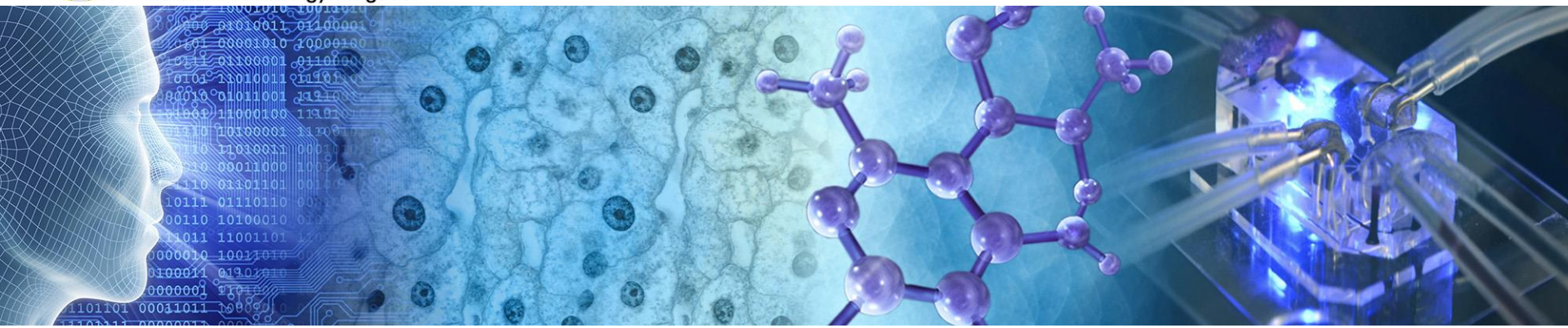
NTP
National Toxicology Program



OECD Defined Approaches for Skin Sensitization

Judy Strickland, PhD, DABT
Inotiv, Inc., Contractor Supporting the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

Society for Chemical Hazard Communication
September 22, 2022

NICEATM

NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), organized as an office under the Division of the NTP, part of NIEHS

Inotiv-RTP provides technical support for NICEATM under an NIEHS contract

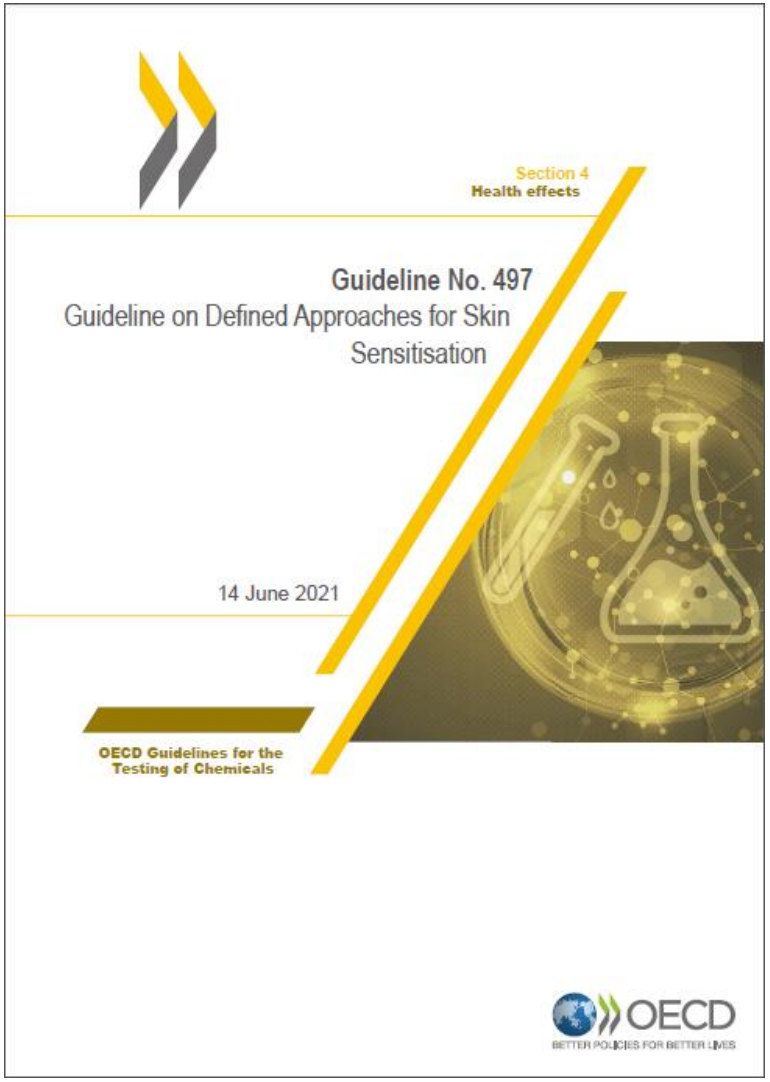
- **ICCVAM support**
- **Tox 21 validation support**
- **International harmonization efforts**

