Givaudan

ENGAGING
THE
SENSES

Read across, in silico, in vitro and human testing for skin sensitizer identification and potency determination: current state of knowledge and thought starter

Graham Ellis

IDEA Workshop : Validity of the QRA Methodology & Possibilities of Further Refinement

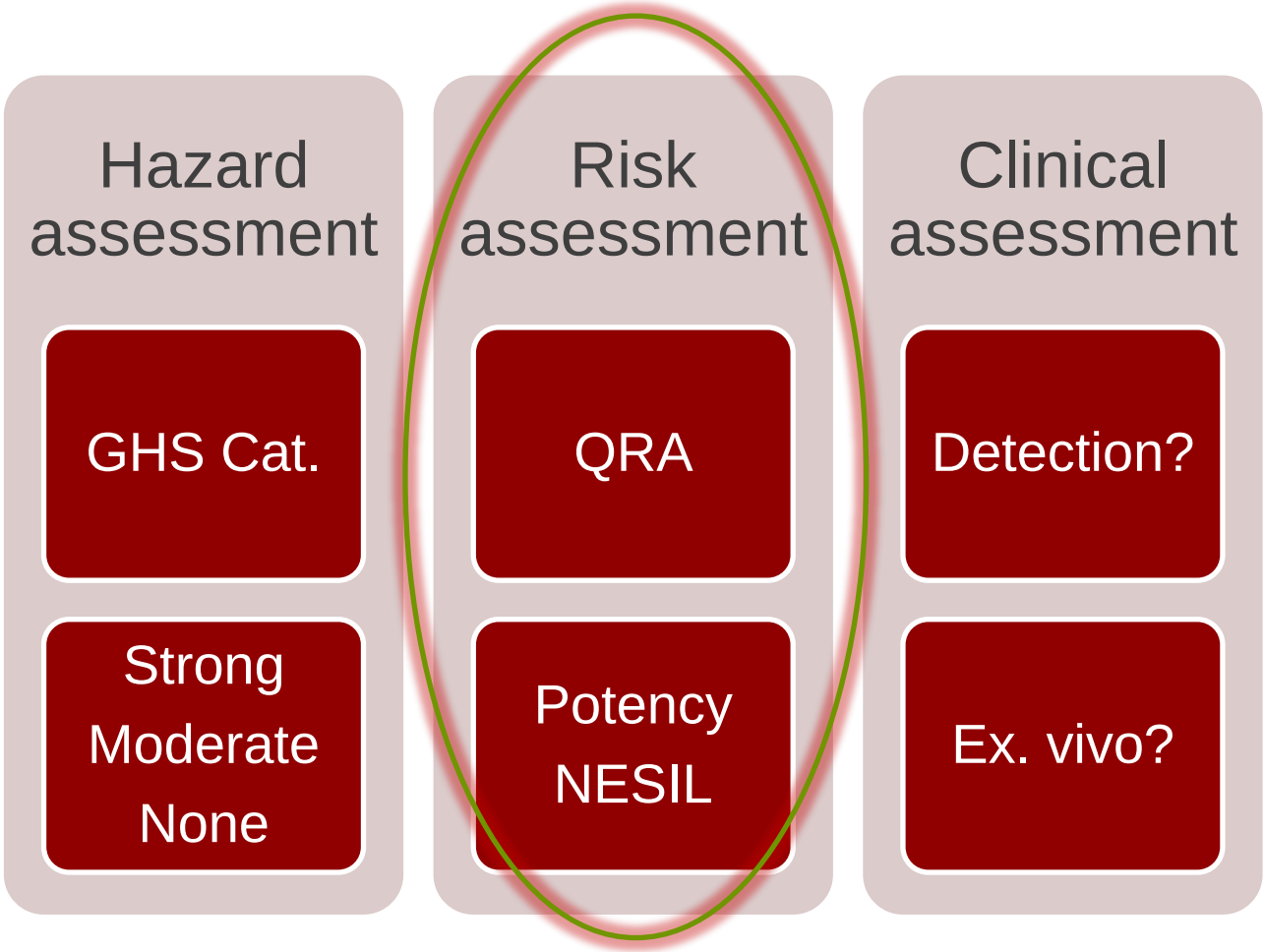
May 13-15th, 2014

Presentation content

- Overall considerations
- Some general considerations on Integrated Testing Strategies (ITS) for skin sensitization
- Key results from the ITS-2 project led by P&G, integrating multiple data sources
- Key results from multiple regression study on 244 chemicals within Givaudan
- Development of “Local” models to improve predictivity
- What is still required for future development for sensitisation prediction use in QRA

Using non-animal data for sensitisation

Which target to discuss?



Givaudan

How to assess sensitisation potency for QRA?

- Traditionally via NESIL or Potency class
- NESIL: Point value based on LLNA +other data in WoE and confirmatory HRIPT
- Potency class: «Deafult» values based on LLNA Classes
  - None – Weak – Moderate – Strong/Extreme

- Animal data (LLNA) will continue to be generated for chemical regulation and assessment purposes around the world (e.g. REACH)

- One big question is: How to understand and manage uncertainty in (the absence of) animal studies?

