Data From In Vivo and In Vitro Tests Used To Identify Skin Sensitisers

David Basketter
DABMEB Consultancy Ltd, Sharnbrook, UK

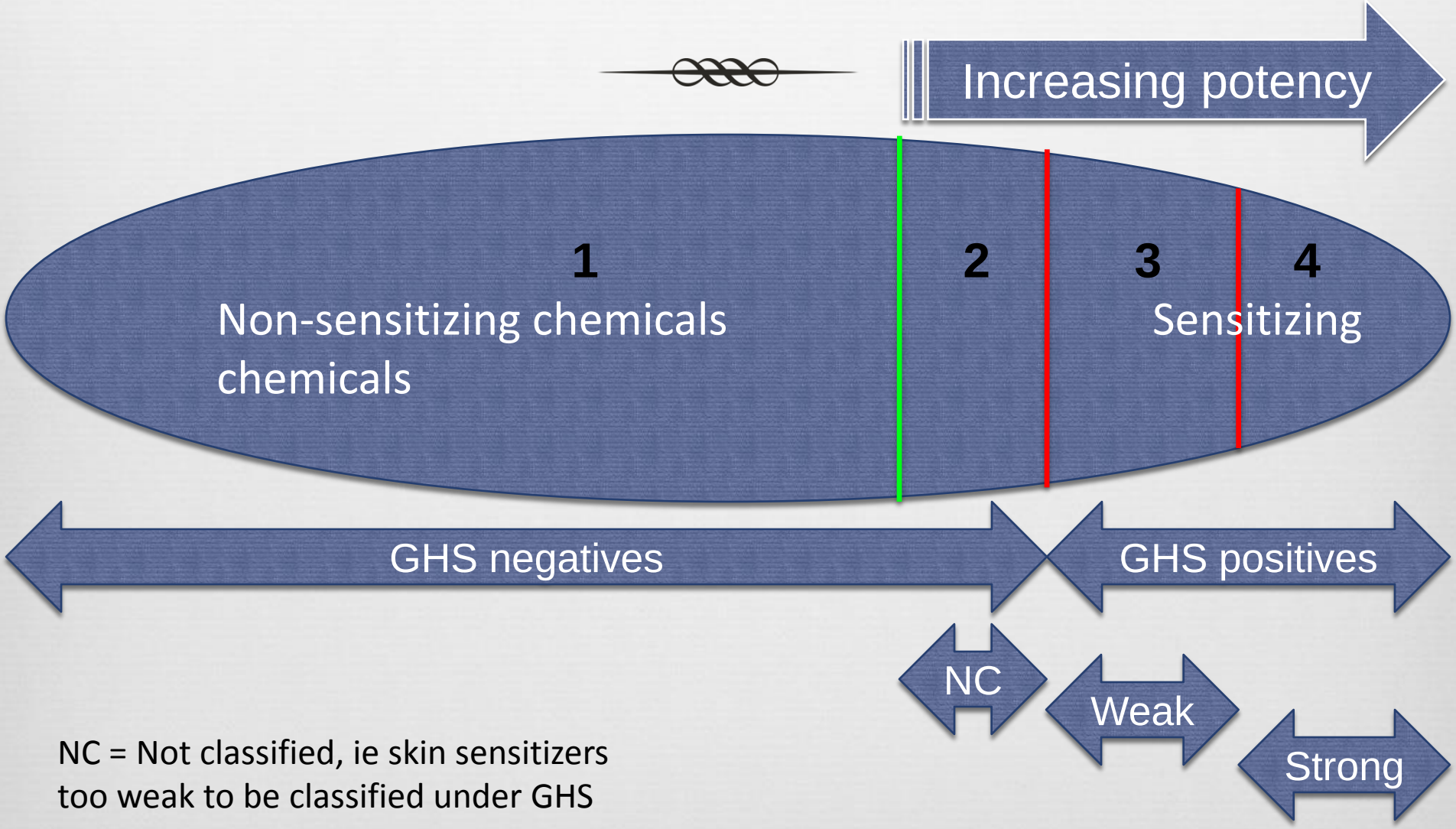
List of contents

- Definitions
- Mechanism
- Classification
- In vivo methods
- Human data
- In vitro methods
- Summary