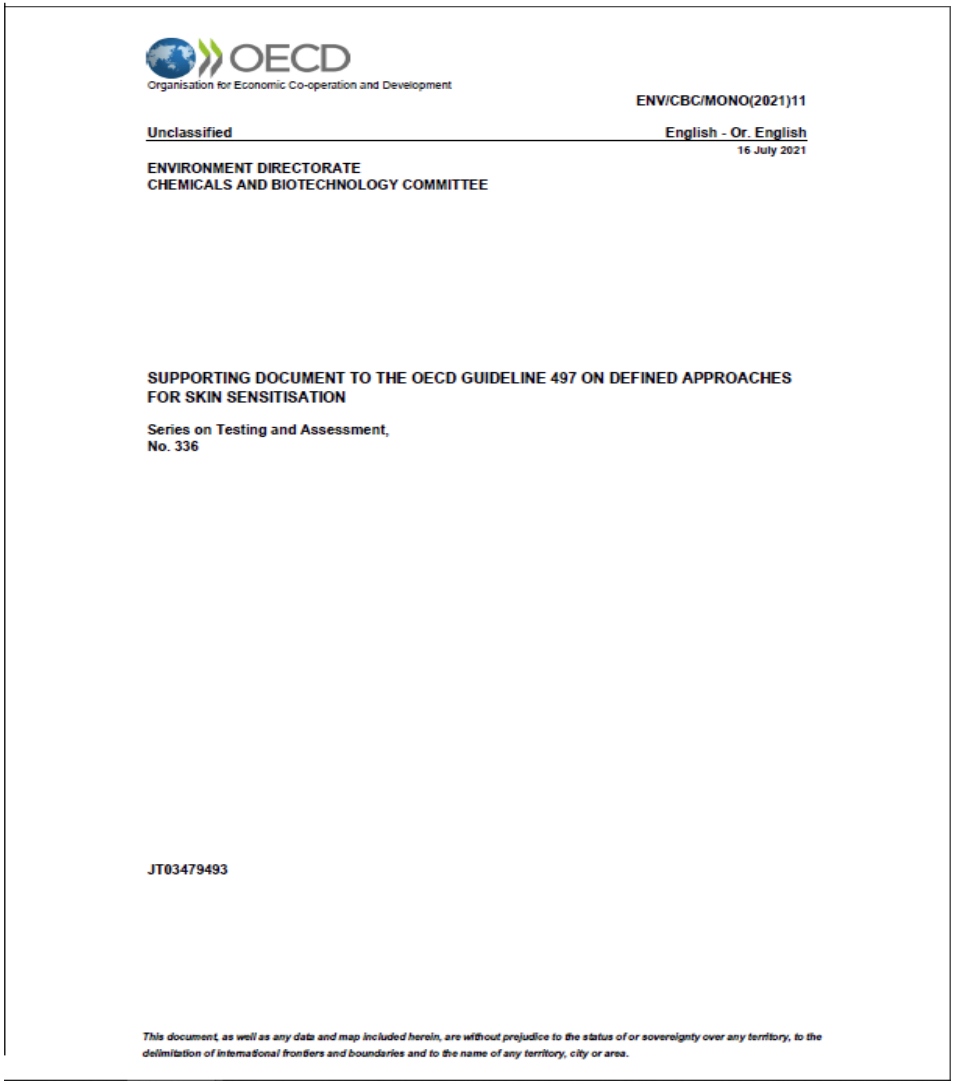
Disclaimer: Inotiv-RTP staff provide technical support for NICEATM, but do not represent NIEHS, NTP, or the official positions of any federal agency

Outline

- Overview of the guideline and documentation
- Development of the guideline
- Performance of the defined approaches

