How are the variability and knowledge gaps addressed in the LLNA and current DAs – what could be done and learnt for quantifying and managing uncertainty?

Roman LISKA

The European Commission's science and knowledge service

Joint Research Centre

Skin sensitisation – overall picture

single methods

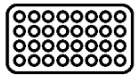
h-CLAT
KeratinoSens
DPRA
others…

Defined Approches

2of3
STS
others…

European Commission

Skin sensitisation – ideal world

Chemical	 Human	 LLNA	 in vitro
1	S	S	NS
2	NS	NS	NS
3	S	NS	S
4	S	S	S
.			
.			
.			
n-1	NS	S	NS
n	S	S	NS

Relevance (sensitivity, specificity, accuracy)

European Commission

Skin sensitisation – reality

Chemical	Human	LLNA	in vitro
1	S	{S,S,S,NS}	{S,S,NS,NS,S,S}
2	NS	{NS,NS}	{NS,NS,NS,NS,NS,NS}
3	?	{NS}	{S,S,NS,S,S,S}
4	S	{Ø}	{S,S,S,S,S,S}
.			
.			
.			
n-1	NS ?	{S,NS,NS}	{NS,NS,S,S,NS,S}
n	S	{NS,NS}	{S,S,S,NS,S,S,NS}

Relevance and Reliability

European Commission

Reliability

Within Laboratory reproducibility

LAB

	run 1	run 2	run 3	concordant
TI 1	S	S	S	1
TI 2	S	S	NS	0
TI 3	NS	NS	NS	1
TI N	S	NS	S	0
				85%

average

Between Laboratory reproducibility

	LAB 1			LAB 2			LAB 3		
	run 1	run 2	run 3	run 1	run 2	run 3	run 1	run 2	run 3
TI 1	S	S	S	S	S	S	S	S	S
TI 2	S	S	NS	S	S	S	S	S	NS
TI 3	NS	NS	NS	NS	NS	NS	NS	NS	NS
TI N	NS	NS	S	NS	NS	NS	S	NS	NS

Majority rule

TI 1	S	S	S
TI 2	S	S	S
TI 3	NS	NS	NS
TI N	NS	NS	NS

concordant
1
1
1
1
95%

average

European Commission

Skin sensitisation – reality

Aggregation – majority rule

Chemical	Human	LLNA	in vitro
1	S	{S,S,S,NS} → S	{S,S,NS,NS,S,S} → S
2	NS	{NS,NS} → NS	{NS,NS,NS,NS,NS,NS} → NS
3	?	{NS} → NS	{S,S,NS,S,S,S} → S
4	S	{Ø}	{S,S,S,S,S,S} → S
.			
.			
.			
n-1	NS ?	{S,NS,NS} → NS	{NS,NS,S,S,NS,S} → ?
n	S	{NS,NS} → NS	{S,S,S,NS,S,S,NS} → S

Relevance and Reliability

European Commission

Relevance

Predictive capacity

- Sensitivity
- Specificity
- Accuracy

LAB

True classification		run 1	run 2	run 3	majority rule	
TN	TI 1	NS	NS	S	NS	Specificity = % of negative classifications
TN	TI 2	S	S	NS	S	
	:					
	:					
TN	TI N_1	NS	NS	S	NS	
TP	TI N_1+1	S	S	S	NS	Sensitivity = % of positive classifications
TP	TI N_1+2	S	S	NS	S	
	:					
	:					
TP	TI N_1+N_2	NS	NS	S	NS	

(2) The EURL ECVAM study showed that the DPRA is transferable to suitably equipped laboratories that are proficient in high performance liquid chromatography (HPLC) analysis and the results obtained demonstrated within- and between-laboratory reproducibility of 87% and 75%, respectively.

(3) Full evaluation of the predictive capacity and applicability domain of the DPRA were outside the scope of the EURL ECVAM study. However, based on the study results and excluding metal compounds for which the test is not applicable, the accuracy of the DPRA for distinguishing sensitisers from non-sensitisers was 82% (sensitivity of 76%, specificity of 92%) which is in agreement with published information from previous studies (Gerberick *et al.*, 2007; Bauch et al., 2012; Natsch et al., 2013).