Givaudan

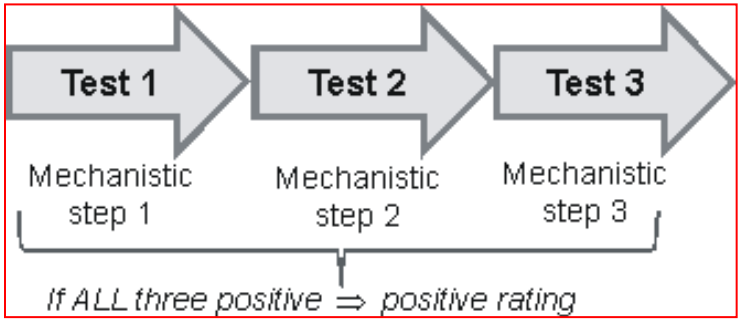
Recognise the limitations of non animal studies

- Regulatory approval ongoing
  - Peptide binding assay and Keratinosens approved by ECVAM
  - Draft OECD guidelines available – final in 2015
  - h-CLAT (human Cell Line Activation Test) to follow
  - Other methods also under development
- Any further alternative methods will take time to be fully developed and approved
  - Keratinosens and Peptide binding took ca7 years
- Limitations in metabolism in in vitro studies to date
- Complex mixtures e.g .essential oils are not validated for in vitro methods (nor LLNA)
  - Understanding of components will remain important

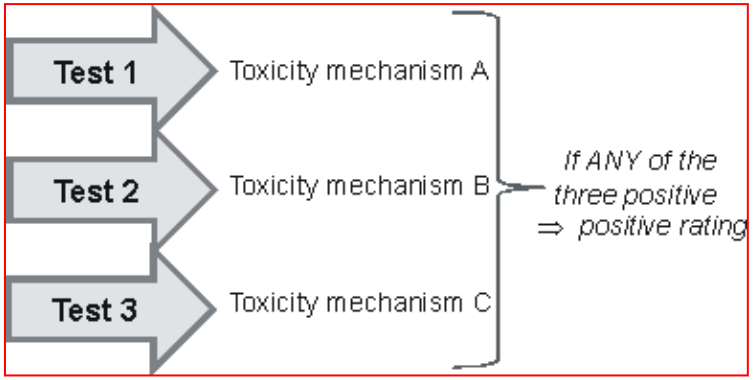
Givaudan

Part I: general considerations on Integrated Testing Strategies (ITS) for skin sensitization

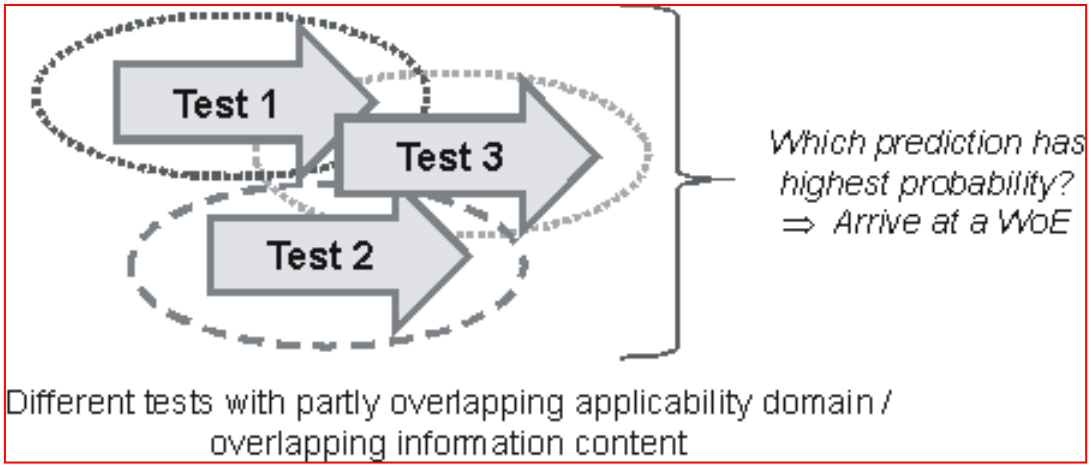
Data integration – Parallel, Serial or weight of evidence?



Serial



Parallel



Weight of evidence

Paradigms proposed for Skin sensitization ITS based on the 'serial chain of events'

OECD adverse outcome pathway for skin sensitization:

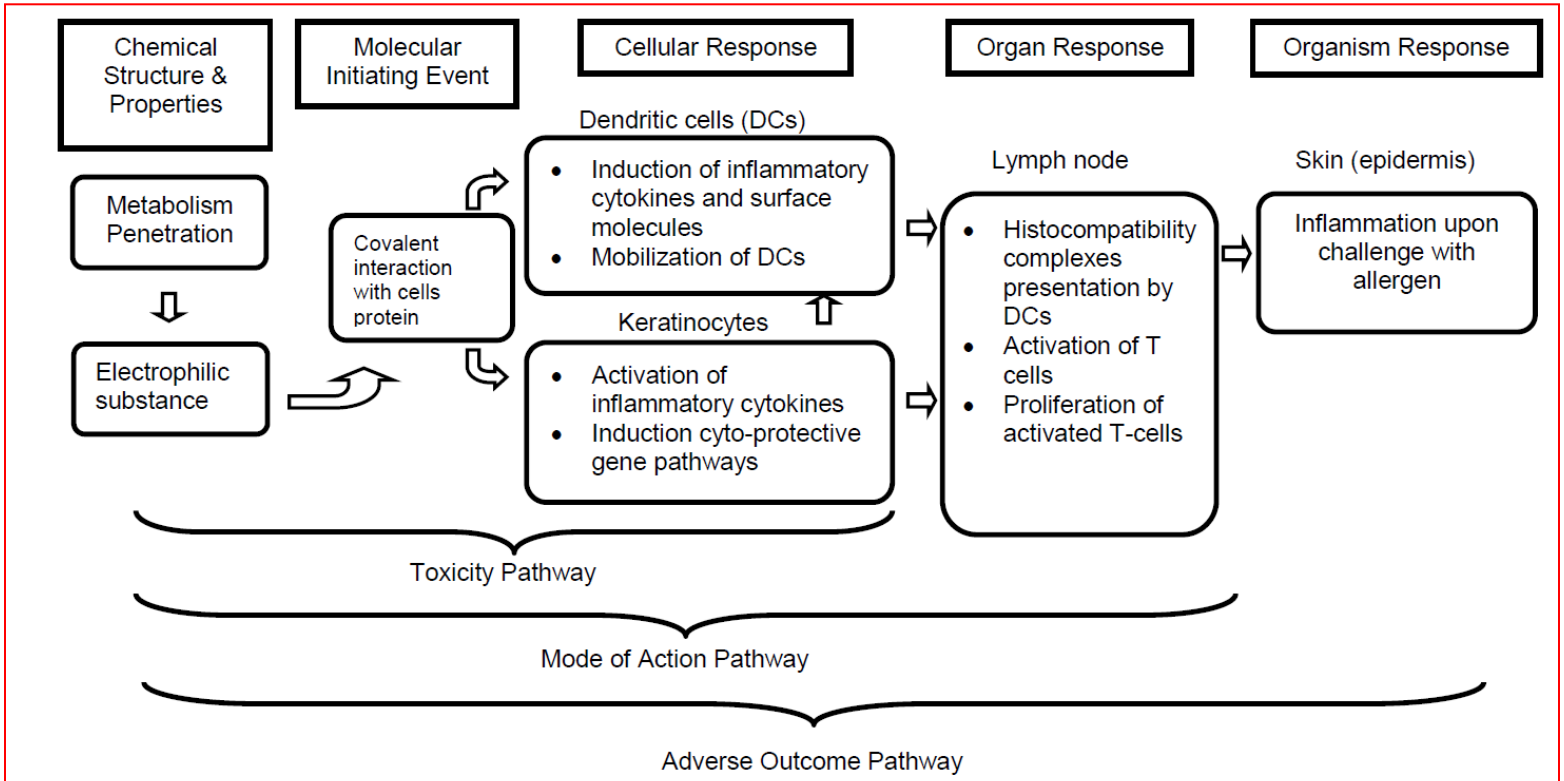
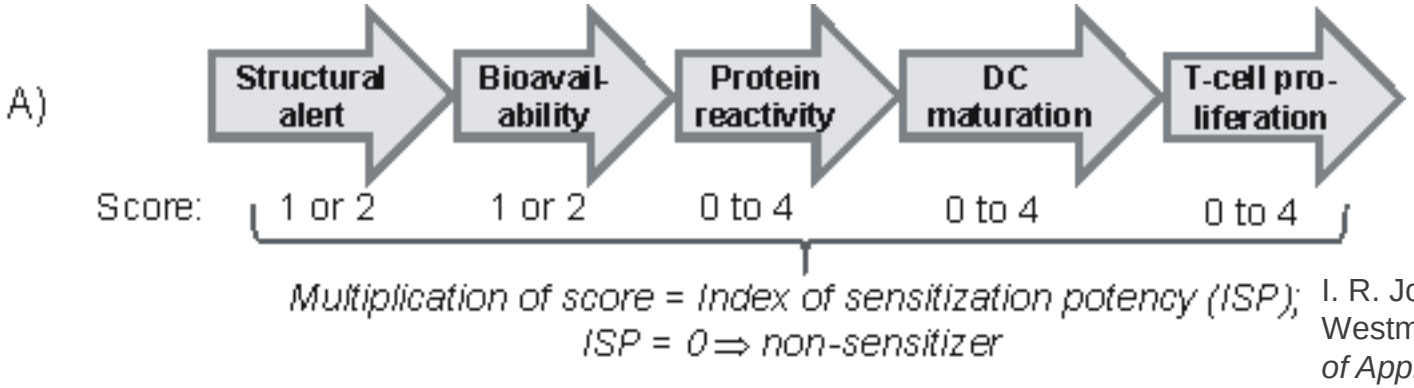


Figure 3. **Flow diagram of the pathways associated with skin sensitisation.**

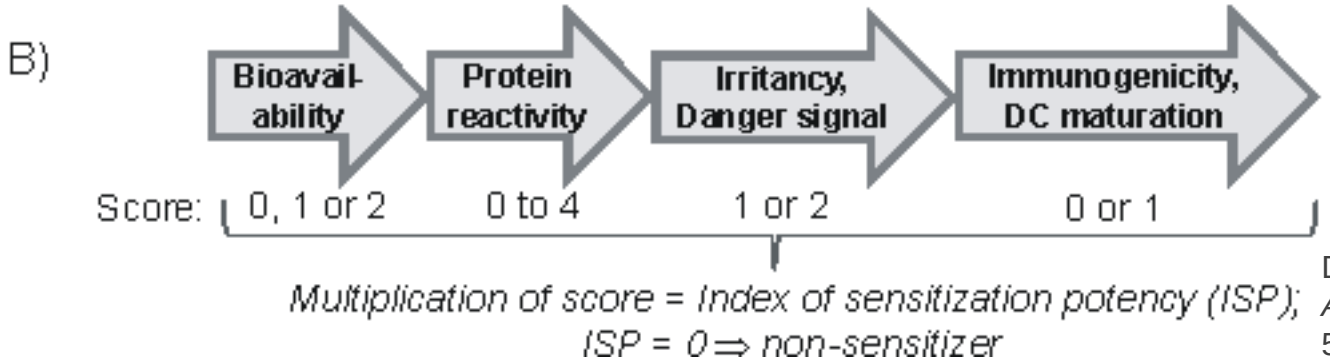
Givaudan

Paradigms proposed for Skin sensitization ITS based on the 'serial chain of events'

- ITS proposed by Jowsey et al. and Basketter & Kimber: Integration based on serial events – molecule must negotiate all hurdles



I. R. Jowsey, D. A. Basketter, C. Westmoreland, I. Kimber, *Journal of Applied Toxicology,* 2006**, 26**, 341-350.



D. A. Basketter, I. Kimber, *Journal of Applied Toxicology,* 2009**, 29**, 545-550.

Part II: ITS 2 – Bayesian integrated testing strategy to assess skin sensitization potency: From theory to practice

