PERFORMANCE BASED VALIDATION APPROACHES AT OECD

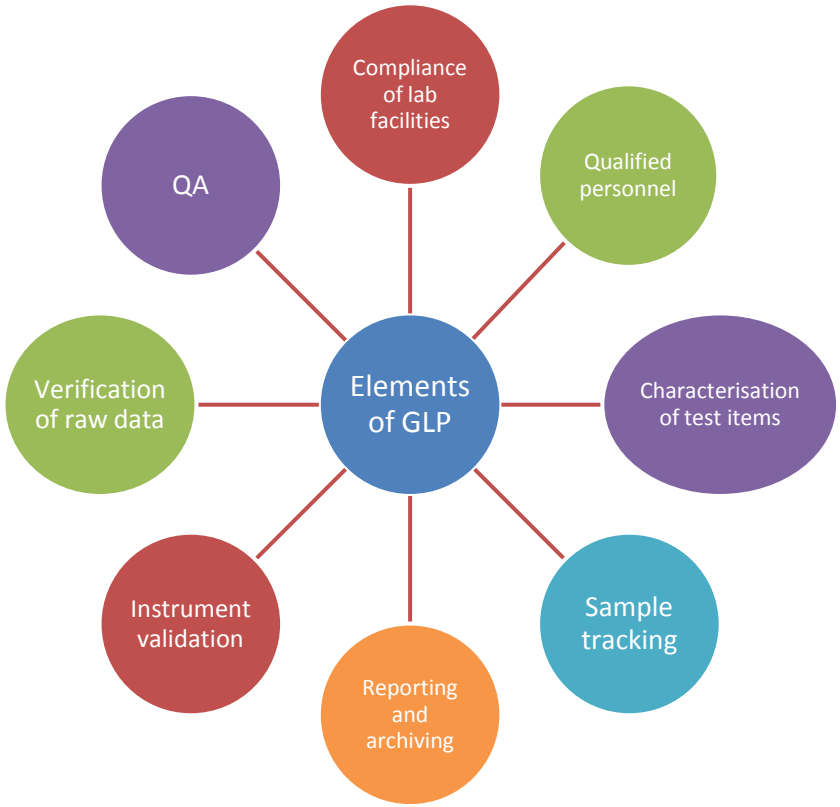
Patience Browne
Test Guidelines Programme, OECD
SOT IVAMSS October 25, 2017

OCDE
DES POLITIQUES MEILLEURES POUR UNE VIE MEILLEURE

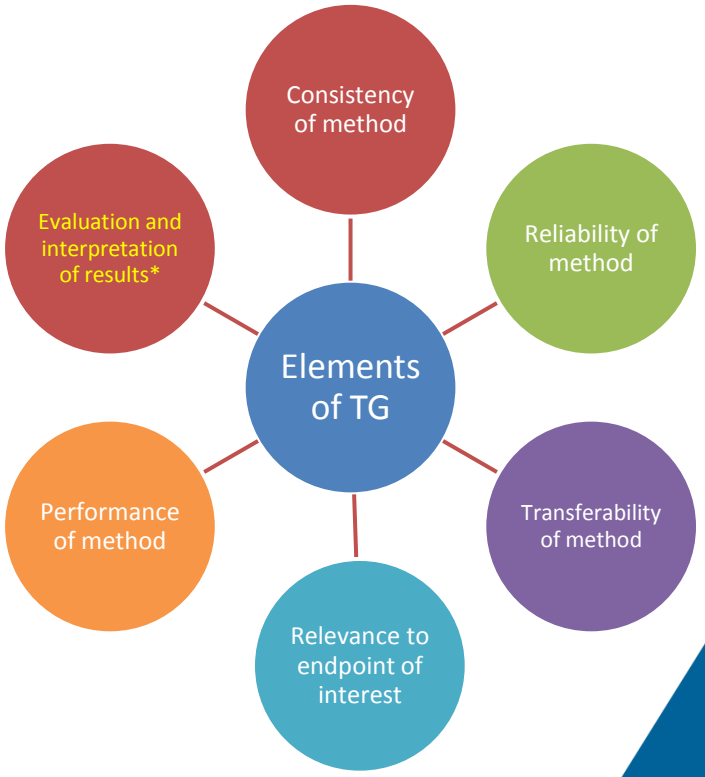
Why is Validation such a big deal at OECD?