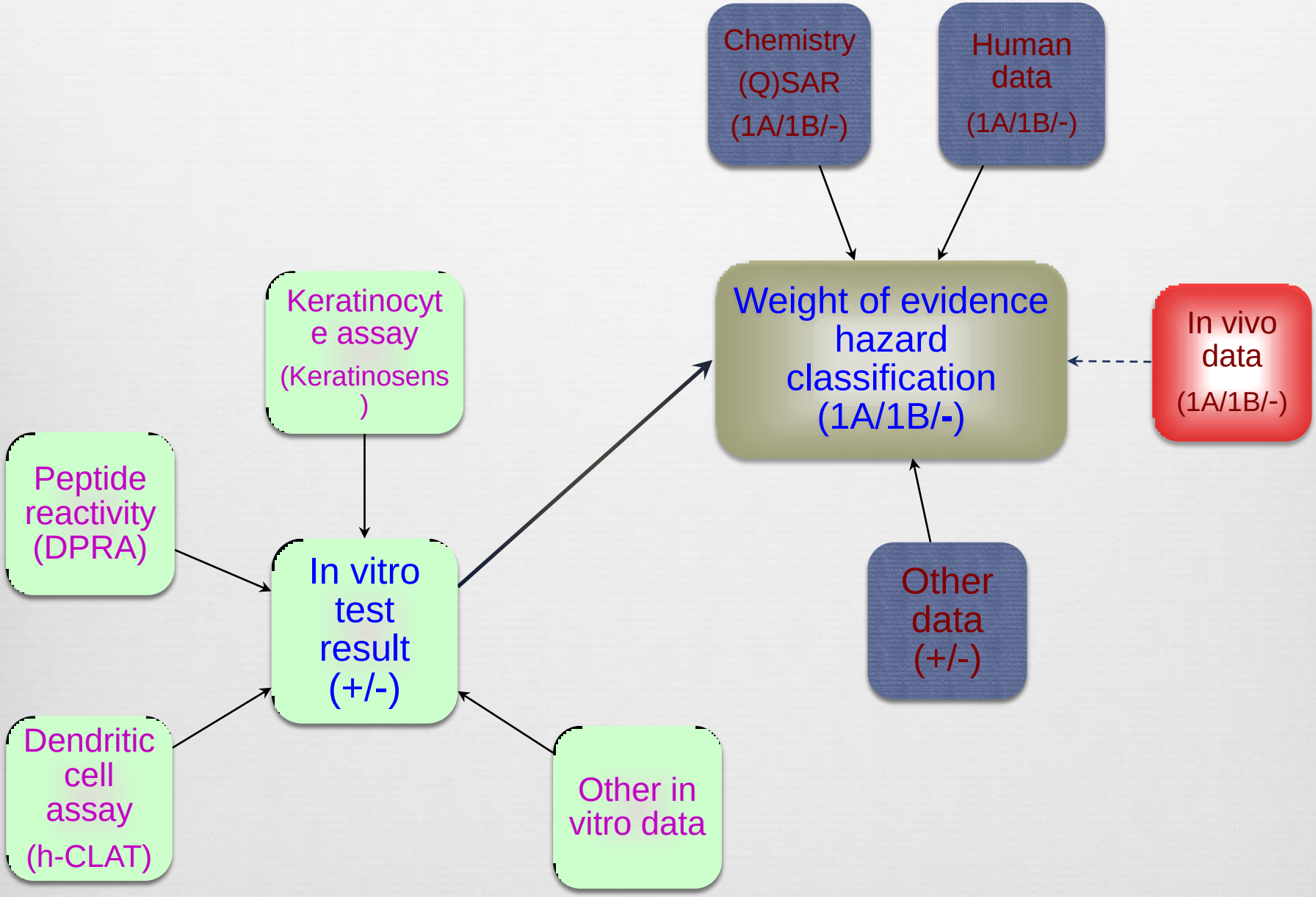
Definitions

- **Skin sensitiser**: a ***chemical*** which, with sufficient skin exposure, can induce…
- **Contact allergy**: the ***asymptomatic condition*** which an individual has when they are sensitised to a specific chemical and which is detected by a…
- **Patch test**: a clinical ***diagnostic procedure*** designed to reveal whether an individual has contact allergy and who is then (permanently) susceptible to…
- **Allergic contact dermatitis**: the ***eczema*** elicited by sufficient skin exposure to the skin sensitiser in an individual who has contact.

Regulatory classification



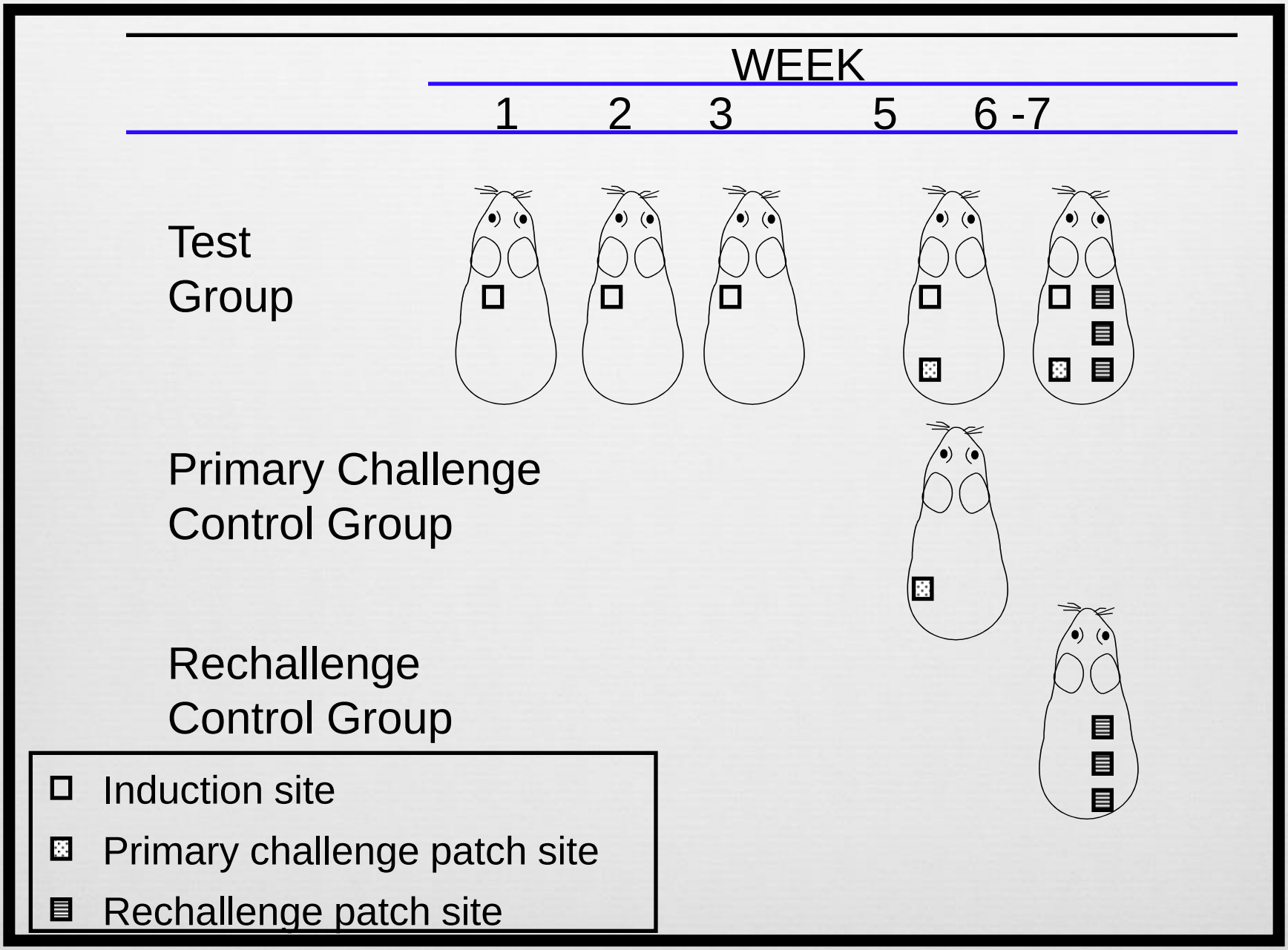
NC = Not classified, ie skin sensitizers too weak to be classified under GHS