Source: ECVAM Validation report - DPRA

(4) The Givaudan-coordinated validation study generated preliminary information on the test method's predictive capacity and it was found that the accuracy of the KeratinoSens™ to discriminate skin sensitisers from non-sensitisers was 90% (sensitivity 87%, specificity 100%; n=21)[1]. The accuracy calculated for an additional set of chemicals (77 sensitisers and 104 non-sensitisers) tested in-house by Givaudan was 75%. These figures are similar to those recently published by Natsch et al. (2013) based on in-house testing of about 145 chemicals (77% accuracy, 79% sensitivity, 72% specificity). Taken together, this information indicates the usefulness of the KeratinoSens™ assay to contribute to the identification of sensitisers and non-sensitisers.

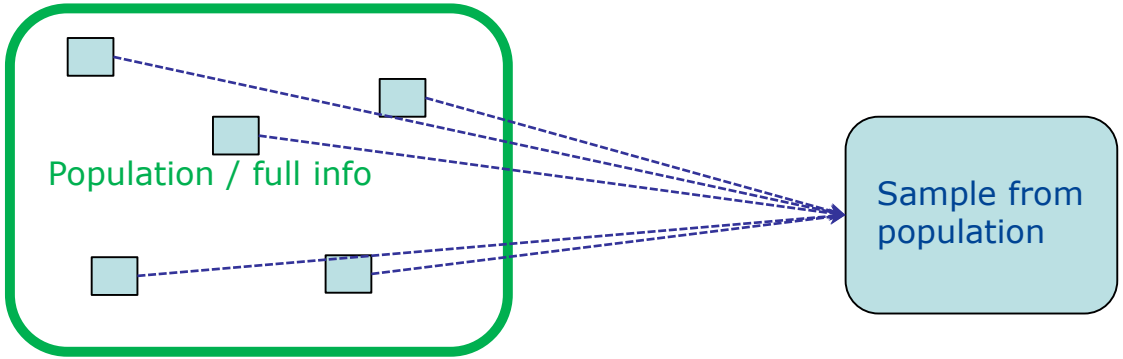
Source: ECVAM Validation report - KeratinoSens

Reliability and Relevance evaluations

- Do not capture the variability at lower levels (loss of information)
- Often based on univariate descriptive measures (not estimates)
- In context of validation
  - Not clear how to "aggregate" information for e.g. BLR evaluation
  - Do not guarantee that BLR ≤ WLR measure

Need for a new approach ….

Important concept in statistics



Full information,
Parameters of interest can be fully known

Limited information,
Estimate parameters
given the sample from population

Full information – example

True classification		run 1	run 2	run 3
TN	TI 1	0	0	1
TN	TI 2	1	1	0
	:		:	
TN	TI N_1	0	0	1
TP	TI N_1+1	1	1	1
TP	TI N_1+2	1	1	0
	:		:	
TP	TI N_1+N_2	0	0	1

European Commission

Alternative performance assessment

- Reproducibility measures (WLR, BLR, GR)

  Based on probability(m out of N chemicals have the same prediction in 3 independent runs), m=0, 1, …, N

- Predictive capacity measures(Sensitivity, Specificity)

  Based on probability(m out of N chemicals has correctly predicted the true outcome) , m=0, 1, …, N

