Bayesian Network Integrated Testing Strategy

Decision support system for qWoE

Joanna Jaworska

Procter & Gamble

Modeling & Simulation

The power to transform.

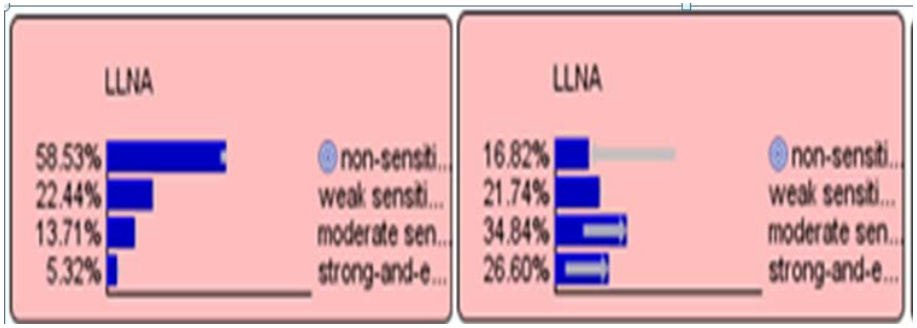
Brussels Innovation Center, Belgium

Cindy Ryan, Els Deconinck, Greg Dameron, Petra Kern, Frank Gerberick (P&G)

Andreas Natsch (Givaudan); Masaaki Miyazawa (Kao); Takao Ashikaga (Shiseido)

Endpoint and Purpose

- Endpoint : skin sensitization potency in the LLNA, TG 429, expressed as probability distribution of LLNA pEC3, very closely following 4 potency classes: nonsensitizers (NS), weak (W), moderate (M), and combined strong and extreme (S) sensitizers.



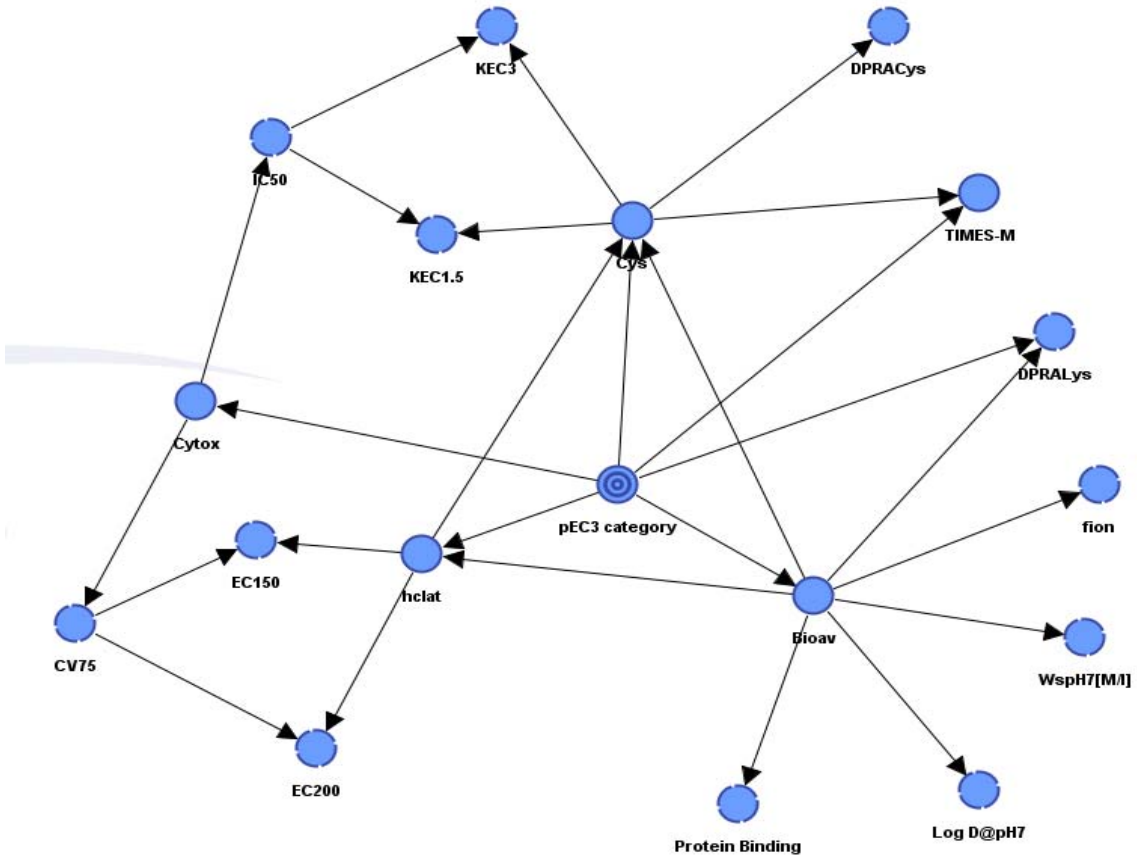
P(LLNA=NS, W, M, S| evidence)