Chemistry (Q)SAR (1A/1B/-)
Human data (1A/1B/-)
Keratinocyte assay (Keratinosens)
Weight of evidence hazard classification (1A/1B/-)
In vivo data (1A/1B/-)
Peptide reactivity (DPRA)
In vitro test result (+/-)
Other data (+/-)
Dendritic cell assay (h-CLAT)
Other in vitro data

Skin sensitization testing timeline

- 1944 – Draize test
- 1965 – Buehler test
- 1970 – M&K test
- 1982 – OECD 406
- 1982 – QSAR paper
- 1989 – LLNA paper
- 1992 – OECD update
- 1995 – Expert SAR system
- 1996 – In vitro pressure!
- 1999 – LLNA validated
- 2000 – LLNA training
- 2002 – OECD 429 LLNA
- 2004 – Peptide binding (DPRA)
- 2006 – h-CLAT papers
- 2007 – DPRA papers
- 2008 – LLNA under fire
- 2009 – Validation battery paradigm
- 2009 – ECVAM pre-validation
- 2010 – Pre-validation underway
- 2013 – EU Cosmetics deadline

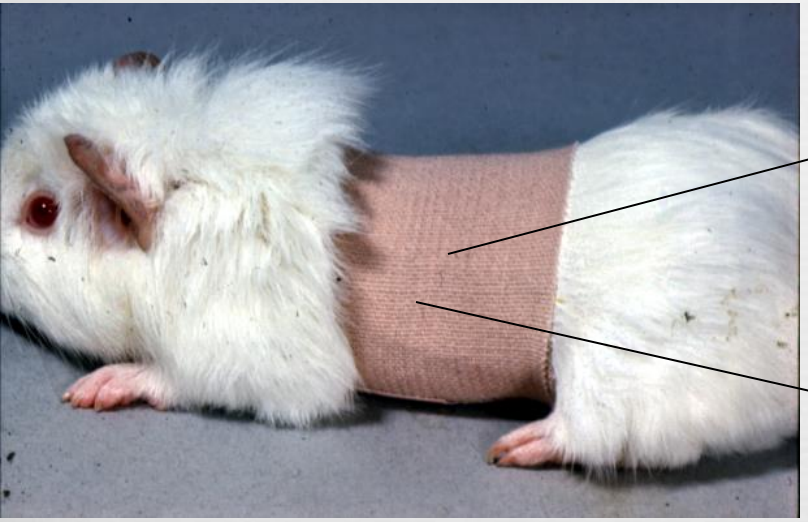
BUEHLER GUINEA PIG TEST

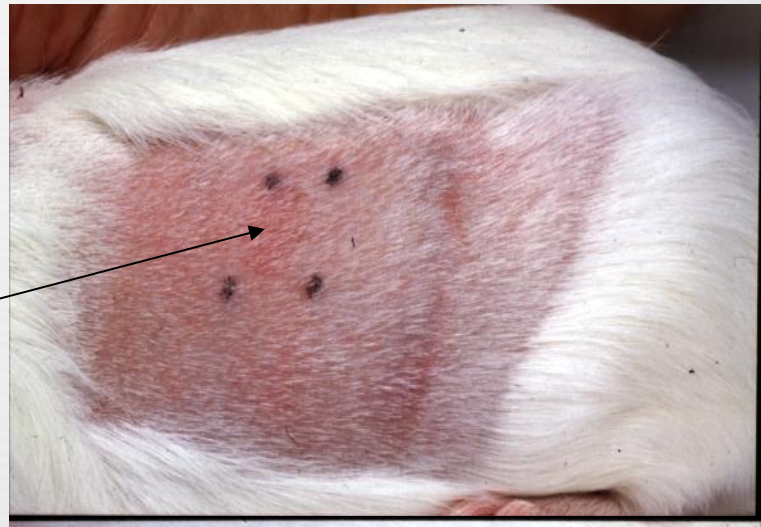


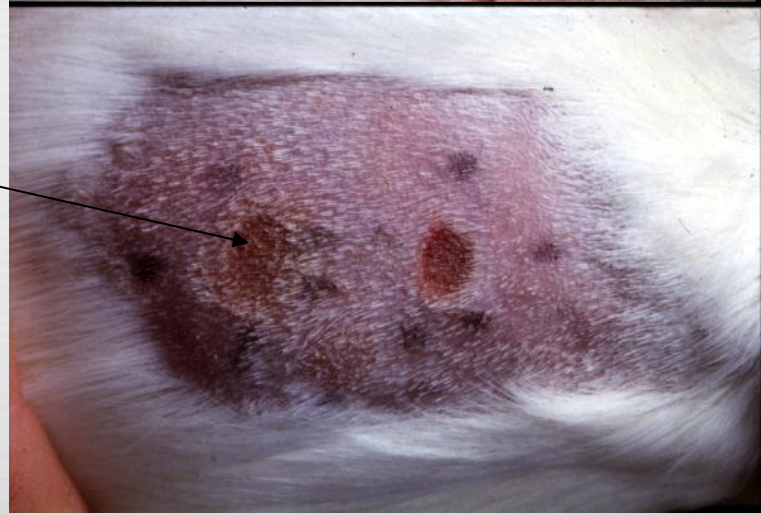
M&K Guinea Pig Maximization Test

- Week 1 - injection induction at the highest mild to moderately irritating concentration
- Week 2 - topical induction by 48h occluded patch at the highest mild to moderately irritating concentration
- Week 3 - rest
- Week 4 - 24h occluded patch challenge at highest non-irritating test concentration
- Week 6 - rechallenge?

