Bayesian Integrated Testing Strategy (ITS) for skin sensitization potency assessment

– a decision support system for quantitative weight of evidence and adaptive testing strategy

(BN ITS-3)

Joanna S. Jaworska[1],

Andreas Natsch[2], Cindy Ryan[1], Judy Strickland[4], Takao Ashikaga[5], Masaaki Miyazawa[6]

[1] Procter and Gamble Company

[2] Givaudan Schweiz AG,

[4] ILS/Contractor Supporting NICEATM, Research Triangle Park, NC 27709, USA.

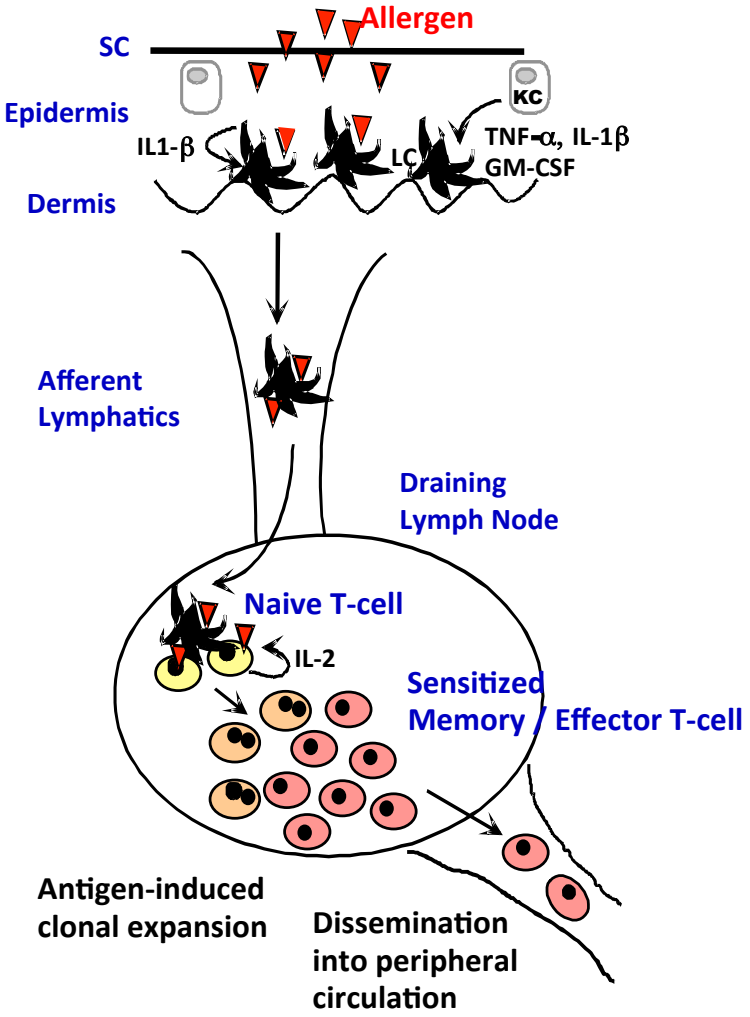
[5] Shisheido Company Limited,

[6] Kao Corporation.