- Harmonised international testing methods and principles of Good Laboratory Practices for chemical safety assessment to **ensure high quality data, reduce duplicative testing**, and **facilitate data** sharing across governments.

Good Laboratory Practice



Harmonised Test Guidelines



Mutual Acceptance of Data (MAD)

MAD is legally binding for OECD Member Countries

OECD GD 34: Modular Approach to Validation

- Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment (2005)
  - Following conference designed to reach consensus on guidance & principles for validating test methods for regulatory hazard assessment
- Modules include
  - test definition (including purpose, need and scientific basis);
  - intra-laboratory repeatability and reproducibility;
  - inter-laboratory transferability;
  - inter-laboratory reproducibility;
  - predictive capacity (accuracy);
  - applicability domain; and,
  - performance standards.

Validation using Performance Standards

- Initially intended to limit monopolies for test methods that included patents/intellectual property/protected elements and increase pace of validation
  - Such TG accepted at OECD, but with performance standards to which similar assays could be compared
  - Performance standards include
    - Essential test method components (controls, procedural details, etc.)
    - Reference chemicals
    - Measures of accuracy and reliability of method predict
  - Generally relevant to in vitro methods

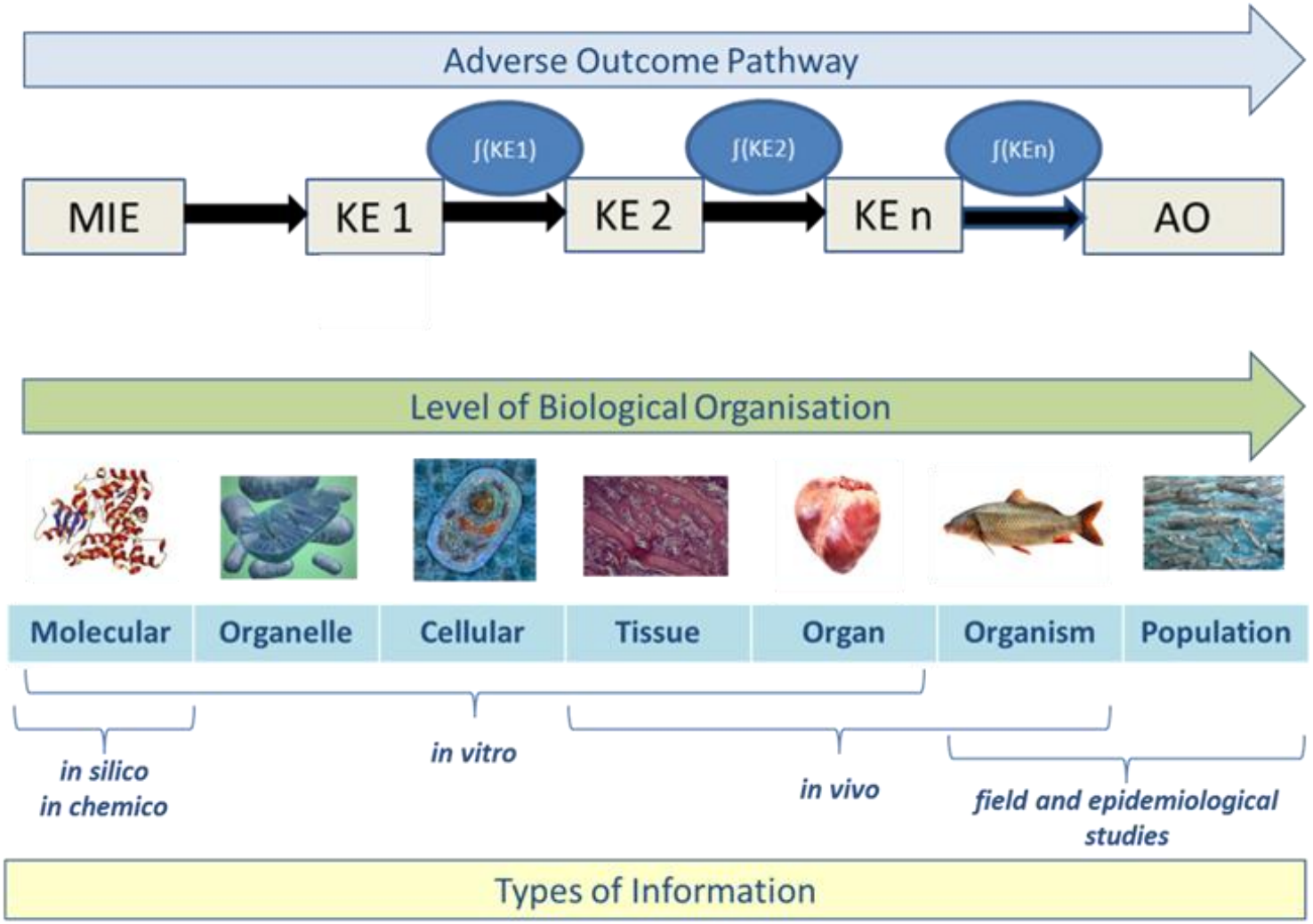
Performance-based validation approaches depend on Reference Chemicals