M&K Maximization Test: Challenge





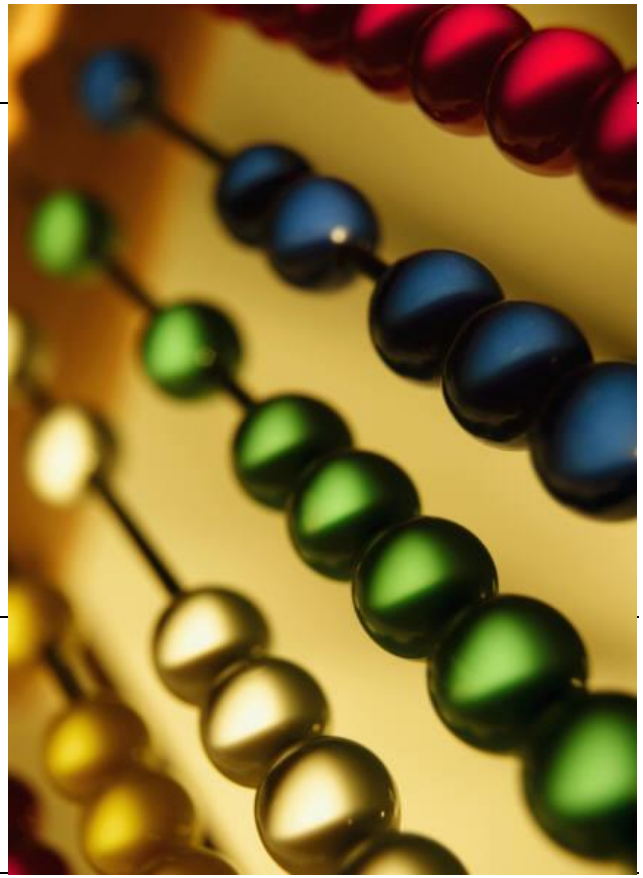


WEEK 4

Challenge

Table 1 An example of borderline data in guinea pig sensitisation testing: Substance X

Guinea pig no.	Primary challenge	
	24h	48h
1 (T)	0	0
2 (T)	0	0
3 (T)	0	1
4 (T)	1	1
5 (T)	0	0
6 (T)	0	0
7 (T)	0	0
8 (T)	0	0
9 (T)	1	2
10 (T)	0	0
11 (C)	0	0
12 (C)	0	0
13 (C)	0	0
14 (C)	0	0
15 (C)	1	0



T = test; C = control. Grading scale: 0 = no reaction, 1 = weak, 2 = moderate and 3 = strong

Challenge/Rechallenge

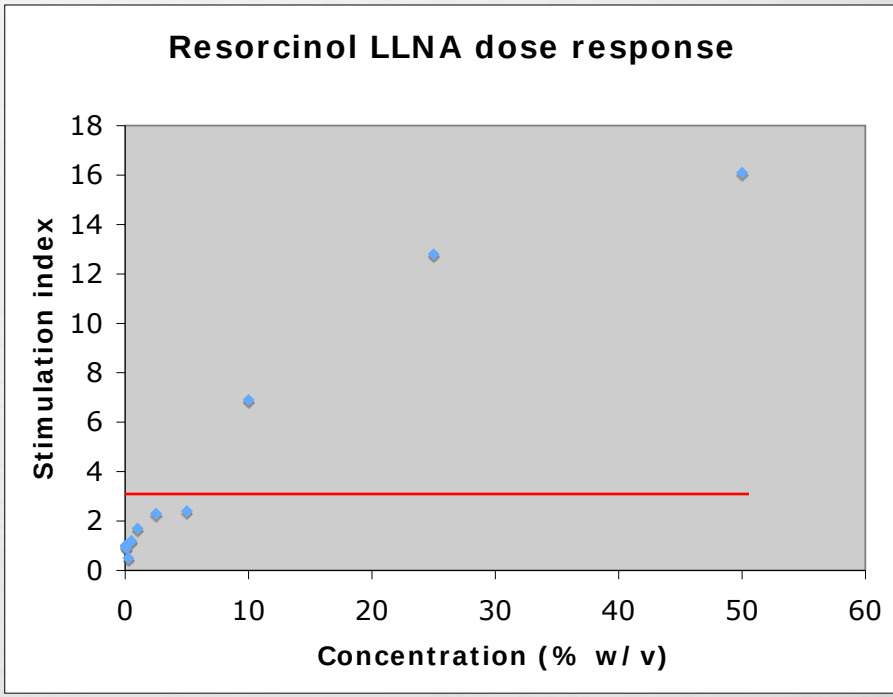
Table 1 An example of borderline data in guinea pig sensitisation testing: Substance X

Guinea pig no.	Primary challenge		Repeat challenge	
	24h	48h	24h	48h
1 (T)	0	0	1	0
2 (T)	0	0	0	0
3 (T)	0	1	0	0
4 (T)	1	1	1	1
5 (T)	0	0	0	0
6 (T)	0	0	0	0
7 (T)	0	0	0	0
8 (T)	0	0	0	0
9 (T)	1	2	1	1
10 (T)	0	0	1	1
11 (C)	0	0	0	0
12 (C)	0	0	0	0
13 (C)	0	0	0	0
14 (C)	0	0	0	0
15 (C)	1	0	0	0

T = test; C = control. Grading scale: 0 = no reaction, 1 = weak, 2 = moderate and 3 = strong