Project led by P&G, Givaudan contributed data

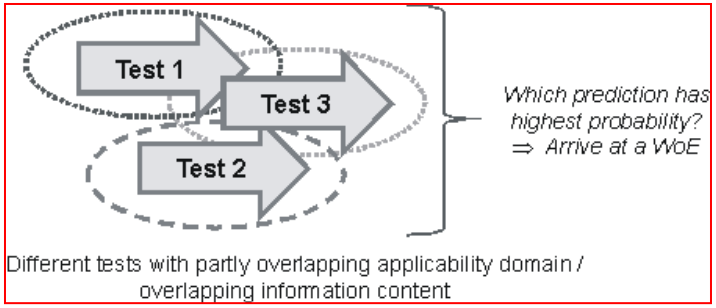
- **Goal: Predict four LLNA classes (NS / weak / Moderate / strong –extreme)**
  - Success criteria (*inter alia*): able to predict better than individual tests on an external test set
  - Predict equally well on the training and on external tests set

Jaworska, J., Y. Dancik, et al. (2013). "Bayesian integrated testing strategy to assess skin sensitization potency: From theory to practice." Journal of Applied Toxicology **33**(11): 1353-1364.

Natsch, A., C. A. Ryan, et al. (2013). "A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation." Journal of Applied Toxicology **33**(11): 1337-1352.

ITS-2: The probabilistic approach to Weight of evidence

- Flow diagram/ serial paradigms / deterministic models more difficult for potency classifications
- Can lead to conflicts in data

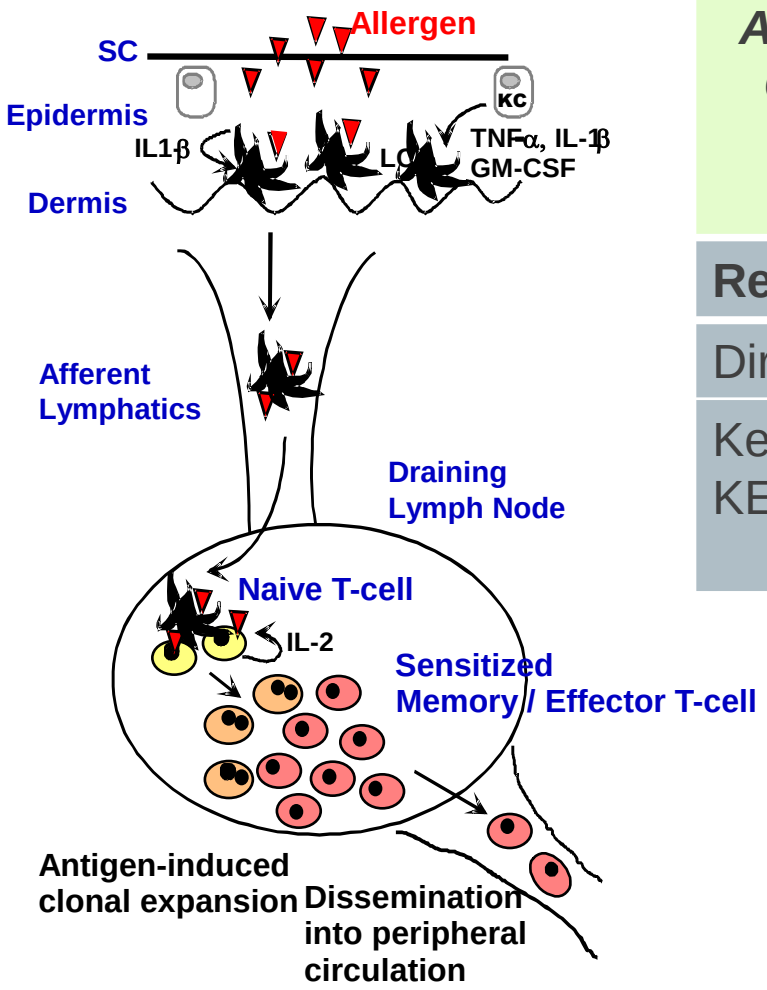


- **For potency predictions different, partly overlapping tests each inform on the target**

- Results of the different tests change probability distribution of the target variable (e.g. LLNA potency)
  - A Bayesian network may calculate probability for different LLNA states based on all available evidence
  - No binary decision: Chemical is attributed to group with highest probability
  - Probability distribution informs on the quality of the prediction
  - Can handle partial evidence

Database available

Sensitization Phase



Log K_{ow}, *AUC*120, $C_{tot,free}$	**Bioavailability data:** lipophilicity and kinetic parameters in epidermis from a simulation of exposure in a LLNA test developed by prof. Kasting (U. Cincinnati)