OECD Guideline 497





OECD
Organisation for Economic Co-operation and Development

ENV/CBC/MONO(2021)11

Unclassified — English - Or. English
16 July 2021

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

SUPPORTING DOCUMENT TO THE OECD GUIDELINE 497 ON DEFINED APPROACHES FOR SKIN SENSITISATION

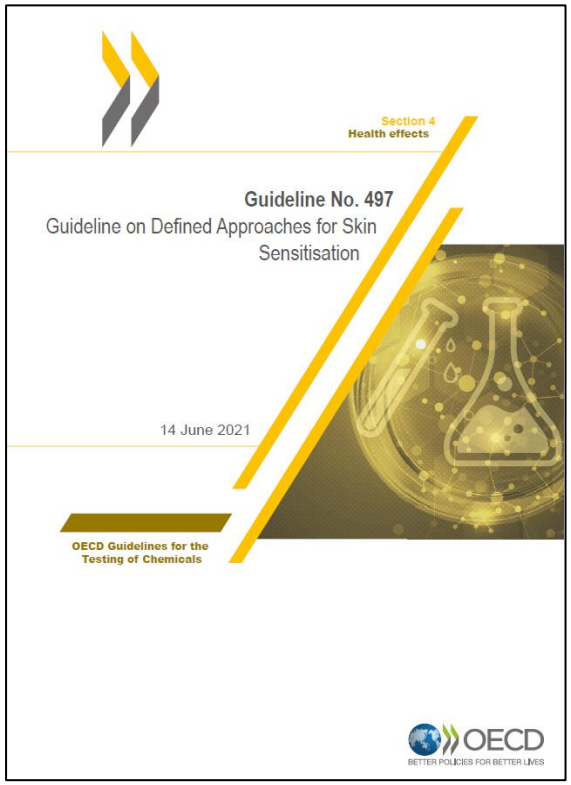
Series on Testing and Assessment,
No. 336

JT03479493

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

https://www.oecd.org/chemicalsafety/guideline-no-497-defined-approaches-on-skin-sensitisation-b92879a4-en.htm
https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/CBC/MONO(2021)11&docLanguage=En

Guideline Structure



Section 1
- General introduction
- DAs and use scenarios
- Limitations

Section 2 – DAs for hazard identification
- “2 out of 3” Defined Approach

Section 3 – DAs for potency categorisation
- Integrated Testing Strategy (ITS) v1
- ITS v2

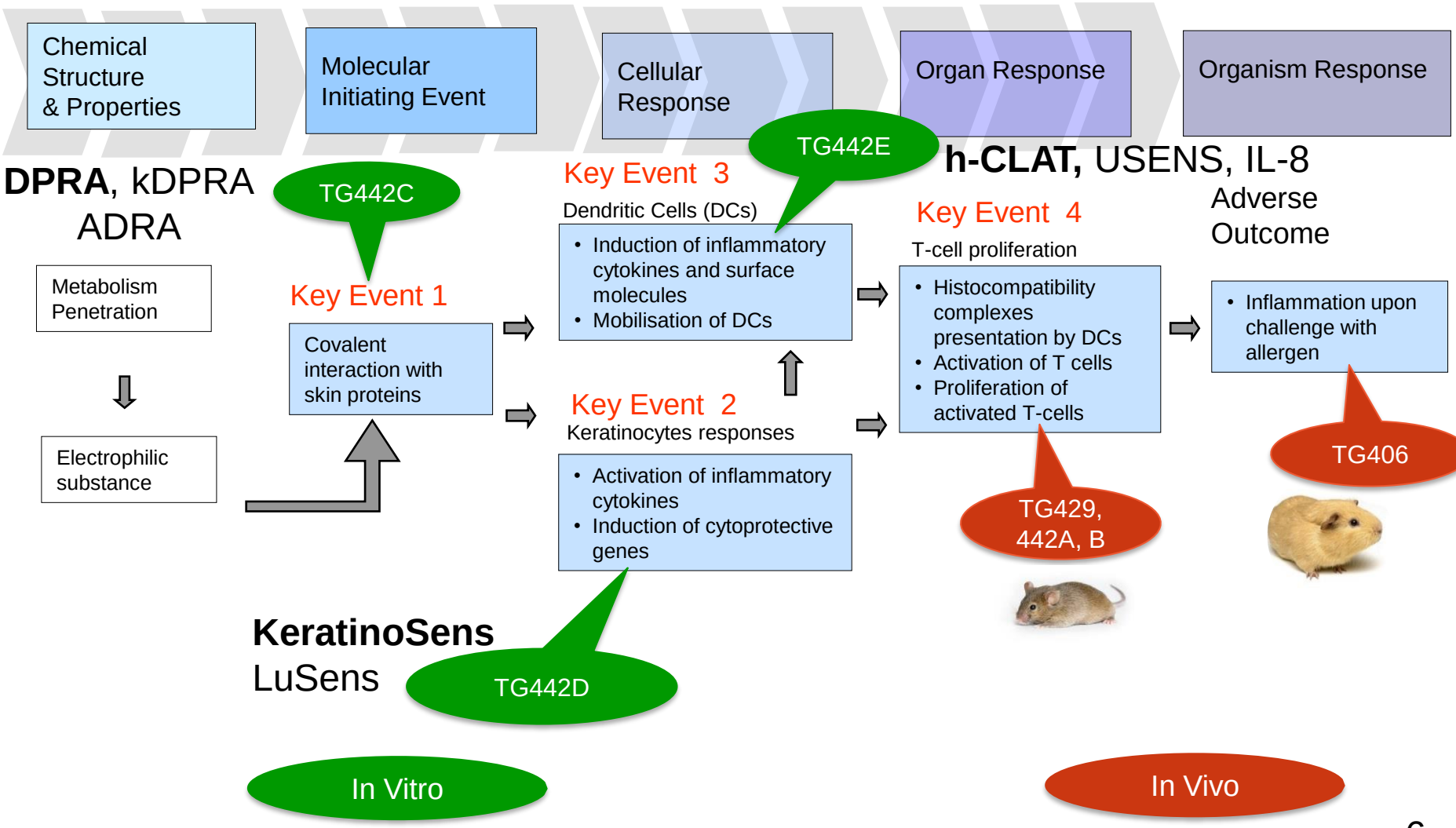
Annexes provide additional information on in silico protocols and assessing confidence in the DAs