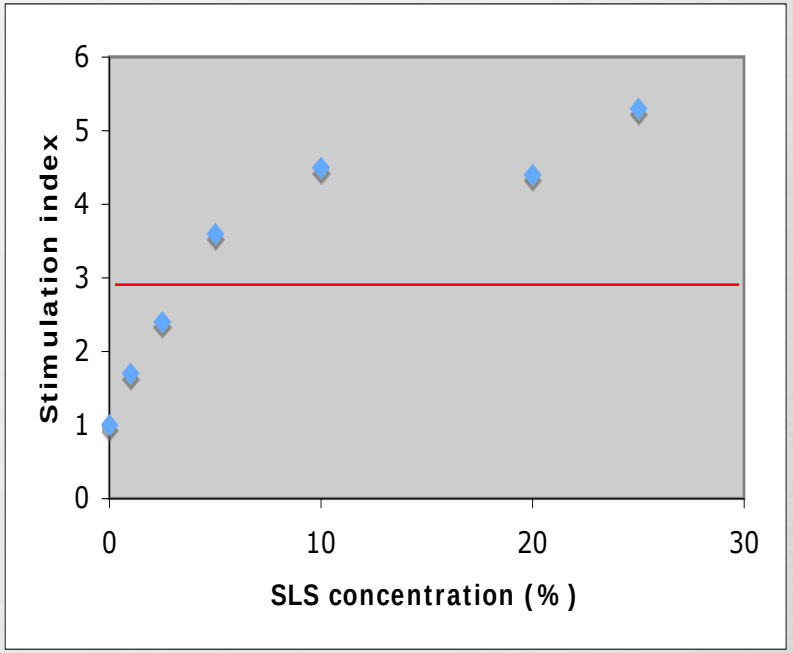
False positives in the LLNA?

- Resorcinol
  - The graph shows data combined from 2 separate experiments. At 25%, this weak sensitiser gave a SI of 12.8.
  - Human evidence of skin sensitisation has been reported.
  - Resorcinol has a plausible chemical mechanism for sensitisation.



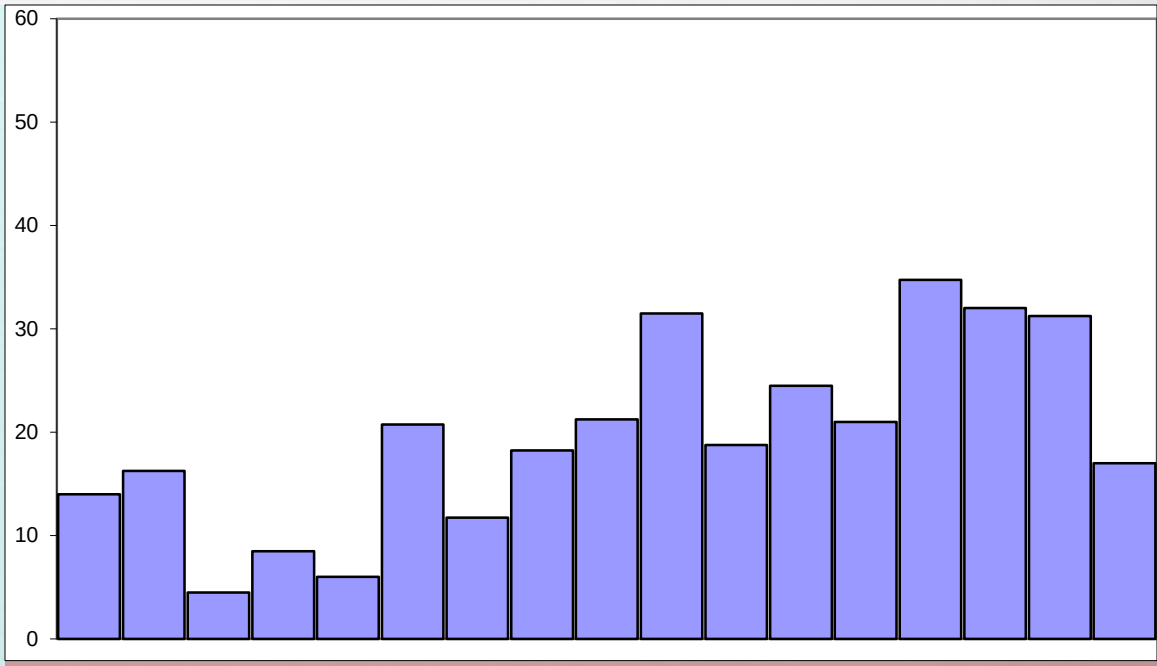
False positive in the LLNA?

- SLS: a true false positive.
  - The graph shows data combined from 2 separate experiments. At 25%, this strong irritant gave a SI of just 5.3.
  - Despite extensive exposure there is no human evidence of sensitisation.
  - SLS has no structural alerts
  - SLS is positive by B220



Reproducibility of GPMT

- 17 GPMTs - 5 years
- OECD 406 method
- Standardised doses
- Standardised vehicle
- Two HCA samples
- From 10% to 100% guinea pigs positive



1985 - 1989 OECD +ve control data for HCA

Even in a single GLP laboratory, the GPMT is variable

Reproducibility of GPMT

- OECD positive control hexylcinnamaldehyde actually from 0% to 100% across laboratories
- PPD reported in the range 10% to 100%
- Two highly respected laboratories in Denmark and Sweden struggled to get reproducibility with formaldehyde 50% v 95% +ve (Andersen et al, 1985); Grotan BK gave 20% v 75%
- Massive change in results with isoeugenol arose from minor alteration of test conduct (within OECD 406) (Basketter, 1994)

Intra and inter laboratory variation in the GPMT is very high; the Buehler test is similar

Buehler test variability

Study	Induction	Challenge	Response
1	10%	1%	70%
2	10%	1%	45%
3	10%	1%	40%
4	10%	1%	28%
5	10%	1%	26%
6	10%	1%	16%
7	10%	1%	11%