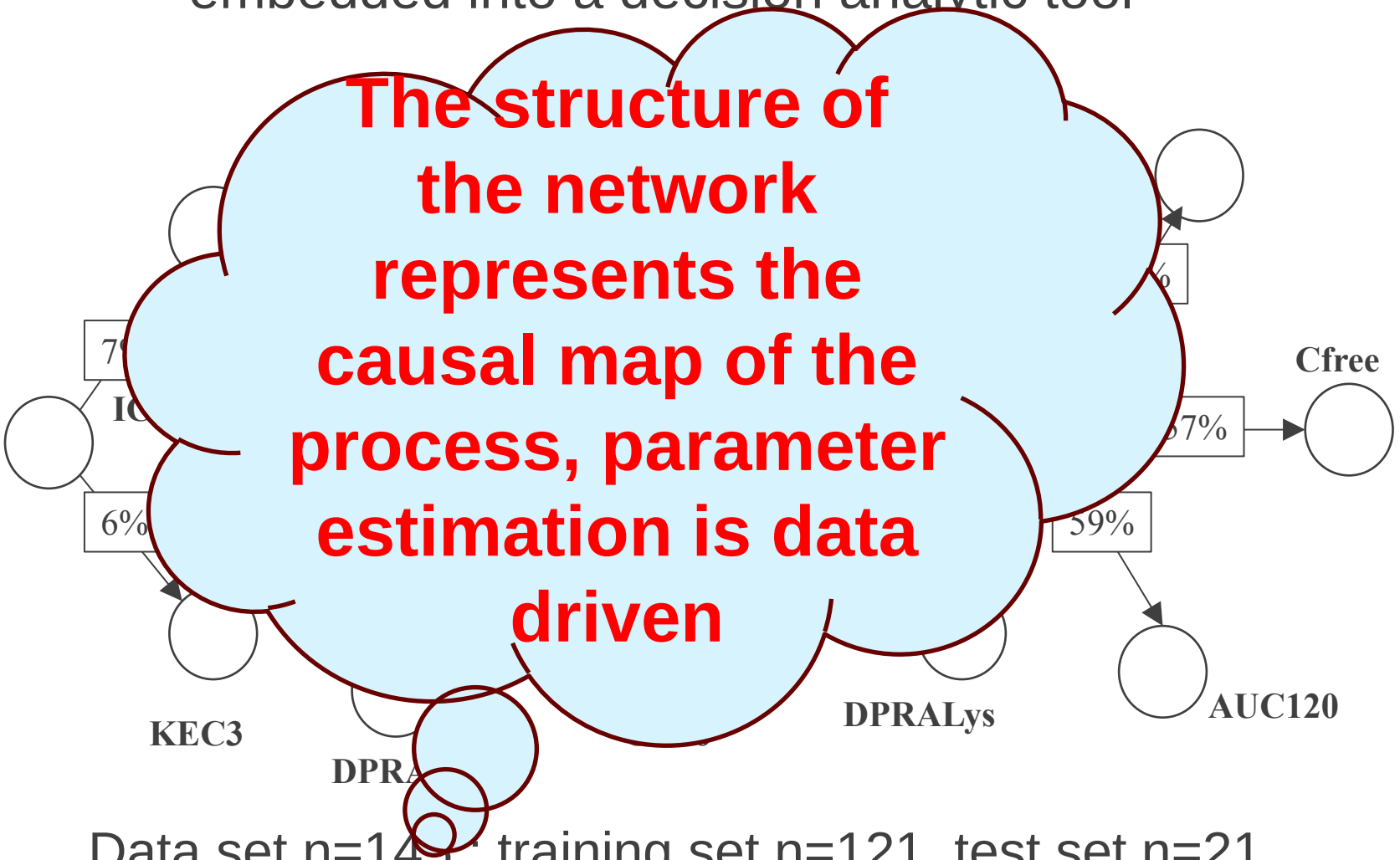
Reactivity data
Direct Peptide Reactivity Assay (DPRA)
KeratinoSens™ Assay (Ksens) (KEC 1.5, KEC3, IC50)

Dendritic cells activation: CD86

TIMES-M

J. Jaworska; P&G

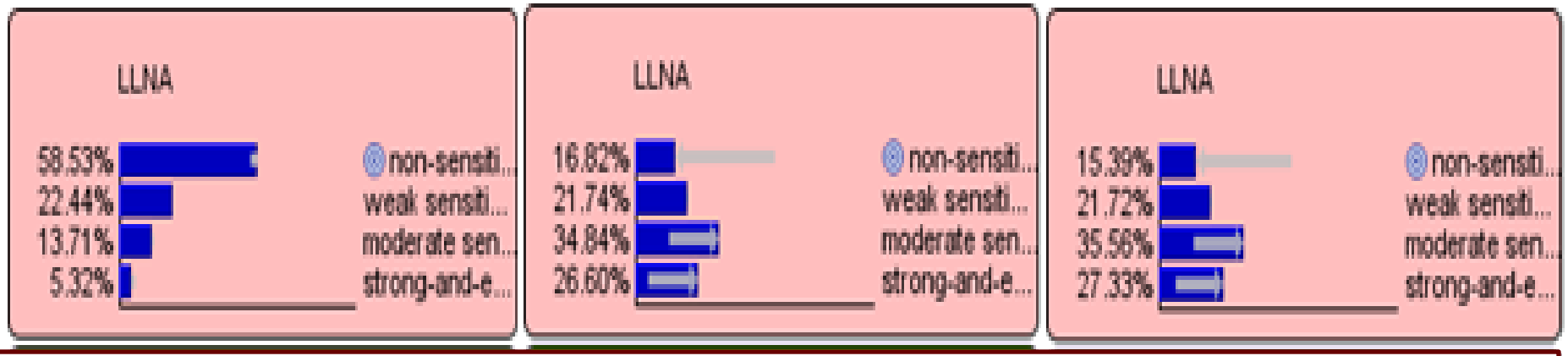
BN ITS 2 - abstracted skin sensitization process embedded into a decision analytic tool



Data set n=149 : training set n=121, test set n=21

BN ITS topology and AOP concepts are very much alike ! J. Jaworska; P&G

How does the final answer look like ?