Archives of Toxicology (2015), 89:2355–2383

Mechanism of Contact Sensitization

Sensitization Phase



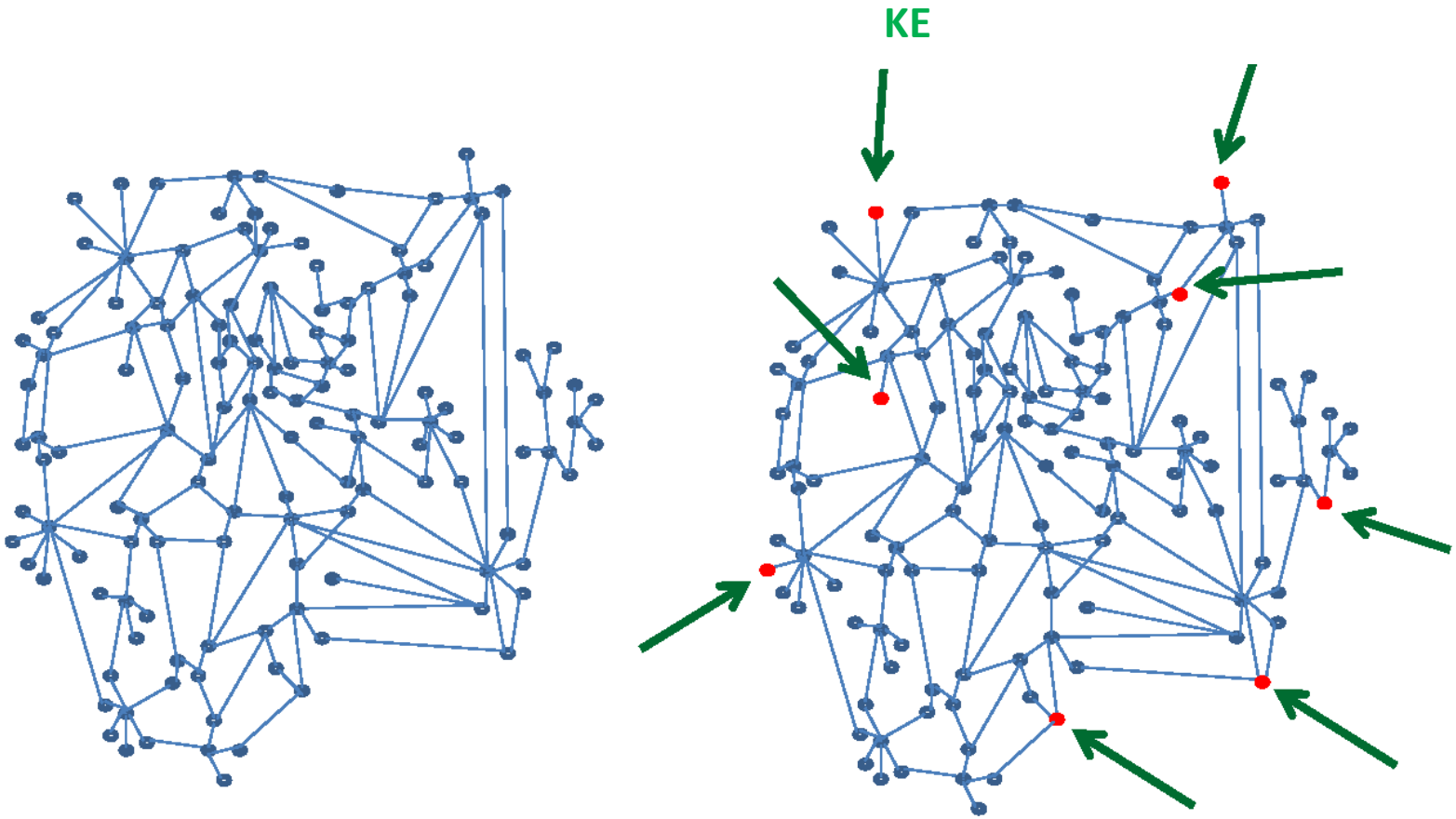
Key Event 1: covalent binding to skin proteins

Key Event 2: activation of inflammatory cytokines and induction of cytoprotective genes in the keratinocyte

Key Event 3: activation (induction of inflammatory cytokines and surface molecules) and mobilization of dendritic cells in the skin

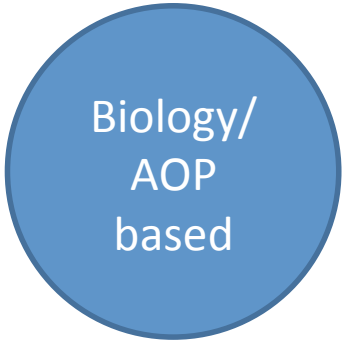
Key Event 4: activation and proliferation of antigenspecific T-cells

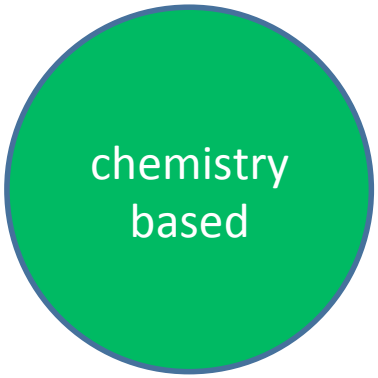
OECD AOP 2012


KE

Predicting skin sensitization potency

- Skin sensitization in vivo potency spans 4+ orders of magnitude
  - Substantial variability (Hoffmann 2015)
  - Well established in vivo assay - LLNA