Supporting document provides detailed information on the curation of in vivo reference classifications, predictive performance and uncertainty in the DAs and their individual data information sources

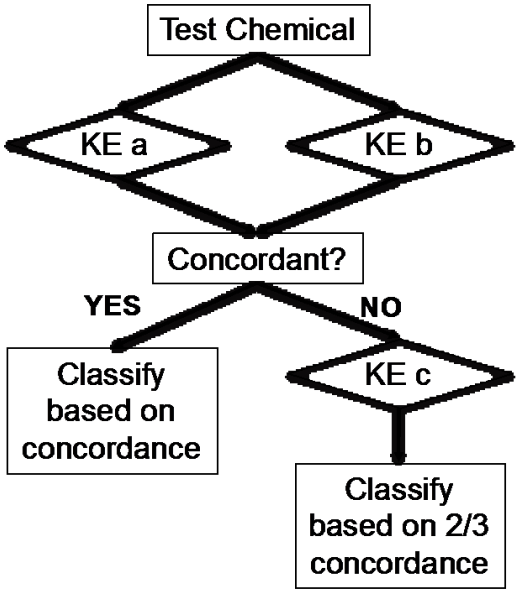
https://www.oecd.org/chemicalsafety/guideline-no-497-defined-approaches-on-skin-sensitisation-b92879a4-en.htm
https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/CBC/MONO(2021)11&docLanguage=En

Test Methods Mapped to AOP

OECD (2014)



Defined Approaches Included in GL 497



2 out of 3

Score	h-CLAT MIT (µg/mL)	DPRA depletion (%)	In silico
3	≤10	≥42.47	
2	>10, ≤150	≥22.62, <42.47	
1	>150, ≤15000	≥6.38, <22.62	Positive
0	Negative	<6.38	Negative
Potency: Total battery score		**GHS 1A**	**6-7**
		GHS 1B	**2-5**
		Not classified	**0-1**

Integrated Testing Strategy (ITS) v1 / v2

Hazard classification (S/NS)

- No differential weighting of individual test methods, or defined sequential order of testing
- Usually KE1 (DPRA) and KE2 (KeratinoSens) performed first since less expensive
- Third test is KE3 (h-CLAT)
- No potency information

Hazard classification + 3 potency classes: NS, GHS 1A, GHS 1B

- Score-based system
- Uses h-CLAT, DPRA, and in silico results
- In silico input for v1 is Derek and for v2 is QSAR Toolbox

Basis of the Guideline

CRITICAL REVIEWS IN TOXICOLOGY, 2018
VOL. 48, NO. 5, 344–358
https://doi.org/10.1080/10408444.2018.1429385