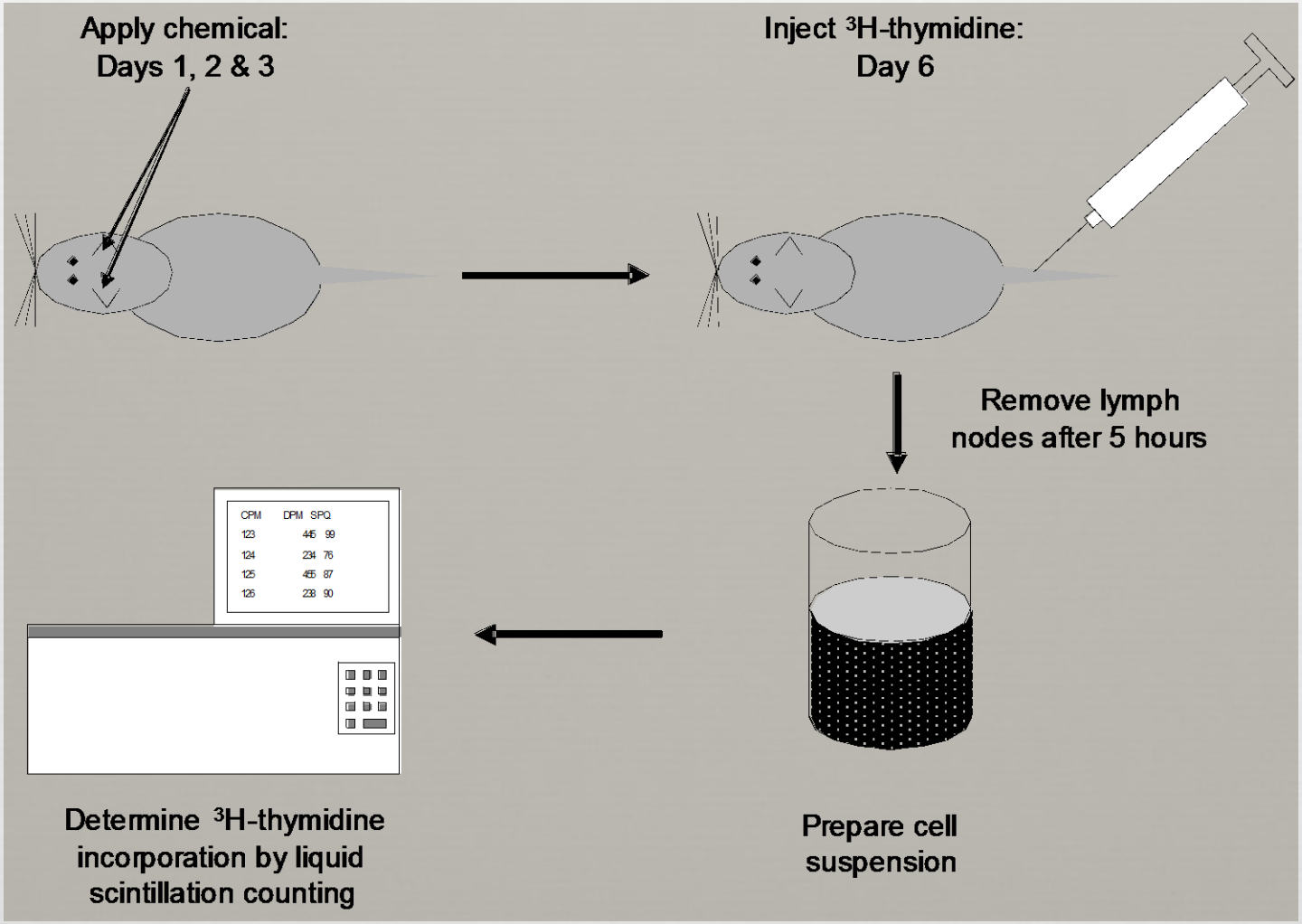
Things to consider….

- Test variability
- Subjective endpoint
- Opportunity to do the test badly
- Criticism of Freund's complete adjuvant in the M&K
- Criticism of the Buehler test sensitivity

- Elicitation dose response
- Opportunity to rechallenge
- Cross challenge
- Effect of vehicle on elicitation
- Sensitivity of the M&K versus the Buehler test
- False negatives/positives

…but remember that these tests have global acceptance and years of experience...

The Local Lymph Node Assay



LLNA output

- The output is quantitative data on ^{3}HTdR incorporation into the draining lymph nodes.
- Test data at the various concentrations are compared with concurrent vehicle control data.
- Where there is a 3 fold or greater stimulation in test versus control, the chemical is regarded as a skin sensitizer. This triggers classification and labelling in the EU (OECD 429/EU B42).

Local Lymph Node Assay (v)
Data Analysis

Example: Cpd X!

Test item concentration % (w/v)		Measurement dpm	Calculation			Result
			dpm - BG[a)]	number of lymph nodes	dpm per lymph node[b)]	S.I.
--	BG I	0	--	--	--	--
--	BG II	0	--	--	--	--
--	CG 1	3885	3885	8	486	--
1	TG 2	5575	5575	8	697	**1.4**
5	TG 3	9349	9349	8	1169	**2.4**
10	TG 4	26859	26859	8	3357	**6.9**
25	TG 5	49695	49695	8	6212	**12.8**
50	TG 6	62635	62635	8	7829	**16.1**