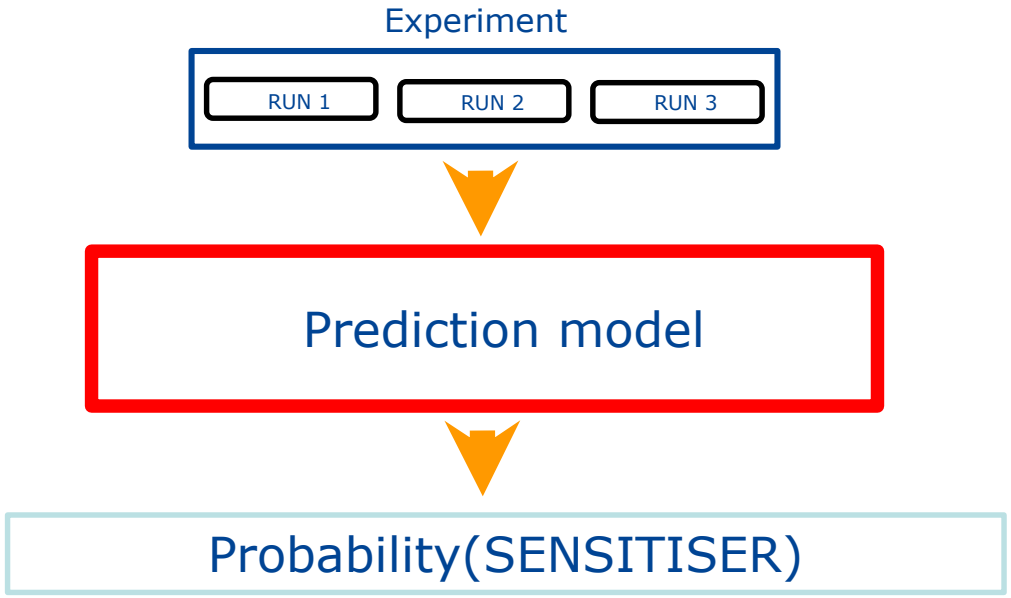
Illustrative Example

Example: h-Clat data

		Chemical	Reference classification		Kao								
				LLNA	Experiment 1			Experiment 2			Experiment 3		
					Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
Group 1	1	Benzoquinone	SENSITISER	POSITIVE	P12	P2	P12						
	2	PPD	SENSITISER	POSITIVE	P1	P1	P12						
	3	Dihydroeugenol	SENSITISER	POSITIVE	P12	P12	N						
	4	Thioglycerol	SENSITISER	POSITIVE	P1	P1	P12						
	5	Imidazolidinylurea	SENSITISER	POSITIVE	P12	P12	P12						
	6	Methylmethacrylate	SENSITISER	POSITIVE	N	N	N						
	7	Glycerol	NON SENSITISER	NEGATIVE	N	N	N						
	8	2,4-Dichloronitrobenzene	NON SENSITISER	NEGATIVE	P12	P12	P12						
	9	Benzyl alcohol	NON SENSITISER	NEGATIVE	P1	N	P1						
Group 2	10	Kathon CG (CMI/MI)	SENSITISER	POSITIVE	P2	P12	P2	P2	P2	P1	P1	P12	P2
	11	Beryllium sulfate	SENSITISER	POSITIVE	N	N	N	N	N	N	N	N	N
	12	Formaldehyde	SENSITISER	POSITIVE	P12	P12	P12	P12	P12	P12	P1	P1	P12
	13	Chloramine T	SENSITISER	POSITIVE	P12	P2	P12	P2	P12	P12	P12	P2	P2
	14	Chlorpromazine HCl	SENSITISER	POSITIVE	P12	P12	P12	P2	P2	P2	P2	P1	P12
	15	2-Mercaptobenzothiazole	SENSITISER	POSITIVE	P2	P2	P2	P2	P2	P2	P12	P2	P2
	16	Benzyl salicylate	SENSITISER	POSITIVE	N	N	N	P2	P2	N	N	P2	N
	17	Benzyl cinnamate	SENSITISER	POSITIVE	N	N	N	N	N	N	N	N	N
	18	R(+)- Limonene	SENSITISER	POSITIVE	P12	P12	P2	P12	P2	P12	P2	P2	P2
	19	Methyl salicylate	NON SENSITISER	NEGATIVE	P1	P1	N	N	N	N	N	P2	P12
	20	Isopropanol	NON SENSITISER	NEGATIVE	N	N	N	N	N	N	N	N	N
	21	Dimethyl isophtalate	NON SENSITISER	NEGATIVE	N	P2	N	N	N	N	N	P2	N
	22	4-Aminobenzoic acid	NON SENSITISER	NEGATIVE	N	N	N	P12	N	N	N	N	N
	23	Nickel chloride	SENSITISER	NEGATIVE	P12	P12	P12	P12	P2	P12	P12	P12	P12
	24	Xylene	NON SENSITISER	POSITIVE	N	P2	N	N	N	N	N	N	N

