0, 1);
193
194   // HPPT
195   s_raw ~ weibull(5, 1);
196   s_scale ~ normal(0, 1);
197   logit_prob_scale_raw ~ weibull(5, 1);
198   logit_prob_scale_scale ~ normal(0, 2);
199   logit_prob_raw ~ normal(0, 1);
200
201   ////////////////////////////////////////////////////////////////////////////////////////////
202
203   // Likelihood of HPPT data given dose response curve parameters
204   {
205     vector[N_HPPT_obs] logit_prob = g((HPPT_dose - theta[HPPT_idx]) ./ s[HPPT_idx] - g_inv_log99)
206     + logit_prob_scale_scale * logit_prob_scale_raw[HPPT_idx] .* logit_prob_raw;
207     HPPT_k ~ binomial_logit(HPPT_n, logit_prob);
208   }
```