- Ideally, have reference chemicals data in:
  - Humans
  - High quality in vivo GL method
  - Proposed alternative method
- Ideally, 25-50+ robust, well-characterised reference chemicals including a range of potencies and structures
  - Easier for well-studied, common targets
  - Challenging for novel assays or lesser studied targets
- Traditionally, identified through exhaustive literature searches
  - May identify hundreds of chemical/study combinations
  - After vetting, generally only a handful that meet reference chemical criteria

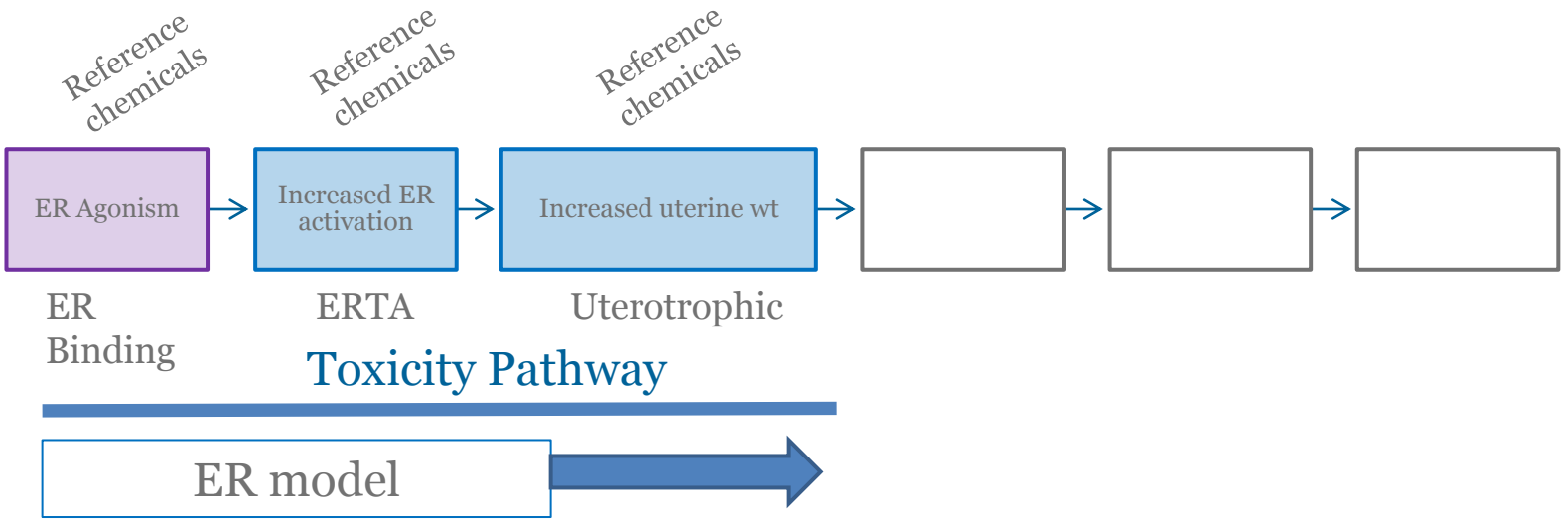
OECD current has PBTG for several endpoints