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database*

Sebastian Hoffmann[a], Nicole Kleinstreuer[b], Nathalie Alépée[c], David Allen[d], Anne Marie Api[e], Takao Ashikaga[f]†, Elodie Clouet[g], Magalie Cluzel[h], Bertrand Desprez[i], Nichola Gellatly[j]‡, Carsten Goebel[k], Petra S. Kern[l], Martina Klaric[i], Jochen Kühnl[m], Jon F. Lalko[e]§, Silvia Martinozzi-Teissier[c], Karsten Mewes[n], Masaaki Miyazawa[o], Rahul Parakhia[e], Erwin van Vliet[p], Qingda Zang[d] and Dirk Petersohn[n]

CRITICAL REVIEWS IN TOXICOLOGY, 2018
https://doi.org/10.1080/10408444.2018.1429386

Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updat

Non-animal methods to predict skin sensitization (II): an assessment of defined approaches**

Nicole C. Kleinstreuer[a], Sebastian Hoffmann[b], Nathalie Alépée[c], David Allen[d], Takao Ashikaga[e]*, Warren Casey[a], Elodie Clouet[f], Magalie Cluzel[g], Bertrand Desprez[h], Nichola Gellatly[i], Carsten Göbel[j], Petra S. Kern[k], Martina Klaric[h], Jochen Kühnl[l], Silvia Martinozzi-Teissier[c], Karsten Mewes[m], Masaaki Miyazawa[n], Judy Strickland[d], Erwin van Vliet[o], Qingda Zang[d] and Dirk Petersohn[m]

Non-animal Approach Evaluation

Most non-animal testing strategies evaluated so far perform **better** than the LLNA at predicting human skin sensitization hazard and potency

(And when compared to the LLNA, are equivalent in performance to the LLNA at predicting itself.)

Hoffmann et al. 2018 Crit Rev Tox
Kleinstreuer et al. 2018 Crit Rev Tox

OECD Defined Approaches SS Guideline Project

- 2017 OECD work plan
  - Lead by US, EC, and Canada, developed with input of the OECD EG on DASS (industry, regulatory agencies, validation bodies, NGOs, industry)
  - Aims to provide a **substitute** for animal testing for skin sensitization based on a combination of methods which, individually, predict key event responses on the AOP
  - Aims for an international guideline covered by the agreement on Mutual Acceptance of Data (MAD)
- To meet regulatory requirements, need:
  - DAs that discriminate skin sensitisers from non-sensitizers
  - DAs that discriminate strong from moderate/weak sensitizers (GHS potency categories)

Expert Group on DA SS

- 68 members covering regulatory authorities, OECD national coordinators, validation experts, animal welfare and industry stakeholders, method developers, etc.
- Focused on resolving scientific issues:
  1. **Curation of reference data**
  2. **DAs to include in the guideline**
  3. **Performance evaluation**
  4. **Applicability domain**
  5. **Confidence and uncertainty**
- National coordinators had special meetings to discuss
- Draft guideline distributed for public comment in Sep 2019 and Dec 2020; final published in Jun 2021

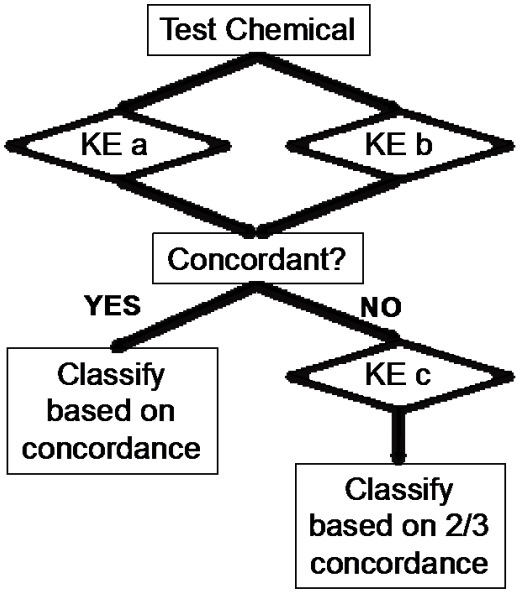
Data

- Reference data
  - LLNA was the primary reference data, but human data from predictive patch tests were also used
  - 168 chemicals have LLNA and 66 have human predictive patch test results
  - Mostly cosmetic ingredients but also other types of chemicals
  - Range of physicochemical characteristics
- Input data to DAs
  - DAs use specific validated methods only: DPRA, KeratinoSens, and h-CLAT
  - ITS DAs use in silico info source: DEREK or QSAR Toolbox
- Performance
  - Hazard (binary) and GHS potency categories (3 classes)