- Purpose
  - Hazard
  - classification and labeling under the GHS C&L scheme
  - in quantitative risk assessment especially when combined with in vivo evidence on analogues.
  - In addition to data integration BN ITS/DS develops an efficient testing strategy. This IDS guides testing by Value of information and measures progress by uncertainty reduction.

Rationale of the Bayesian Network approach

Feature	What it does?
AOP structure – biological knowledge about skin sensitization 1. AOP sequence of events encoded 2. Cysteine and Lysine are treated as two independent molecular initiating events (MIEs). 3. Bioavailability consideration is applied to both in vivo and in vitro assays	Allows interpretation in the biological context chemical specific.
Biology (KE 1,2,3) and chemistry data directly encoded	Optimizes potency prediction, Eliminates uncertainty propagation due to use of individual assays prediction models.
Individual assays co-dependencies in the information they provide are accounted for (TIMES, Cys reactivity, hCLAT)	Reduces false positives and false negative classifications
Can build hypothesis with partial data	Data outside applicability domain can be eliminated Flexible
Quantifies uncertainty	Facilitates consistent decisions, Guides testing strategy using VoI

BN IATA structure



The structure of the BN ITS model represents abstracted AOP with the aim to follow sequence of the mechanistic events in the AOP.

Limitations

BN ITS-3 system requires biological data input of reliable consistent quality. The data need to come from within the applicability domains of the individual assays:

- *In vitro* assays are applicable to test chemicals soluble in either water or DMSO and test chemicals that form a stable dispersion;
- Highly cytotoxic test chemicals cannot be tested in the *in vitro* assays
- Prohaptens: varying metabolic capacity: DPRA < hCLAT < Keratinosens. Possible underestimation of potency.
- Prehaptens: experimental assays (DPRA , KS, hCLAT) results may yield underestimation of potency.
- Metals fall out of the applicability domain of the DPRA, since they are known to react with proteins with mechanisms other than covalent binding

Predictive capacity

GHS C&L	Observed ->									
	Training set n=147					Test set n=59				
	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	**16**	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	88

Process to derive prediction 1 gathering evidence

- Prediction of physico-chemical properties of chemicals
- 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:
  - Pre or prohapten data DPRA , KS and hCLAT data are examined with caution. Hypothesis w/o these data is considered.
  - 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

Process to derive prediction 2 prediction

- Integration of all the in domain evidence and prediction of the pEC3 probability distribution
  - Analysis of individual evidence and its combinations predictions
- Post processing step of probability distribution correction for Michael acceptors, if applicable.
- Conversion of probability distribution to Bayes Factors for final interpretation and decision.