P(LLNA=NS, W, M, S| evidence)

«Each bit of evidence (*in silico, in vitro) changes the overall probability that a chemical falls in one particular LLNA class – the most likely class is our prediction*»

J. Jaworska; P&G

Probability attributions to different LLNA classes: Test set

Table 4. The Integrated Testing Strategy (ITS)-2 probability predictions for each p
cells denote the experimental local lymph node assay (LLNA) potency class. Boldec

		P (LLNA=			
Chemical	MA	**N**	W	M	S
Chlorobenzene	0	**93**	4	0	3
Isopropanol	0	**98**	1	0	1
Lactic acid	0	**98**	1	0	1
Methyl salicylate	0	**97**	1	0	2
Salicylic acid	0	**84**	3	0	13
2-Hydroxypropyl methacrylate	1	**95**	3	2	0
Hexyl cinnamic aldehyde	1	0	**99**	1	0
Imidazolidinyl urea	0	0	**95**	5	0
Eugenol	0	4	**69**	27	0
1-Iodohexane	0	11	**87**	2	0
Cinnamyl Alcohol	0	**78**	8	13	0
Citral	0	0	3	**93**	4
Dihydroeugenol	0	4	9	**62**	25
Isoeugenol	0	0	16	**67**	17
2-Mercapto-benzothiazole	0	0	**46**	37	17
2,4-Heptadienal	1	0	1	39	**60**
Fluorescein-5-isothiocyanate	0	0	2	23	**75**
1-Chloro-2,4-dinitrobenzene	0	0	0	45	**55**
1,4-Phenylenediamine	0	0	0	**99**	1
Tetrachloro-salicylanilide	0	0	0	16	**84**
5-Chloro-2-methyl-4-isothiazolin-3-one	0	0	0	20	**80**

J. Jaworska; P&G

ITS-2 performance

		Training set n=121					Test set n=21			
		observed					observed			
		NS (36)	W (28)	M (35)	S(25)		NS (6)	W (5)	M (5)	S(5)
pred	NS (37)	**31**	2	1	3	NS (7)	**6**	1	0	0
	W (27)	2	**22**	2	1	W (5)	0	**4**	1	0
	M (27)	2	3	**19**	3	M (4)	0	0	**3**	1
	S (33)	1	1	13	**18**	S (5)	0	0	1	**4**

Summary BN ITS-2

- In the external validation (n=21) ITS-2 predictions were 86% correct for potency, 95% for hazard. Why ?

1) ITS network structure that follows mechanistic steps of skin sensitization induction process (including bioavailability);
   - BN ITS topology and AOP are very similar (but: We account for the fact that the boxes are not nicely separated)
2) a large dataset used to parameterize the ITS;
3) Probabilistic framework for inference and testing strategy development.

- Flexible and Adaptive Testing Strategy
  - BN ITS-2 consistently resolves conflicting evidence, deals with different set of evidence, missing data
  - BN ITS-2 can guide testing and identify impact of generating new data before testing

Going forward ITS-2

- Short-term improvements for which solutions are already emerging include:
  - better discrimination between Moderate and Strong sensitizers
    - Increase the dynamic range of reactivity assay: kinetic profiling , PPRA
  - improved detection of prehaptens and prohaptens
    - PPRA
  - Calculate network with h-CLAT instead of U937 data, as h-CLAT is closer to prevalidation
- Longer term improvements relate to advancing mechanistic knowledge, especially mechanisms that are not well characterized by a combination of reactivity and DC assays.

- Currently BN ITS-2 predicts probability distribution of LLNA classes i.e 'most likely EC3 range'
  - ideally we would predict not the 'most likely EC3 range' but a 'most likely EC3 value with x % certainty (i.e. 90% certainty) ' – critical for Quantitative Risk Assessment

Part III: Key results from multiple regression study on 244 chemicals within Givaudan

Parameters contributing to prediction: analysing a database with reactivity data and KeratinoSens™ data on 244 chemicals

- Database with modified peptide binding assay comprising kinetic measure for highly reactive chemicals
  - Natsch, A. and H. Gfeller (2008). "LC-MS-based characterization of the peptide reactivity of chemicals to improve the in vitro prediction of the skin sensitization potential." Toxicol. Sci. 106(2): 464-478.
- Database: 244 chemicals tested in this assay and in KeratinoSens™, including calculated PhysChem parameters
- Contribution of individual parameters to LLNA potency, determined based on multiple regression

Givaudan

Multiple regression

- Best simple formula to predict LLNA EC3:

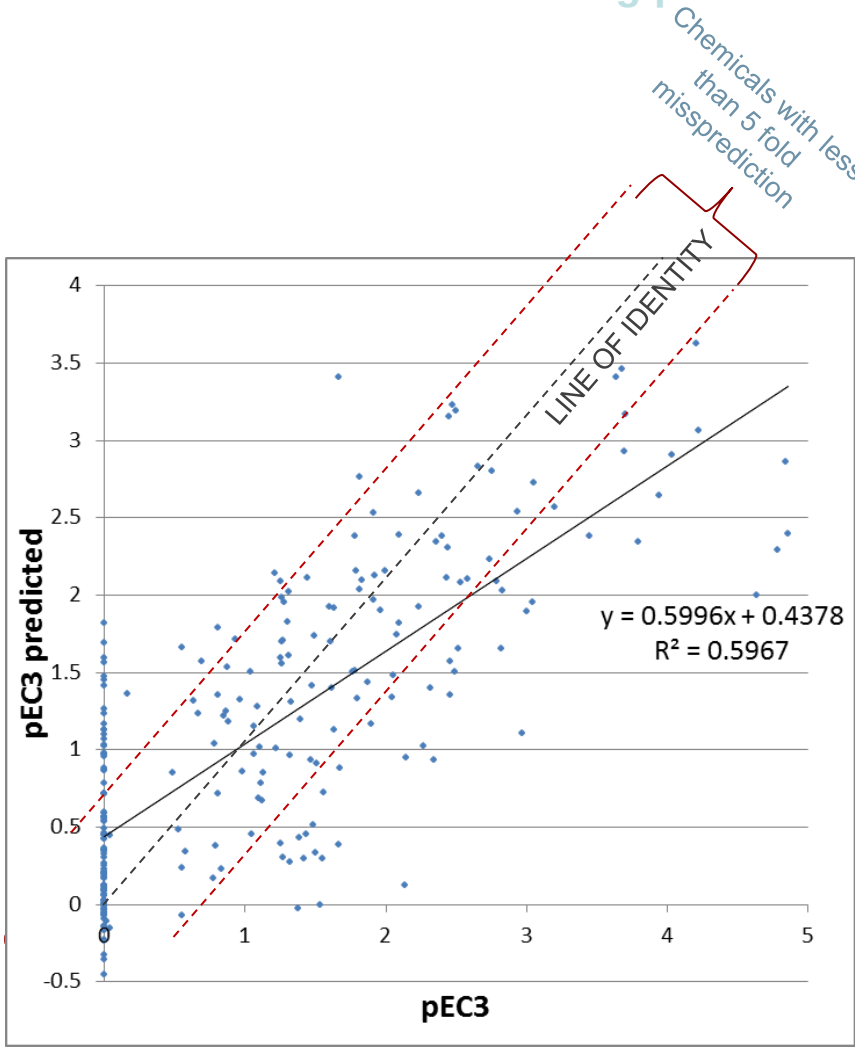
pEC3 = 0.34 + 0.09 × ln EC 1.5 + 0.11 × Ln IC50 + 0.16 × ln K + - 2.01 LOG boiling point