Guideline 497 includes:

- An approach for describing the applicability domain of the DAs, including in vitro results combined with in silico predictions
- Decision trees for each DASS to include the uncertainty in the data information sources and confidence in the overall prediction
- An approach for standardizing in silico predictions to assure reliable and reproducible results
- Details necessary on in silico models and predictions to include in a test report used for regulatory decision-making
- Recognition of areas where additional research may help to elucidate chemistries where the human response is not well predicted by the animal model

Defined Approaches Included in GL 497



2 out of 3

Score	h-CLAT MIT (µg/mL)	DPRA depletion (%)	In silico
3	≤10	≥42.47	
2	>10, ≤150	≥22.62, <42.47	
1	>150, ≤15000	≥6.38, <22.62	Positive
0	Negative	<6.38	Negative
Potency: Total battery score		**GHS 1A**	**6-7**
		GHS 1B	**2-5**
		Not classified	**0-1**

Integrated Testing Strategy (ITS) v1 / v2

Hazard classification (S/NS)

- No differential weighting of individual test methods, or defined sequential order of testing
- Usually KE1 (DPRA) and KE2 (KeratinoSens) performed first since less expensive
- Third test is KE3 (h-CLAT)
- No potency information

Hazard classification + 3 potency classes: NS, GHS 1A, GHS 1B

- Score-based system
- Uses h-CLAT, DPRA, and in silico results
- In silico input for v1 is Derek and for v2 is QSAR Toolbox

NTP National Toxicology Program

Performance of DAs

DA/Method	Information Sources	Capability (Hazard and/or Potency)	Hazard Performance vs. LLNA N~168	Hazard Performance vs. Human N~63	GHS Potency Performance vs. LLNA (Accuracy)	GHS Potency Performance vs. Human (Accuracy)
2o3 DA	DPRA, KeratinoSens, h-CLAT	Hazard	84% BA, 82% Sens, 85% Spec	88% BA, 89% Sens, 88% Spec	-	-
ITSv1 DA	DPRA, h-CLAT, DEREK Nexus v6.1.0	Hazard, Potency (GHS)	81% BA, 92% Sens, 70% Spec	69% BA, 93% Sens, 44% Spec	70% NC, 71% 1B, 74% 1A	44% NC, 77% 1B, 65% 1A
ITSv2 DA	DPRA, h-CLAT, OECD QSAR Toolbox v4.5	Hazard, Potency (GHS)	80% BA, 93% Sens, 67% Spec	69% BA, 94% Sens, 44% Spec	67% NC, 72% 1B, 72% 1A	44% NC, 80% 1B, 67% 1A
LLNA (provided for comparison)	*in vivo*	Hazard, Potency	-	58% BA, 94% Sens, 22% Spec	-	25% NC, 74% 1B, 56% 1A

*BA = Balanced Accuracy, average of Sensitivity and Specificity

Considerations for Use

- Consider all information known about a chemical before testing to determine whether guideline is applicable
- Substances must be within applicability domain of individual methods
  - Review the limitations of DPRA, KeratinoSens, and h-CLAT methods (metals, mixtures, log P > 3.5 for h-CLAT, etc.)
  - High confidence vs. low confidence results based on individual methods
- Low confidence results will produce an inconclusive DA prediction, but it may be useable with other supporting information

Conclusion

- OECD guideline 497 is the first guideline of its kind
  - A standardized procedure to integrate data from multiple non-animal methods
  - Amenable to Mutual Acceptance of Data agreement
  - Intended to replace the use of animals for skin sensitization assessments; provides information equivalent to the LLNA (i.e., hazard and potency classification)
- Now underway, recently added to OECD workplan:
  - Evaluate DAs with these same rule-based structures but substitute other methods that align with the specified key events of the AOP (US leads)
  - Evaluate feasibility of adding a method that addresses regulatory needs for quantitative risk assessment, the Skin Allergy Risk Assessment model (US and UK lead)

Partners

OECD

- Patience Browne
- Anne Gourmelon
- EG DA SS (~70 members!)