- Current OECD PBTGs for
  - Skin corrosion
  - Skin irritation
  - Estrogen receptor transactivation
  - Estrogen receptor binding
  - *More in validation process*
- PBTGs include methodologically and functionally assays

AOPs for more broad Performance-Based evaluations



e.g. US EPA performance-based validation of ER model



- 18 ER in vitro HTS assays included in ToxCast
  - Includes a variety of assay methods/technologies/functional endpoints
  - Judson et al. 2015 description of an integrated ER model
- Comparison of model performance against 65 well-characterised reference chemicals

e.g. US EPA performance-based validation of ER model

- 40 ***in vitro*** ER reference chemicals with independently confirmed activity (OECD 2012)

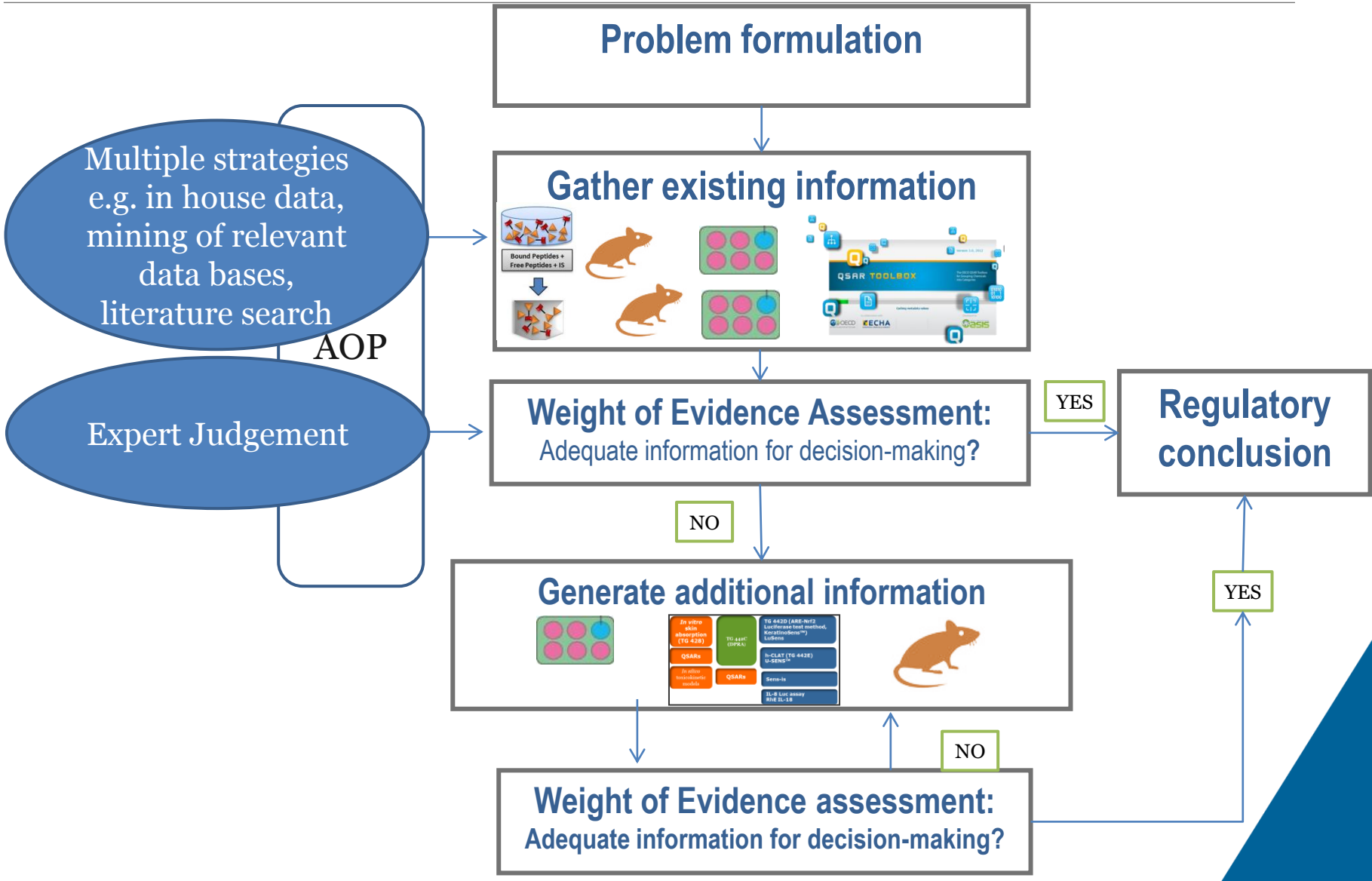
Performance Metrics	Value
# True Pos	26 (25)
# True Neg	11 (11)
# False Pos	1 (0)
# False Neg	2 (2)
Accuracy	0.93 (0.95)
Sensitivity	0.93 (0.93)
Specificity	0.92 (1.0)