Predictor	Coef	SE Coef	T	P
Constant	0.3417	0.1204	2.84	0.005
ln EC 1.5 (KS)	0.09353	0.02529	3.70	0.000
ln K (reactivity)	0.15716	0.01764	**8.91**	0.000
Ln IC50 (KS)	0.11358	0.03259	3.49	0.001
LOG boiling point	-2.0150	0.5567	-3.62	0.000

N= 244; R-Sq = 59.7%

Median misprediction: 2.3-fold

GeoMean misprediction: 3.3-fold



Key learnings regression analysis

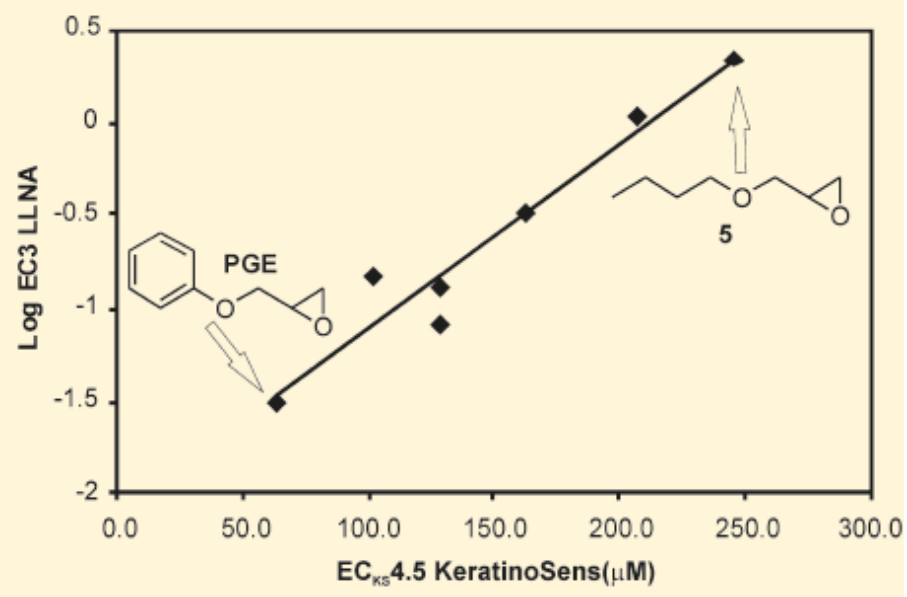
- Peptide binding rate has strongest overall contribution to potency
- From KeratinoSens™ both the EC1.5 and the IC50 have highly significant contribution to overall regression
  - IC50 may be a simple mimic of “danger signal”
  - Cytotoxicity has been reported before as important contributor
- cLogP has no contribution overall, but contribution of boiling point (volatility) is highly significant, negative coefficient
  - This reflects the open application of LLNA. Highly reactive but highly volatile chemicals are weak in LLNA
- Overall 60% of LLNA variation can be explained by these four parameters
  - Keep in mind that LLNA data and *in vitro* data themselves have an intrinsic variability

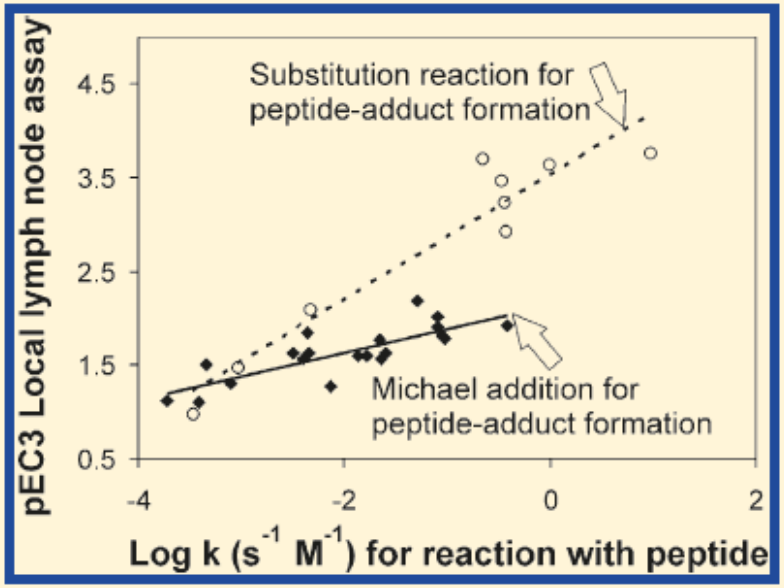
Givaudan

Part IV: Development of “Local” models to improve predictivity

Local Models - Predicting within mechanistic classes

- Within well defined mechanistic classes, prediction can often be made even with a single test
- Can probably be improved by class-specific ITS
- BUT (a BUT in capitals!): Only a limited number of chemicals falls in well defined classes with sufficient chemicals with animal / human data available!