Cosmetics Europe

- Sebastian Hoffmann
- Many industry partners

EURL ECVAM

- Silvia Casati
- David Asturiol

Health Canada

- Michele Regimbald-Krnel
- Cameron Bowes
- Pierre Therriault

US EPA

- Anna Lowit
- Tim McMahon
- OPP Staff

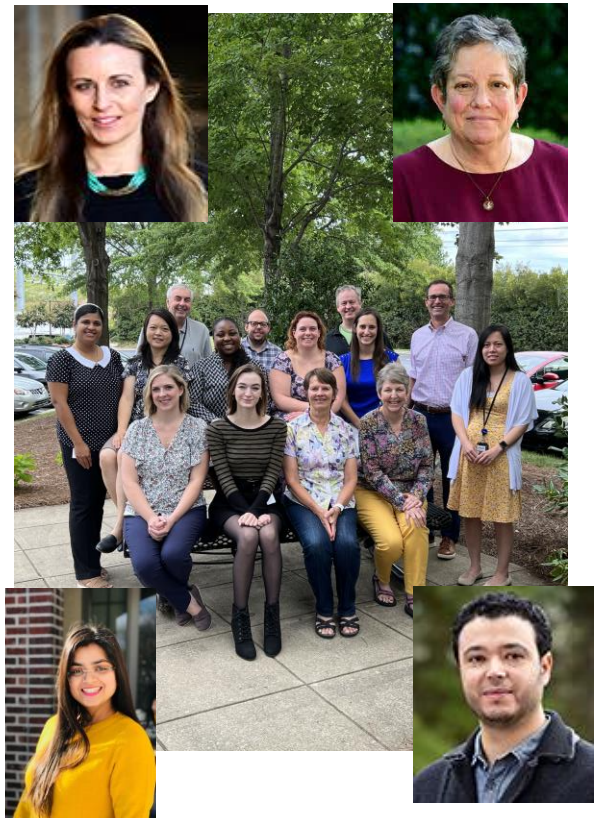
DNTP

- Dori Germolec
- Warren Casey

ICCVAM

- Skin Sensitization EG

And the NICEATM group….



Extra Slides

2 out of 3 DA for Hazard Classification

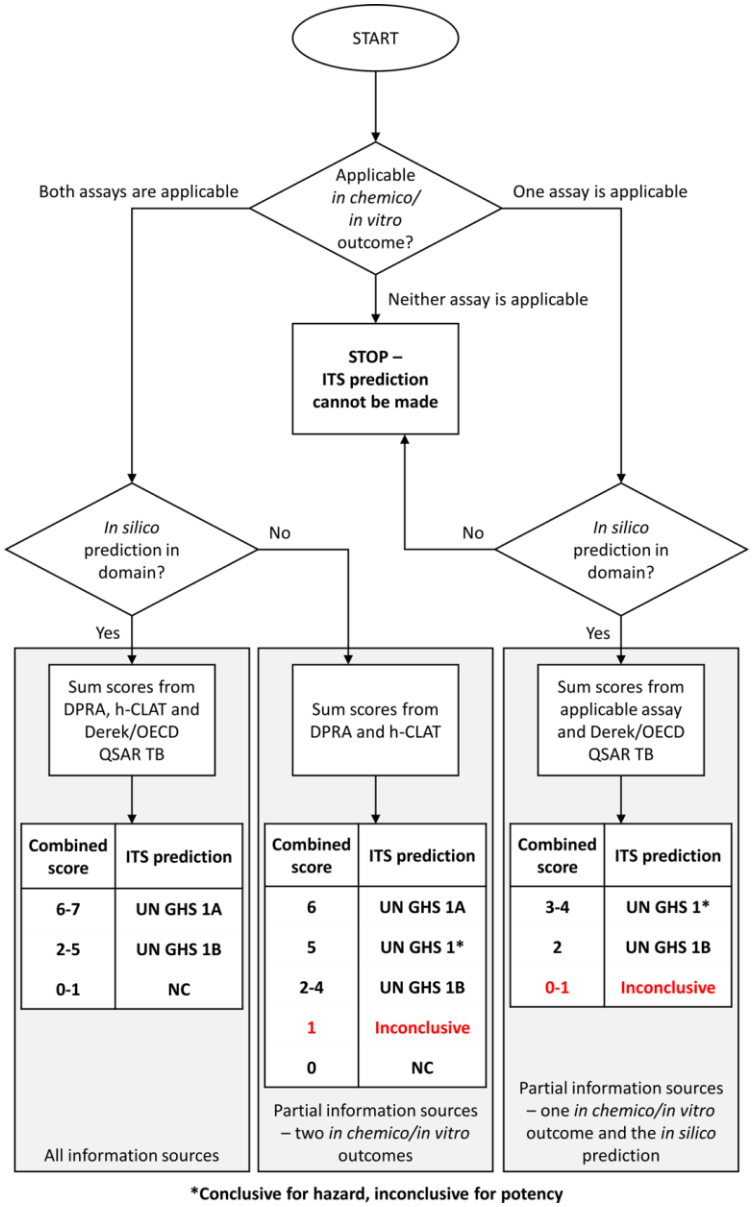
- Borderline ranges
  - DPRA, KeratinoSens, and h-CLAT
  - High confidence prediction can be made only if 2 of 3 concordant results are outside the borderline ranges
  - If one of the two concordant results is in the borderline range, the DA prediction is inconclusive
    - Depending on context or regulatory authority, borderline positive results may be used
    - DA is also inconclusive if one of the two concordant results is a negative h-CLAT for a substance with log P > 3.5
- Inconclusive DA results can be used in a weight-of-evidence approach with other information sources