Example: h-Clat

one test item



Example: h-Clat data

HUMAN	Test item #	S	NS
1	1	5	7
1	2	11	1
1	3	4	0
1	4	4	0
1	5	11	1
1	6	12	0
1	7	10	2
1	8	12	0
1	9	5	7
1	10	4	0
1	11	4	0
1	12	12	0
1	13	4	8
1	14	4	0
1	15	12	0
1	16	0	4
0	17	0	4
0	18	4	0
0	19	4	0
0	20	6	6
0	21	0	12
0	22	1	11
0	23	0	12
0	24	4	8

Example: h-Clat

	average	Kao	Shiseido	Bioassay	EURL ECVAM
h-CLAT report WLR	80.0%	86.7%	80.0%	73.3%	80.0%
WLR_0	84.8%	88.6%	82.4%	85.5%	82.7%
$\widehat{WLR}_1$	83.5%	88.2%	81.4%	81.6%	82.9%
$\widehat{WLR}_{1BC}$	82.4%	88.0%	79.2%	81.4%	80.8%
95% bootstrap CI for WLR_1		(82.1%, 94.1%)	(71.4%, 87.6%)	(74.1%, 89.4%)	(74.4%, 87.6%)

(a) WLR estimates

	average	Kao Shiseido Bioassay	Kao Shiseido EURL ECVAM	Kao Bioassay EURL ECVAM	Shiseido Bioassay EURL ECVAM
h-CLAT report BLR	82.3%	87.5%	79.2%	83.3%	79.2%
BLR_0	82.4%	91.1%	78.4%	81.7%	78.4%
$\widehat{BLR}_1$	75.2%	81.1%	71.5%	75.8%	72.4%
$\widehat{BLR}_{1BC}$	76.9%	84.0%	73.1%	77.1%	73.6%
95% bootstrap CI for BLR_1		(78.8%, 89.2%)	(68.4%, 78.0%)	(72.2%,82.7%)	(68.8%, 78.9%)

(b) BLR estimates

Table 1: h-CLAT. WLR and BLR estimates and its 95% confidence intervals(CI). Notation: $\widehat{WLR}_{1BC}$ is the bootstrap corrected estimate of WLR_1.(similarly for $\widehat{BLR}_{1BC}$)

Example: h-Clat

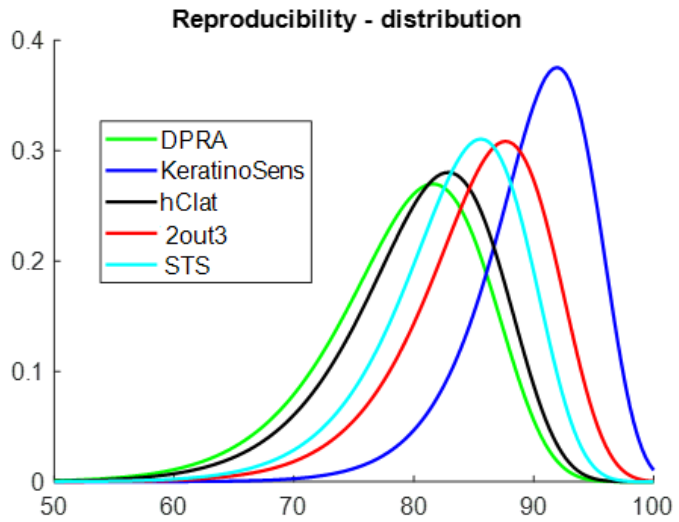
Probability(SENSITISER)