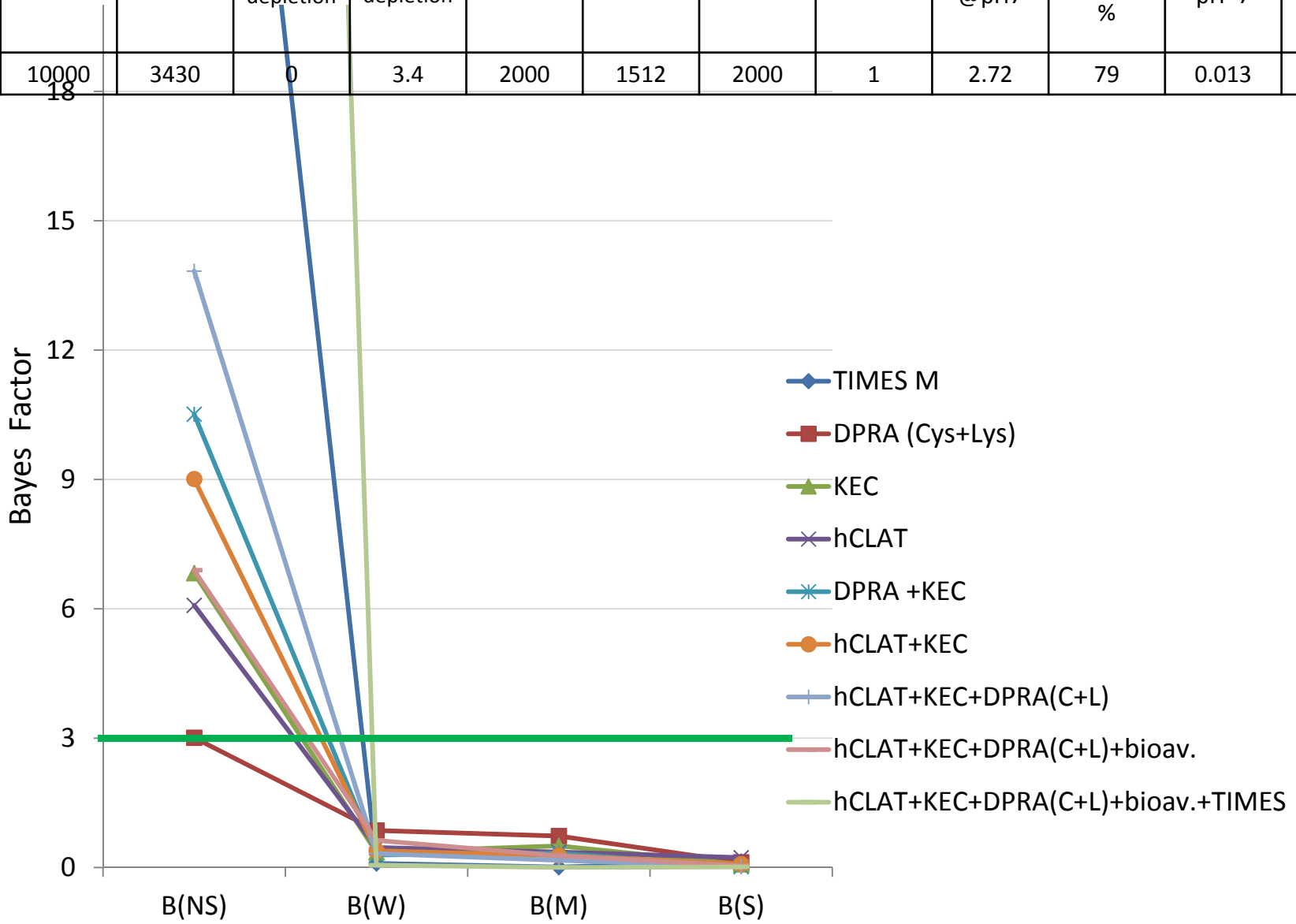
$$B = \frac{P(H = x|e)/P(H = not\_x|e)}{P(H|x)/P(H = not\_x)} = \frac{posterior\ odds}{prior\ odds}$$

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

Jeffereys, 1961

Prior distribution as in the training set				Posterior distribution predicted by BN ITS-3				Bayes factors			
Pr (NS)	Pr (W)	Pr (M)	Pr (S)	Pr (NS)	Pr (W)	Pr (M)	Pr (S)	B (NS)	B (W)	B (M)	B (S)
0.27	0.27	0.27	0.20	0.04	0.29	**0.37**	0.30	0.11	1.11	1.60	**1.75**

Octanenitrile 124-12-9-C(#N)CCCCCCC LLNA EC3% ND, nonsensitizer												
EC150	EC200	CV75	DPRACys depletion	DPRALys depletion	KEC1.5	KEC3	IC50	TIMES-M	Log D @pH7	Protein Binding %	Ws@ pH=7	fion
10000	10000	3430	0	3.4	2000	1512	2000	1	2.72	79	0.013	0.06



Octanenitrile-124-12-9 C(#N)CCCCCCCO LLNA EC3% ND, nonsensitizer