BG = Background (1 ml 5 % trichloroacetic acid) in duplicate

CG = Control Group

TG = Test Group

S.I. = Stimulation Index

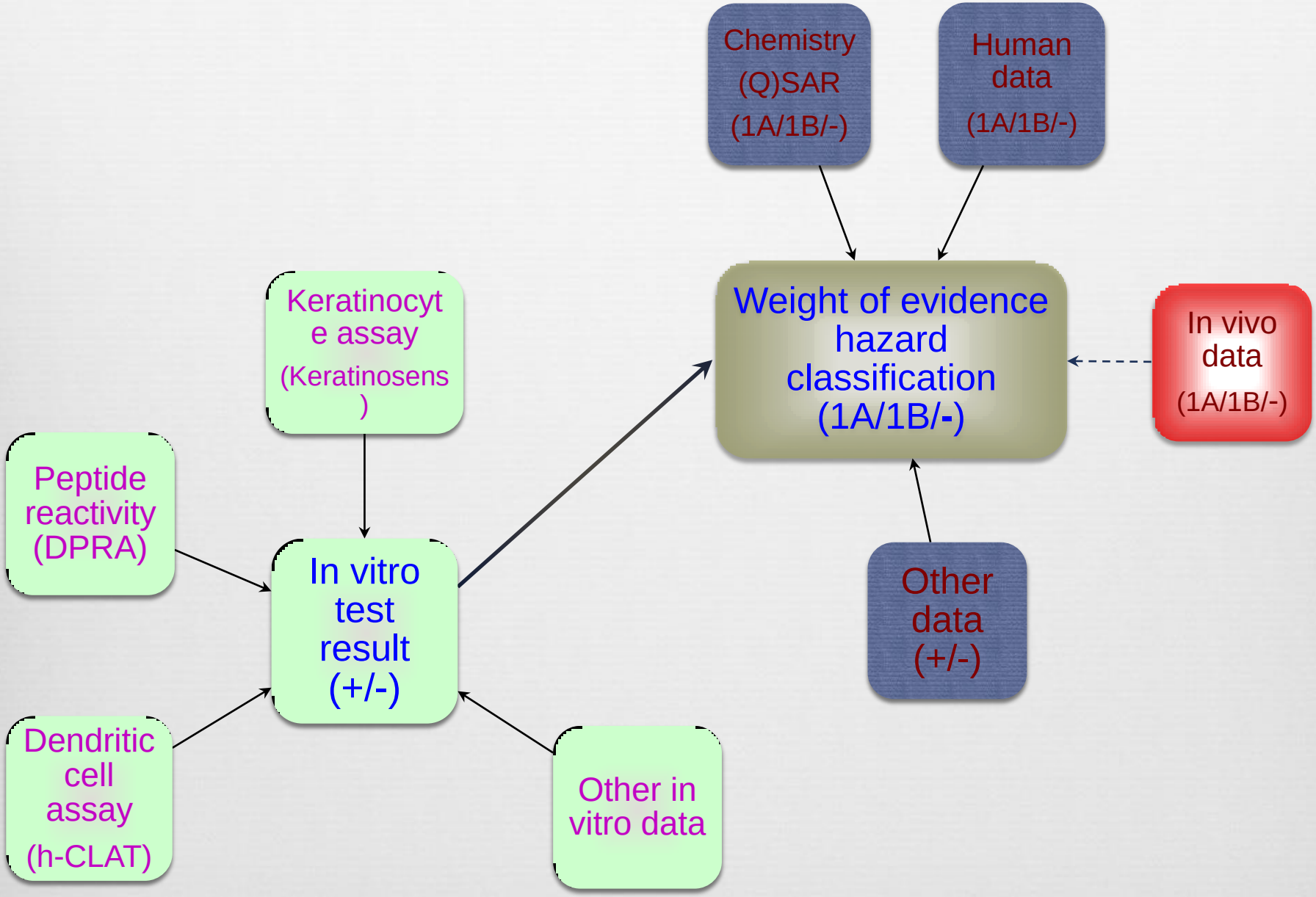
a) = The mean value was taken from the figures BG I and BG II

b) = Since the lymph nodes of the animals of a dose group were pooled, DPM/node was determined by dividing the measured value by the number of lymph nodes pooled

LLNA and HCA Control

	5%	10%	25%
1	2.1	3.3	8.4
2	1.5	4.4	8.8
3	1.1	2.5	10.4
4	2.1	4.4	8.1
5	2.2	2.8	8.2
6	2.1	2.4	7.2
7	1.6	2.5	6.8
8	2.1	2.7	7.8
9	1.4	2.7	5.3
10	1.4	2.0	8.7

- The table shows HCA data from repeated tests in three laboratories.
- Results are very concordant (as are derived EC3 values).
- The vehicle was AOO, with dpm/node values ranging from 159 – 495.


Chemistry (Q)SAR (1A/1B/-)
Human data (1A/1B/-)
Keratinocyte assay (Keratinosens)
Weight of evidence hazard classification (1A/1B/-)
In vivo data (1A/1B/-)
Peptide reactivity (DPRA)
In vitro test result (+/-)
Other data (+/-)
Dendritic cell assay (h-CLAT)
Other in vitro data