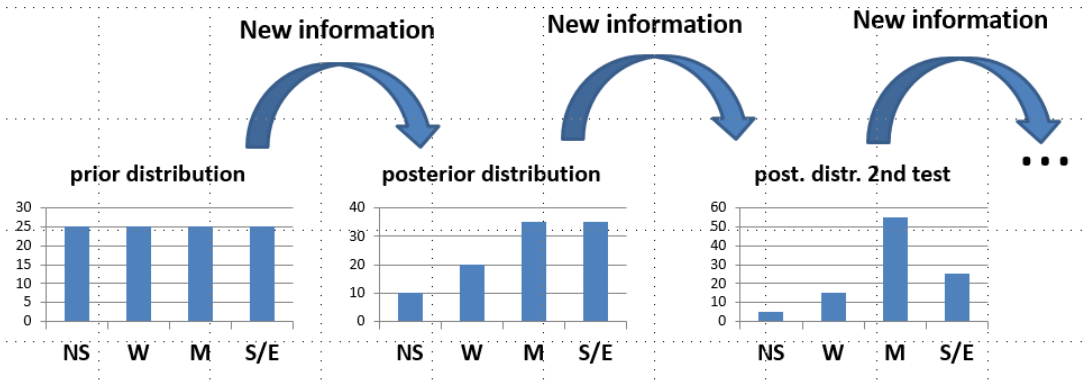
  - KE1: DPRA
  - KE2: Keratinosens
  - KE3: h-CLAT

Kinetics (D. Roberts et al.)

- Units
- IVIVE from KE4 (Unilever)
- ?

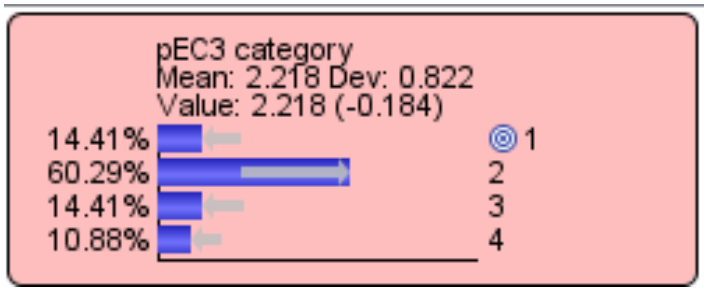
Bayesian Network ITS approach

Feature	Function
The structure of the BN ITS represents the underlying mechanistic processes leading to an in vivo adverse effect (AOP, knowledge maps, systems of ODS) while recognizing the uncertainty of the exact formalism.	Allows interpretation in the biological context and is chemical specific.
In BN ITS framework integration is on the level of data and not existing models of single assays	Eliminates potential inconsistency and uncertainty propagation due to use of the prediction models of multiple individual assays.
Information overlap between individual assays regarding adverse effect is accounted for.	Reduces false positives and false negative classifications,
Can build a hypothesis with partial data in any sequence	Flexible and adaptive. Data outside the applicability domain of individual tests can be eliminated.
Quantifies uncertainty for the hypothesis with any partial data	Facilitates consistent prediction acceptance criteria. Guides testing strategy using Value of Information.



Endpoint and Purpose

- Endpoint : skin sensitization potency in the LLNA
  - expressed as probability distribution of LLNA pEC3 class
  - Can be transformed to EC3% probability percentile (‘ most likely EC3 value’)



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

EC3% 50th or any other percentile

- Purpose/ regulatory decision type
  - Hazard
  - classification and labeling under the GHS C&L scheme
  - **in quantitative risk assessment especially when combined with in vivo evidence on analogues.**
  - Efficient testing strategy: Decision whether additional tests are needed

Development of Integrated Testing Strategies

On This Page

Most traditional toxicity testing methods involve treating a laboratory animal with a test substance and observing adverse effects. This approach is expensive and time-consuming, and the use of animals raises ethical concerns and issues of interspecies extrapolation. Using cell-based, biochemical, or computational methods to predict chemical toxicity would overcome these drawbacks, but it usually takes several nonanimal tests to provide the same level of information as single animal test. An **integrated testing strategy** provides a means by which all available relevant information about a chemical can be considered in a structured and reproducible manner to arrive at a hazard classification decision. Integrated testing strategies could potentially be used to support decisions about whether chemicals require hazard labeling without the use of animal test data.

Development of Integrated Testing Strategies to Identify Potential Sensitizers

There is a growing international need for nonanimal test methods to identify skin sensitizers (substances with the potential to cause allergic contact dermatitis). NICEATM is working with other NTP scientists and industry experts to create an integrated testing strategy to combine information from multiple testing methods to identify potential skin sensitizers.

View reference:
Pirone et al. 2014. Open source software implementation of an integrated testing strategy for skin sensitization potency based on a Bayesian network. ALTEX 31:336-340.

NTP scientists are collaborating with Dr. Joanna Jaworska and colleagues at the Procter & Gamble Company (P&G), who have developed an integrated testing strategy for identifying potential skin sensitizers without conducting animal tests. The testing strategy uses a Bayesian network to analyze data from nonanimal tests and other information about a test substance, such as chemical structure and solubility, to identify potential skin sensitizers. The analysis considers all the available relevant information about a substance and produces a numerical probability that the substance is a sensitizer. This probability could potentially be used to make decisions about whether substances require hazard labeling, without requiring animal testing.