	Chemical	Reference classification	$\hat{q}(c)$ = est. P(chemical c is classified as P)			
			Kao	Shiseido	Bioassay	ECVAM
Group 1						
1	Benzoquinone	SENSITISER	1	1	1	1
2	PPD	SENSITISER	1	1	1	1
3	Dihydroeugenol	SENSITISER	0.74	0.74	1	1
4	Thioglycerol	SENSITISER	1	1	1	0.74
5	Imidazolidinylurea	SENSITISER	1	1	1	1
6	Methylmethacrylate	SENSITISER	0	0	0	0
7	Glycerol	NON SENSITISER	0	0	0	0
8	2,4-Dichloronitrobenzene	NON SENSITISER	1	1	1	1
9	Benzyl alcohol	NON SENSITISER	0.74	0.74	1	1
Group 2						
10	Kathon CG (CMI/MI)	SENSITISER	1	0.92	0.97	0.97
11	Beryllium sulfate	SENSITISER	0	0.58	0.03	0.79
12	Formaldehyde	SENSITISER	1	0.87	0.87	1
13	Chloramine T	SENSITISER	1	0.94	1	1
14	Chlorpromazine HCl	SENSITISER	1	1	0.97	0.87
15	2-Mercaptobenzothiazole	SENSITISER	1	1	1	1
16	Benzyl salicylate	SENSITISER	0.26	0.13	0.16	1
17	Benzyl cinnamate	SENSITISER	0	0.79	0	0.42
18	R(+)- Limonene	SENSITISER	1	1	1	1
19	Methyl salicylate	NON SENSITISER	0.33	0.13	0.25	0.97
20	Isopropanol	NON SENSITISER	0	0.03	0.03	0
21	Dimethyl isophtalate	NON SENSITISER	0.13	0	0.13	0.39
22	4-Aminobenzoic acid	NON SENSITISER	0.03	0	0	0.03
23	Nickel chloride	SENSITISER	1	1	1	1
24	Xylene	NON SENSITISER	0.03	0.03	0.42	1