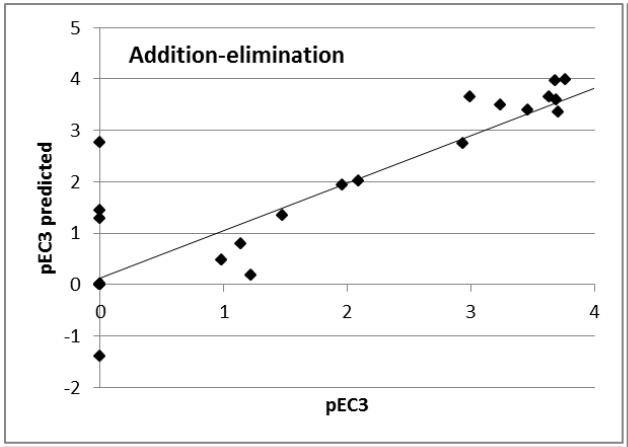
Delaine, T., Niklasson, I. B., Emter, R., Luthman, K., Karlberg, A. T., and Natsch, A. (2011) Structure-Activity Relationship between the in Vivo Skin Sensitizing Potency of Analogues of Phenyl Glycidyl Ether and the Induction of Nrf2-Dependent Luciferase Activity in the KeratinoSens in Vitro Assay. *Chem. Res. Toxicol.*, *24*, 1312-1318.

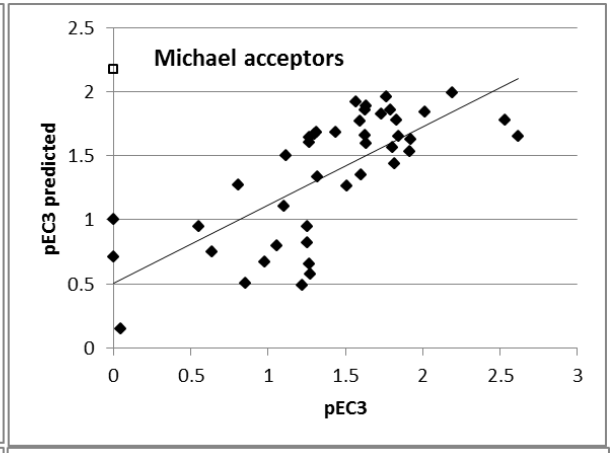
audar

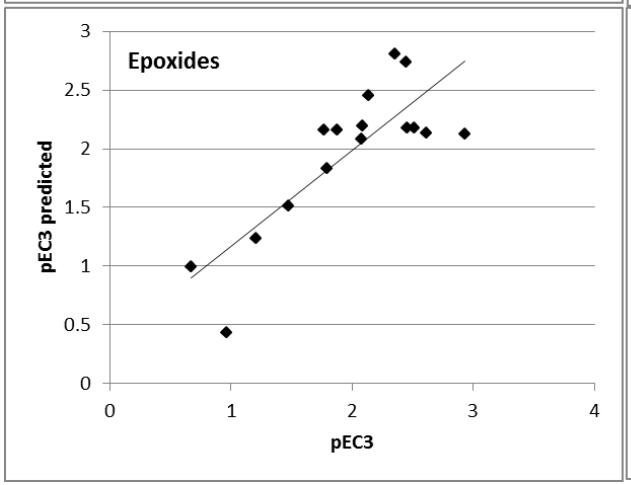
Natsch, A., Haupt, T., and Laue, H. (2011) Relating skin sensitizing potency to chemical reactivity: reactive Michael acceptors inhibit NF-kappaB signaling and are less sensitizing than S(N)Ar- and S(N)2- reactive chemicals. *Chem Res Toxicol*, *24*, 2018-2027.

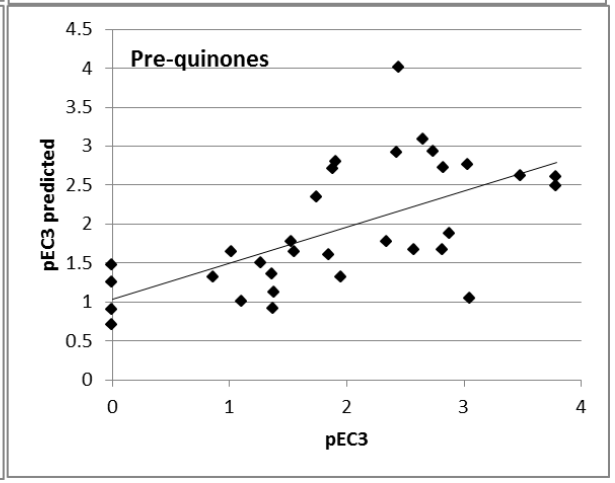
Potency prediction – integrated testing strategy - internal data integration project

- Local models vs LLNA data









Leave-one-out approach: Predict chemicals based on data from related chemicals.

For several domains, better predictions by this approach

Prediction more difficult for chemicals which need activation (pre-quinones)

Givaudan

What is still required for future development for sensitisation prediction use in QRA

- Potency prediction must be able to inform risk assessment
  - Need to understand uncertainty and how to manage this
- Probabilistic / Bayesian approach to predict “most likely” NESIL levels
  - Can be predicted LLNA EC3, better human NESIL
- Compare predictions with global models (i.e. global Bayesian net, global regression, …) and with **local models**
  - Local models: Models to predict within specific structural domains, may be more accurate
  - Will the same *in vitro* test battery be ideal for all chemicals or does testing need to be adapted to the structural class?
- Integrate h-CLAT data
- Evaluate which emerging *in vitro* models add further, non-redundant information for potency

Givaudan

Thank you for your attention

Givaudan

ENGAGING
THE
SENSES