More information about the testing strategy can be found in the following publications:

- Jaworska et al. 2011. Integrating non-animal test information into an adaptive testing strategy — skin sensitization proof of concept case. ALTEX 28:211-225.
- Jaworska et al. 2013. Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice. J Appl Toxicol 33:1353-1364.

Previously two versions of ITS published: What's new in ITS3?

- Refine precision and accuracy of potency prediction, especially for the moderate and strong sensitizers.
  - **Better integrate chemistry and biology**
  - **Refine handling of bioavailability (in vitro, ionized chemicals)**

- Standardization of inputs , Standardized prediction process
  - Replace test for Key event 3: dendritic cell activation **From U937 to hCLAT**
  - **Examine applicability domains** as a way of eliminating non- reliable data
  - Include consistent criteria for accepting prediction based on **uncertainty quantification**

ITS-3 Database

- 207 chemicals with full records for the three ECVAM validated assays
  - 147 training set , 60 test set
- Sources
  - ITS-2 data n=145
  - hCLAT database from Kao and Shiseido
  - New data from P&G, Givaudan, Kao, Shiseido to fill data gaps (n=30)
  - Additional chemicals set from RIFM n=40

Input Type	Endpoint	Unit
Bioavailability	W_S – Water solubility at pH = 7 Log D – Distribution coefficient at pH = 7 PB – Plasma protein binding fraction Fraction ionized at pH = 7	M [-] [-] [-]
In silico prediction of potency in vivo: TIMES	Mechanistic alert for direct reactivity (including direct Michael acceptor) and auto-oxidation Prediction of 3 classes (non-sensitizer, weak, moderate/strong) based on the most potent among parent and metabolites.	Classes (NS, W, S)
Key Event 1: DPRACys, DPRALys	% of the cysteine (Cys) and lysine (Lys) peptide remaining in the DPRA	% peptide remaining
Key Event 2: KeratinoSens™ KEC1.5, KEC3, IC50	Concentration yielding 1.5-fold (KEC1.5); 3-fold (KEC3) induction of Nrf2-dependent luciferase activity; concentration resulting in 50% reduction of cell viability	µM
Key Event 3: h-CLAT EC150, EC200, CV75	Concentrations yielding 150% induction of CD86 cell surface marker; 200% induction in CD54; 75% cell viability	µM