DAs for Potency Classification: ITSv1 and v2

- Uses h-CLAT, DPRA, and in silico
  - ITSv1 uses Derek alerts and ITSv2 uses QSAR Toolbox hazard predictions
- Derek Nexus v6.1.0, from LHASA, Ltd., is an expert knowledge-based software tool that has structural alerts for skin sensitization, which have likelihoods
  - Positive: certain, probable, plausible, and equivocal
  - Negative: doubted, improbable, impossible, non-sensitizer
    - "Contains misclassified and/or unclassified features" means it's out of domain
- OECD QSAR Toolbox v4.5 has an automated workflow that uses read-across or protein-binding alerts to make hazard predictions –
  - Provides an "in domain" or "out of domain" notation
- Both in silico tools consider metabolites and auto-oxidation products

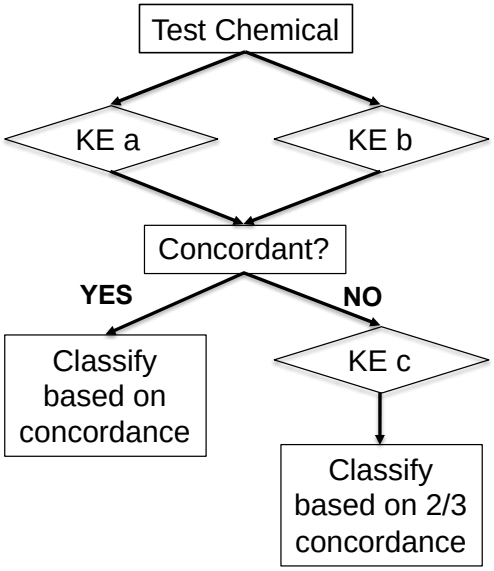
Use of Partial Information for ITS

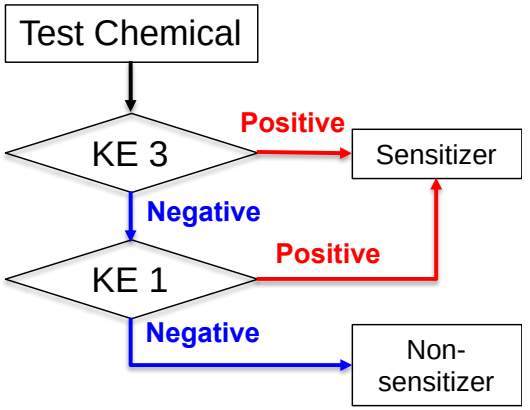
- Left – all information sources and in silico is in domain
- Middle top – neither DPRA and h-CLAT are applicable: no prediction
- Middle bottom – in silico is out of domain; only DPRA and h-CLAT available
- Right – either DPRA or h-CLAT is applicable and in silico prediction is in domain



US EPA Regulatory Progress

- Release of *Draft Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing* (10 April 2018)
  - Joint policy between Office of Pesticide Programs (OPP) and Office of Pollution Prevention and Toxics (OPPT)
  - Applies to pesticide active ingredients, inerts, and single chemicals regulated under amended TSCA
  - Two DAs currently accepted:
    "AOP 2 out of 3" and "KE 3/1 STS"





https://www.epa.gov/pesticides/epa-releases-draft-policy-reduce-animal-testing-skin-sensitization