- 43 ***in vivo*** ER reference chemicals with independently confirmed activity (Kleinstreuer et al. 2015)

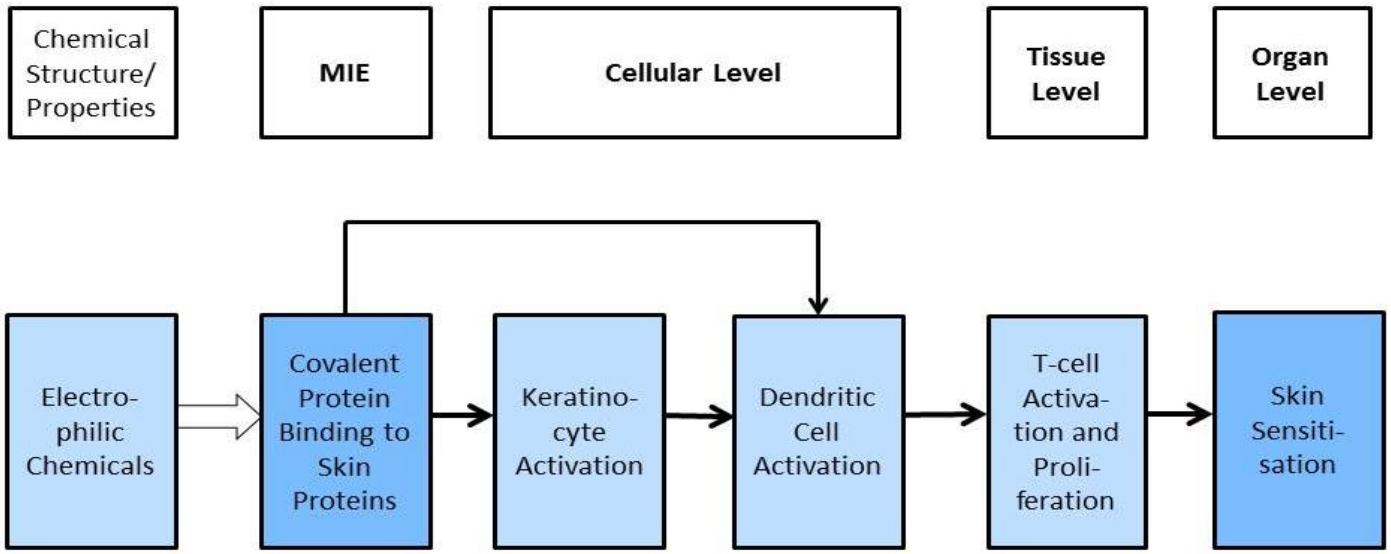
Performance Metrics	Value
# True Pos	29 (29)
# True Neg	8 (8)
# False Pos	5 (1)
# False Neg	1 (1)
Accuracy	0.86 (0.95)
Sensitivity	0.97 (0.97)
Specificity	0.67 (0.89)