Applicability Domain Assessment of In Vitro Assays: Filter Out Non-reliable Data

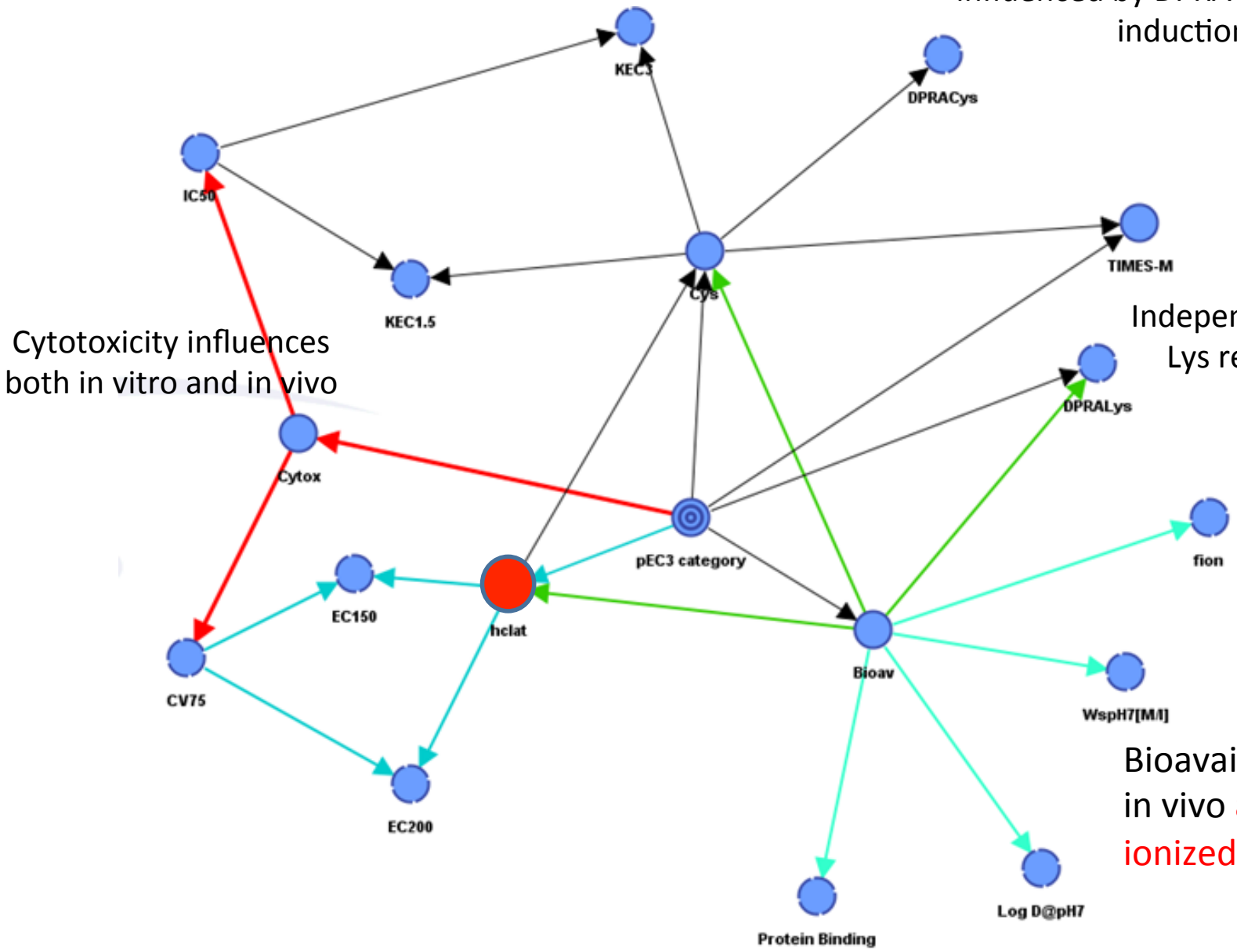
- Pre (auto-oxidation) or prohapten (metabolic activation) data DPRA , KS and hCLAT data are examined with caution. Hypothesis w/o these data is considered.

- Bioavailability in in vitro
  - Ionization: chemicals that are 100% ionized considered not suitable for *in vitro* assays.
  - Water solubility at pH=7 cutoffs for DPRA, KeratinoSens™, hCLAT

Ws at pH=7 [M/l]	DPRA	Keratinosens	hCLAT
<2.5e-08	x	x	x
2.5e-08 - 1.7e-04	ok	x	x
1.7e-04 - 2.1e-04	ok	ok	x
> 2.1e-04	ok	ok	ok