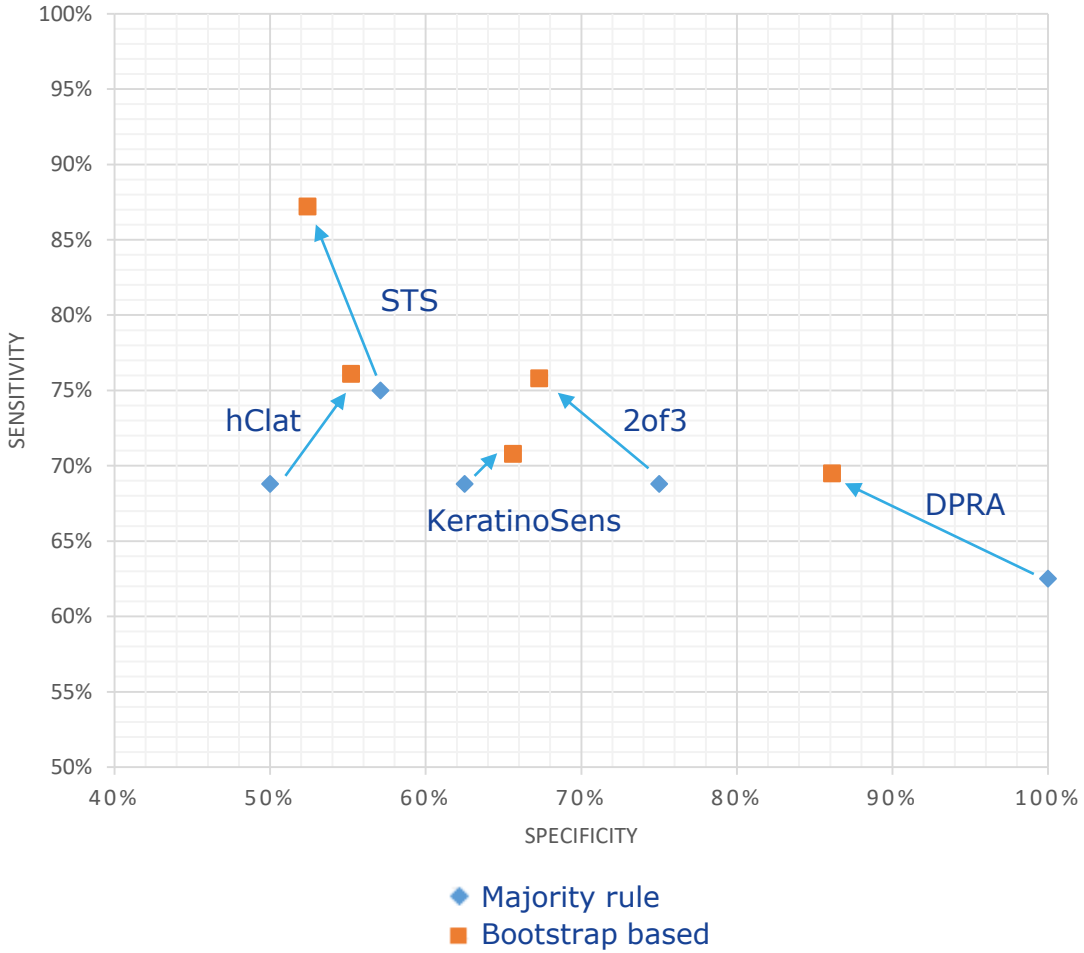
Example: DA skin sensitisation

Performance measures based on 100.000 Bootstrap replicates

		vs. LLNA		
	reproducibility	accuracy	specificity	sensitivity
DPRA	79.3%	75.0%	86.1%	69.5%
KeratinoSens	90.2%	69.1%	65.6%	70.8%
hClat	80.6%	69.1%	55.2%	76.1%
2of3	85.6%	72.9%	67.3%	75.8%
STS	83.6%	75.6%	52.4%	87.2%



European Commission



	specificity	
	majority rule	bootstrap based
DPRA	100.0%	86.1%
KeratinoSens	62.5%	65.6%
hClat	50.0%	55.2%
2of3	75.0%	67.3%
STS	57.1%	52.4%

	sensitivity	
	majority rule	bootstrap based
DPRA	62.5%	69.5%
KeratinoSens	68.8%	70.8%
hClat	68.8%	76.1%
2of3	68.8%	75.8%
STS	75.0%	87.2%

European Commission

Summary

- **Need for better performance evaluations**
  - To compare methods and/or DA
  - Data variability characterization
  - Expert judgement plays still an important role
- **A possible way based on probability measures**
- **Some issues to be resolved**
  - Unbalanced data
  - Missing data (Bayesian approach?)
  - Complexity
- **Purpose/use of a method – decision making**