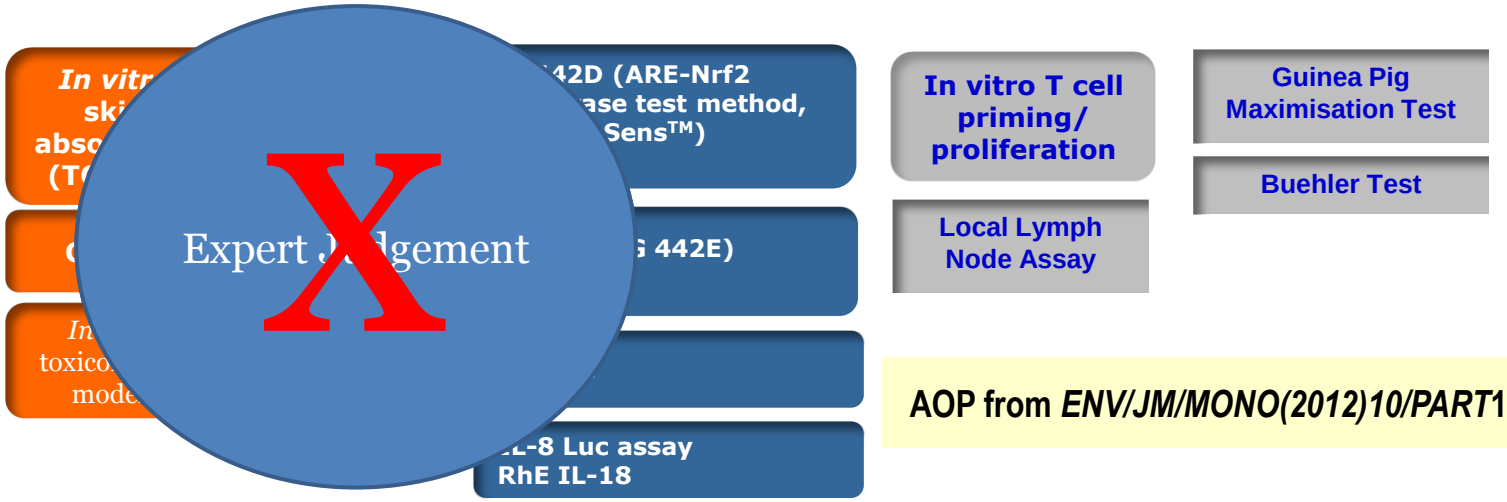
Details in Judson et al. 2015, Kleinstreuer et al. 2015, and Browne et al. 2015

Using AOPs to develop Integrated Approaches to Testing and Assessment (IATA)