BN ITS-3 structure

Cysteine mode of action:
influenced by DPRACys and Nrf2 induction

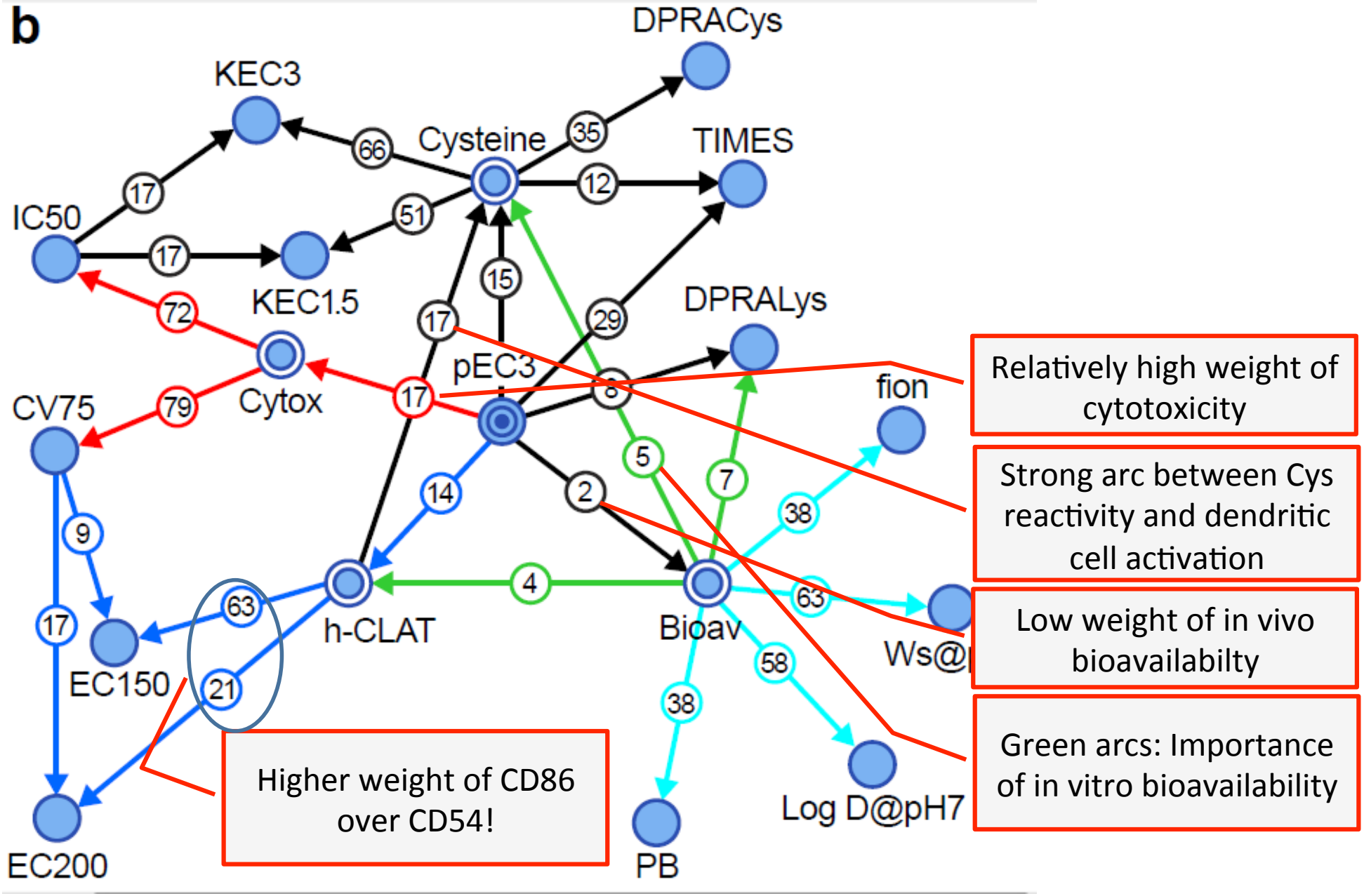


Cytotoxicity influences both in vitro and in vivo

Independent MIE:
Lys reactivity

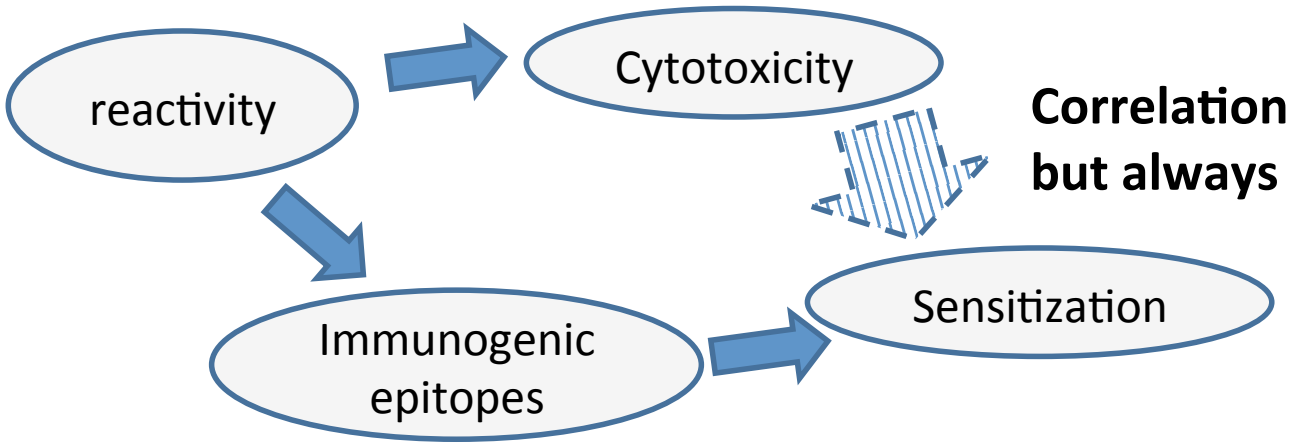
Bioavailability
in vivo and in vitro
ionized

The network arcs and relative mutual information



Weight of cytotoxicity: Is it a key potency determinant? – some considerations

- **<u>LLNA situation</u>**: In LLNA no adjuvans is given – **Molecule must provide danger signal** and reactive, immunogenic modifications (Difference from maximisation tests!)
  - Danger signal = local trauma, ATP and endogenous DNA release triggered by cytotox.
- Cytotoxicity correlates to irritancy – may trigger **<u>false-positives in LLNA</u>** – when training against LLNA we recapitulate that
- **<u>Cys-Reactivity triggers cytotoxicity</u>** – Cytotoxicity is an secondary effect of strong reactivity!