Stay in touch

EU Science Hub: ***ec.europa.eu/jrc***

Twitter: ***@EU_ScienceHub***

Facebook: ***EU Science Hub - Joint Research Centre***

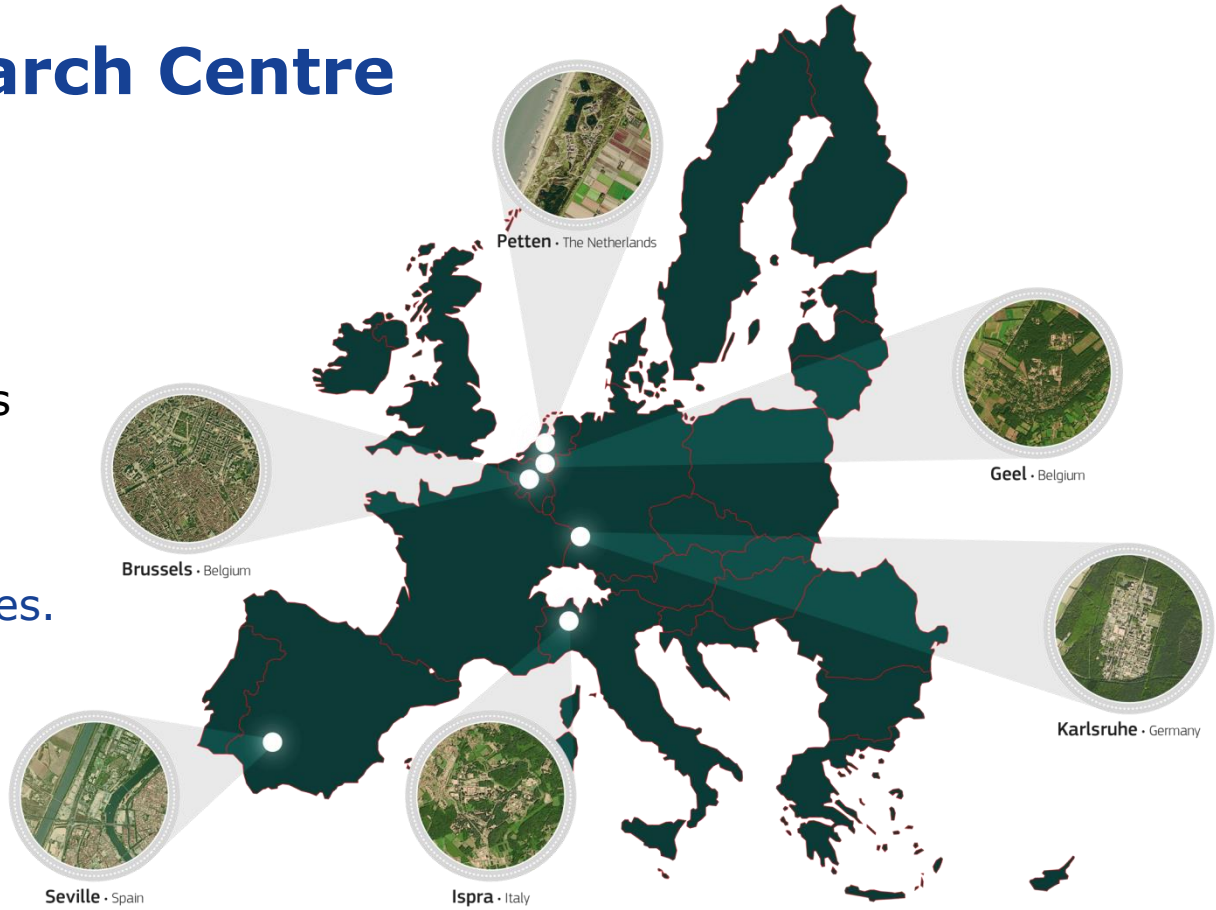
LinkedIn: ***Joint Research Centre***

YouTube: ***EU Science Hub***

European Commission

The Joint Research Centre at a glance

3000 staff
Almost 75% are scientists and researchers.
Headquarters in Brussels and research facilities located in 5 Member States.



European Commission

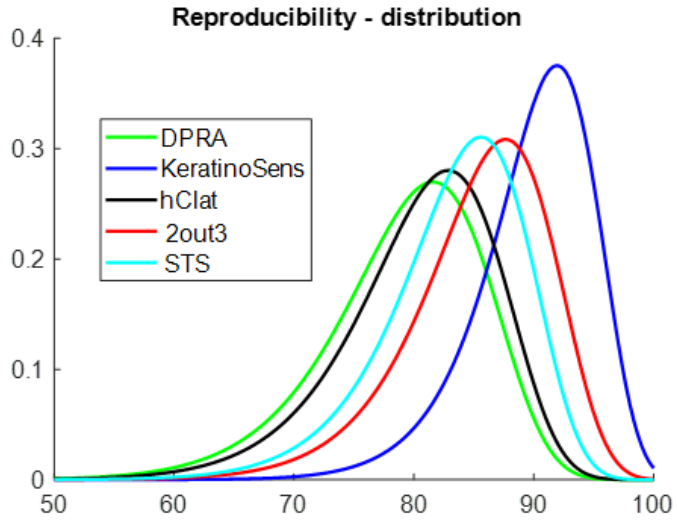
Example: DA skin sensitisation

majority rule of predictions

	vs. LLNA		
	accuracy	specificity	sensitivity
DPRA	75.0%	100.0%	62.5%
KeratinoSens	66.7%	62.5%	68.8%
hClat	62.5%	50.0%	68.8%
2of3	70.8%	75.0%	68.8%
STS	72.0%	57.1%	75.0%

Performance measures based on 100.000 Bootstrap replicates

		vs. LLNA		
	reproducibility	accuracy	specificity	sensitivity
DPRA	79.3%	75.0%	86.1%	69.5%
KeratinoSens	90.2%	69.1%	65.6%	70.8%
hClat	80.6%	69.1%	55.2%	76.1%
2of3	85.6%	72.9%	67.3%	75.8%
STS	83.6%	75.6%	52.4%	87.2%



European Commission