Defined Approaches are tools developed based on AOPs: e.g. Skin sensitisation





AOP from *ENV/JM/MONO(2012)10/PART1*

https://aopwiki.org/wiki/index.php/Aop:40

12

Defined Approaches

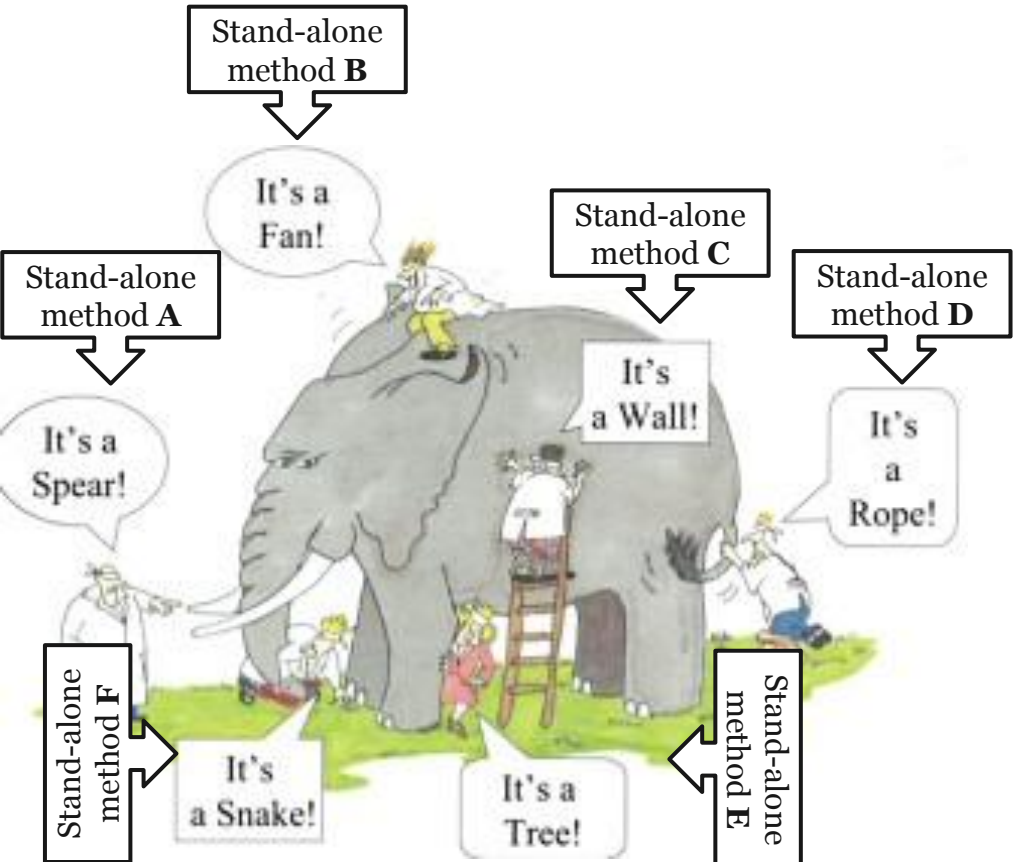
A **defined approach** to testing and assessment consists of a fixed data interpretation procedure (DIP) used to interpret data generated with a defined set of information sources, that can either be used alone, or together with other information sources, to satisfy a specific regulatory need.

- **Guidance Document** on the Reporting of Defined Approaches to be Used within Integrated Approaches to Testing and Assessment ENV/JM/MONO(2016)28

- **Guidance Document** on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation ENV/JM/MONO(2016)
  - Includes **12 skin sensitisation case studies**

Alternative methods used in combination

- Defined Approaches leverage the strengths of individual methods (e.g. some are better at some jobs than others)



Stand-alone methods A + E
or
Stand-alone methods A + C
or
Stand-alone methods C + D+ F

=

It's an elephant!

Challenges to Using Integrated Approaches for regulatory decisions

- Harmonisation
  - IATA are **flexible** and **not amenable**
  - Defined approaches are **not flexible** and **amenable**
- Communication
  - Clear understanding of the differences
- Validation
  - Description of test methods/information sources
  - Fixed interpretation
  - Performance; to what standard?
    - OECD language “TGP committed to using alternative methods when feasible”
    - New TSCA language “alternative methods when they can be demonstrated to perform as well as existing method”

What does international validation of a defined approach look like?

- Application
  - Fit-for-purpose
    - i.e. the solution must suit the regulatory need;
    - varies with regulatory authority/region
      - Activity
      - Potency
      - Classification and labelling
- Information sources
  - Pretty easy if all information sources are OECD in vitro test methods
  - More of a challenge for approaches that include QSARs and test methods that are not harmonised guidelines

General areas for Consideration

- What characteristics are needed to determine adequate scientific confidence (and capture the limitations) of defined approaches?
  - Performance, alone, will not be adequate for internationally harmonization.
- How might validation for assays intended to be used with other methods (i.e. defined approaches) differ from validation of a stand-alone assay ?
  - For example, some assays may be good for limited applicability domains, but do a poor job against “conventional” reference chemicals.
- Are standards for “alternative method” validation substantially higher than those for novel in vivo endpoints designed to predict the same outcome?
  - Most in vivo assays were validated with 3-7 chemicals
  - It this appropriate/prohibitive?
  - How can we separate aspects that are biologically based (e.g. how to account for metabolism in vitro) from a priori perception that in vivo methods are “better”?

Path Forward

- There is agreement these novel approaches (such as DAs) are the future and reflect the broadening scientific understanding
  - How to get international buy-in and acceptance is an active area of discussion at OECD
  - How do we ensure approaches are compatible with MAD
    - Do we need a new concept of validation?
    - Do we need new tiers of validation?
    - What are we validating against?

Thank you! Questions?