Correlation yes,
but always mechanistic link??

Process to derive prediction: gathering evidence

- Prediction of **physico-chemical properties** of chemicals (logD@pH7, Ws@pH7, fraction_ionized, protein binding)

- Prediction of **TIMES SS**:
  - Potency based on the highest potency among parent molecule and predicted metabolites
  - Assessment of potential of metabolic activations (prohapten) and autooxidation (pre-hapten)
  - reactivity alerts, direct Michael Acceptor

- Completeness of evidence on MIEs check: **Cysteine and Lysine reactivity?**
- Assessment of applicability domains/ elimination of data outside domain

Process to derive prediction- prediction

- Integration of all the in domain evidence and prediction of the pEC3 probability distribution
- Post processing step of probability distribution correction for direct Michael acceptors
- Conversion of probability distribution to Bayes' Factors for final interpretation and decision.

$$B = \frac{P(H = x|e)/P(H = not\_x|e)}{P(H|x)/P(H = not\_x)} = \frac{posterior\ odds}{prior\ odds}$$

Bayes Factor	Strength of evidence
<1	Negative (supports alternative)
1-3	Barely worth mentioning (weak)
3-10	Substantial
>30	Strong

Jeffereys, 1961

- Conversion from pEC3 to EC3% - Estimation of EC3% by percentiles: 50th and 90th percentile

Predictive capacity

GHS C&L	Observed ->									
	Training set n=147					Test set n=60				
	Class	NS(39)	W(39)	M(40)	S(29)	Class	NS(14)	W(18)	M(12)	S(15)
none	NS	**36**	2	1	0	NS	**14**	0	0	0
1B	W	2	**32**	3	3	W	0	**17**	3	0
	M	0	3	**38**	5	M	0	2	**9**	2
1A	S	1	2	8	**21**	S	0	0	0	**13**

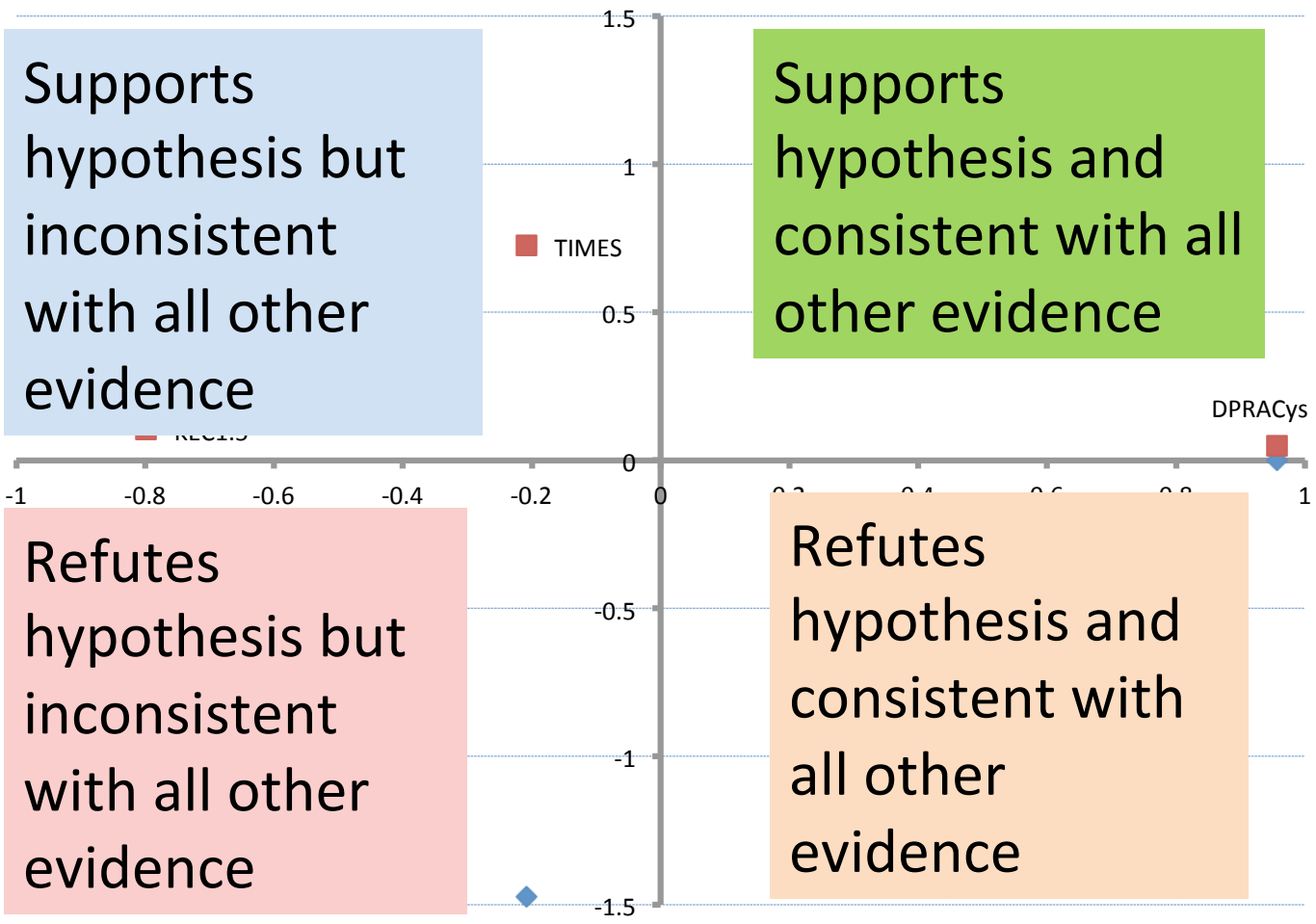
Test set	Hazard %	GSH C&L %	EC3 potency 4 class %
Balanced accuracy $bac = \frac{Se + Sp}{2}$	100	96	89