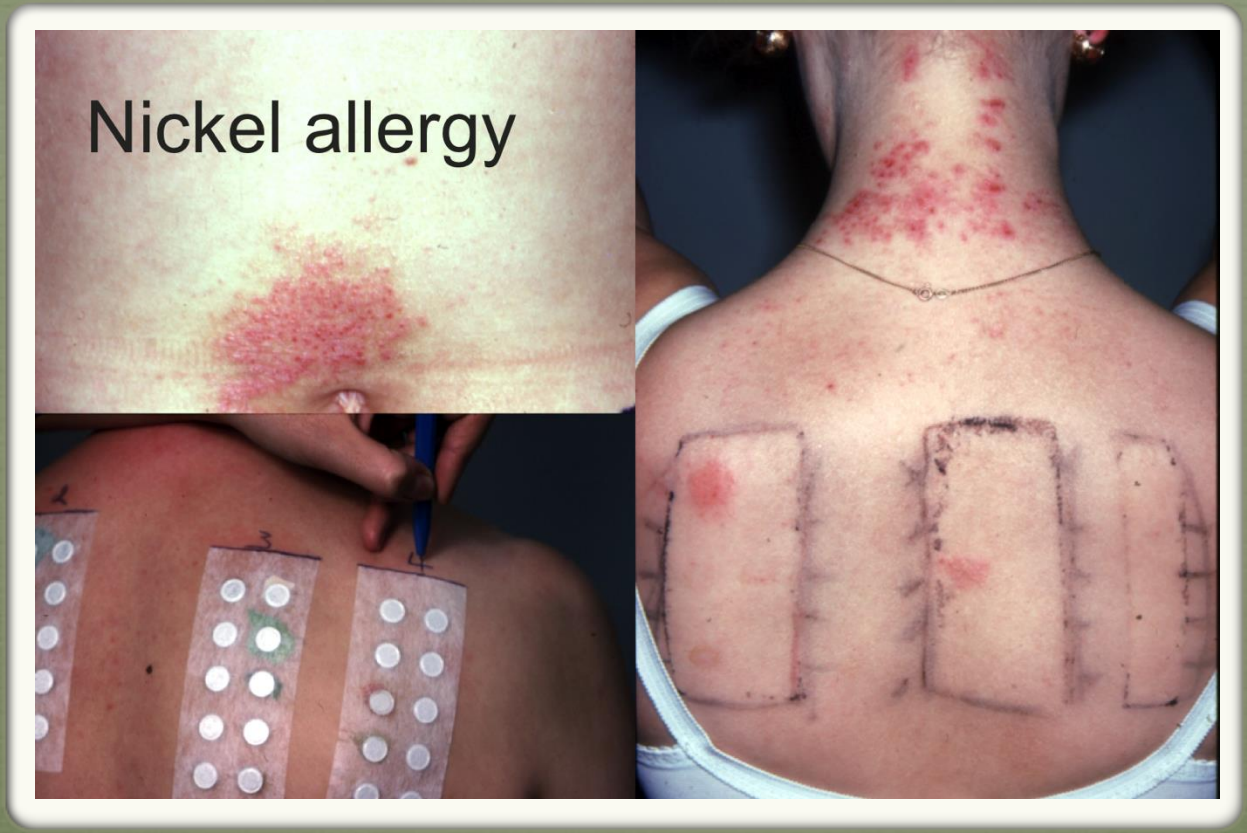
Under what circumstances is human data relevant?

When it gives the right answer!

Considerations...

- Evidence of absence is generally more useful than an absence of evidence
- Positive results from multiple clinics must override negative in vivo and in vitro tests
- Diagnostic patch testing in multiple clinics for months may indicate no sensitisation
- Absence of evidence of skin sensitisation can only be compelling if:
  - there is more than a HRIPT
  - there is extensive dermal exposure for years in many
  - there is (almost) no clinical report of skin allergy
  - there is an understanding of what people were exposed to

All human data should be subject to scrutiny for scientific credibility, just like any other.



human data types include clinical and experimental

Human experimental data

- For existing substances, there is a published body of work using the HMT (human maximization test) (n=87) & the (HRIPT) human repeated insult patch test (n=25?)
- These tests, carried out properly, have a defined level of sensitivity. The HMT compares favourably to the GPMT; the HRIPT is more like the Buehler test
- **HMT**: 5 x 48h in occluded exposures over 2 weeks to inflamed skin at a moderately irritating concentration
- 25 healthy volunteers
- **HRIPT**: 9 x 24/48h (semi-) occluded exposures over 3 wk with a mildly irritant dose
- 100-200 healthy volunteers

Clinical data: case histories

- Diagnostic patch testing is carried out weekly in hundreds of dermatology clinics around the world
- Collations of these results are published, in addition to specific investigations
- The information tells us which substances are inducing contact allergy (i.e. are human skin sensitisers), but often cannot identify the causative exposures
- Groups of patients are sometimes collected so that elicitation dose response work can be done and substance specific thresholds identified

Clinical data benchmarks

- Nickel: too obvious/unique
- MCI/MI and chromium - 1A
  - great examples of strong human skin sensitisers since they cause contact allergy in lots of people at low exposure
- MDGN: clinical evidence - 1A
  - originally in vivo negative, but clinical patch test positive
- HICC: clinical evidence for
- Hexyl cinnamal: clinically 1B
- EGDMA and resorcinol - 1B
  - examples of well known contact allergens with a fair degree of exposure, but only a modest amount of contact sensitisation observed
- Citral and imidazolidinyl urea may also be good examples

Clinical data benchmarks

- Some substances can be placed in the not classified category:
  - propylene glycol
  - benzalkonium chloride
  - isopropanol
  - isopropyl myristate

- All of these have extensive skin exposures, all have positive patch test results, but all at such low frequency they do not classify. None of them are non-sensitisers.

- Sodium lauryl sulphate, nitrogen & benzene are non-

Human data overrides in vivo

- Nickel – negative in vivo, but human data shows it is positive
- Quaternium 15, MDGN – negative in vivo, but human data overrides
- Sodium lauryl sulphate – positive in vivo, but human data proves it is negative, not classifiable
- Isopropyl myristate, xylene – positive in vivo, but human data proves they do not classify

It is a basic scientific, toxicological and regulatory error to argue that negative human data cannot override positive in vivo/in vitro data. For sensitisation, as for other toxicology endpoints, it has already done so, it is doing so and will continue to do so. Therefore, what is vital is that we agree standards and benchmarks, both for positive and negative decision making!

Human data can help to make positive (incl sub-categorisation) **and** negative classification decisions, but only if there is known exposure.

Summary

- In vivo methods aim to identify the intrinsic property of a chemical in respect of the skin sensitisation endpoint
- With imperfections, they have done this well for decades
- In vitro methods are set to supplant hazard identification
- Human (and other) data can be used to refine decisions
- Fragrance substances are simply one part of the broad spectrum of chemicals which possess skin sensitising properties
- In vivo and in vitro methods have been extensively evaluated with fragrance chemicals and thus represent perhaps the most reliable area (AD) in which the test methods operate


Thank you for your attention