	all			C3 and C4 (similar to M+S)		
	n	Bac %	95% CI	n	Bac %	95% CI
ITS-2	21	85	70 - 100	10	80	55 - 100
ITS-3	60	89	81 - 97	27	82	68 - 96

Evidence Interpretation Chart

Impact given all other evidence

y(Xi)=log2(P(H|E)/P(H|E\Xi)), where E\Xi is the set of evidence excluding the piece of evidence Xi.

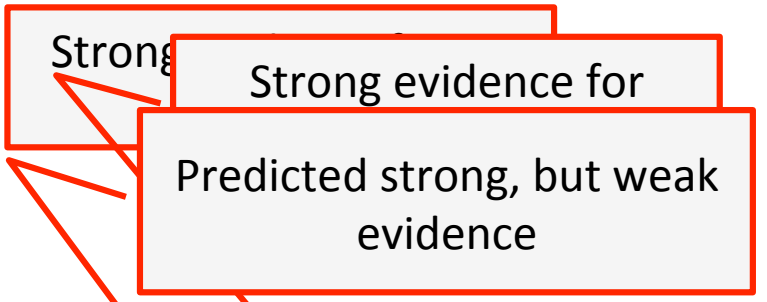


Consistency with all other evidence

x(Xi)=log2(P(Xi |E\Xi)/P(Xi)).

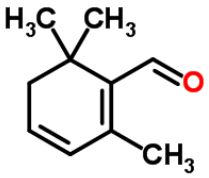
Application of BN-ITS-3

- **Bayes factors (B):**
  - **Which class is most probable and how reliable is this prediction**



	B („NS“)	B („weak“)	B („moderate“)	B(„srong/ extreme“)
Octannitrile	**129.1**	0.1	0.0	0.0
2-methyl-4H-3,1-benzoxazin-4-one	1.1	0.5	0.0	**5.1**
benzo(a)pyrene	0.11	1.11	1.60	1.75

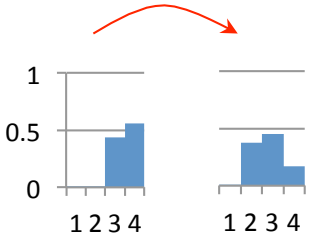
Example 1 - Safranal



hCLAT			Ksens			DPRA			Bioavailability			
EC150	EC200	CV75	KEC1.5	KEC3	IC50	DPRACys depletion	DPRALys depletion	TIMES-M	Log D @pH7	Protein Binding %	Ws@ pH=7	fion
176.2	256.2	456.8	5.4	33.5	337.3	91.8	0	3	2.8	40	0.008	0

1. TIMES-M
   a) parent – M or S, metabolite – M or S
   b) Di-substituted α,β-unsaturated aldehyde - direct MA
2. DPRA Lys ✓
3. Assessment of applicability domains- all tests are in domain ✓

MA correction



Bayes' factors for the ITS-3

	B(NS)	B(W)	B(M)	B(S)
All evidence	0.00	0.01	2.08	**5.10**
All evidence+ MAcorrection	0.01	1.67	**2.22**	0.82

EC3% 50th percentile 5.2; Safranal is at most M with p=83%

EC3% experimental – 7.5