Evolving validation to better address regulatory use

João Barroso

Joint Research Centre, EURL ECVAM

NAMs workshop, 31 May-1 June 2023

The definition and principles of validation OECD Guidance Document 34

Unclassified **ENV/JM/MONO(2005)14**

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

18-Aug-2005

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

OECD SERIES ON TESTING AND ASSESSMENT
Number 34

GUIDANCE DOCUMENT ON THE VALIDATION AND INTERNATIONAL ACCEPTANCE OF NEW OR UPDATED TEST METHODS FOR HAZARD ASSESSMENT

The principles and process of validation

- PRINCIPLES are universal and valid
- PROCESS for validation and international acceptance described in GD34 no longer reflects current state-of-the art
- Revision needed to encourage timely uptake of NAMs!

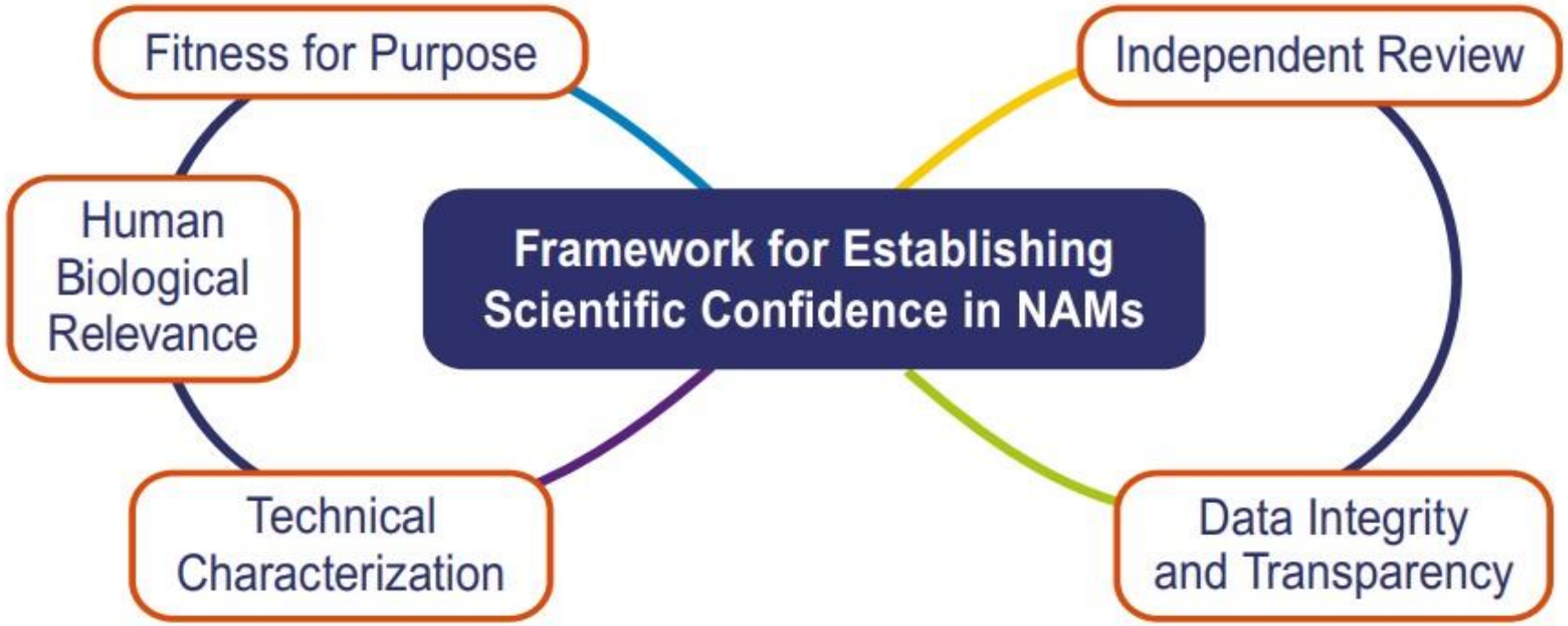
Archives of Toxicology
https://doi.org/10.1007/s00204-022-03365-4

August 2022

REVIEW ARTICLE

A framework for establishing scientific confidence in new approach methodologies

Anna J. van der Zalm[1] · João Barroso[2] · Patience Browne[3] · Warren Casey[4] · John Gordon[5] · Tala R. Henry[6] · Nicole C. Kleinstreuer[7] · Anna B. Lowit[6] · Monique Perron[8] · Amy J. Clippinger[1]



Data integrity and transparency

- According to OECD GD 34, validation studies should follow the principles of GLP
- Mostly not done in the past but not a problem because studies were coordinated by independent parties
- Now managed by commercial parties
- Important to demonstrate the integrity and credibility of the results, from the raw data through to the final report

Independent scientific review

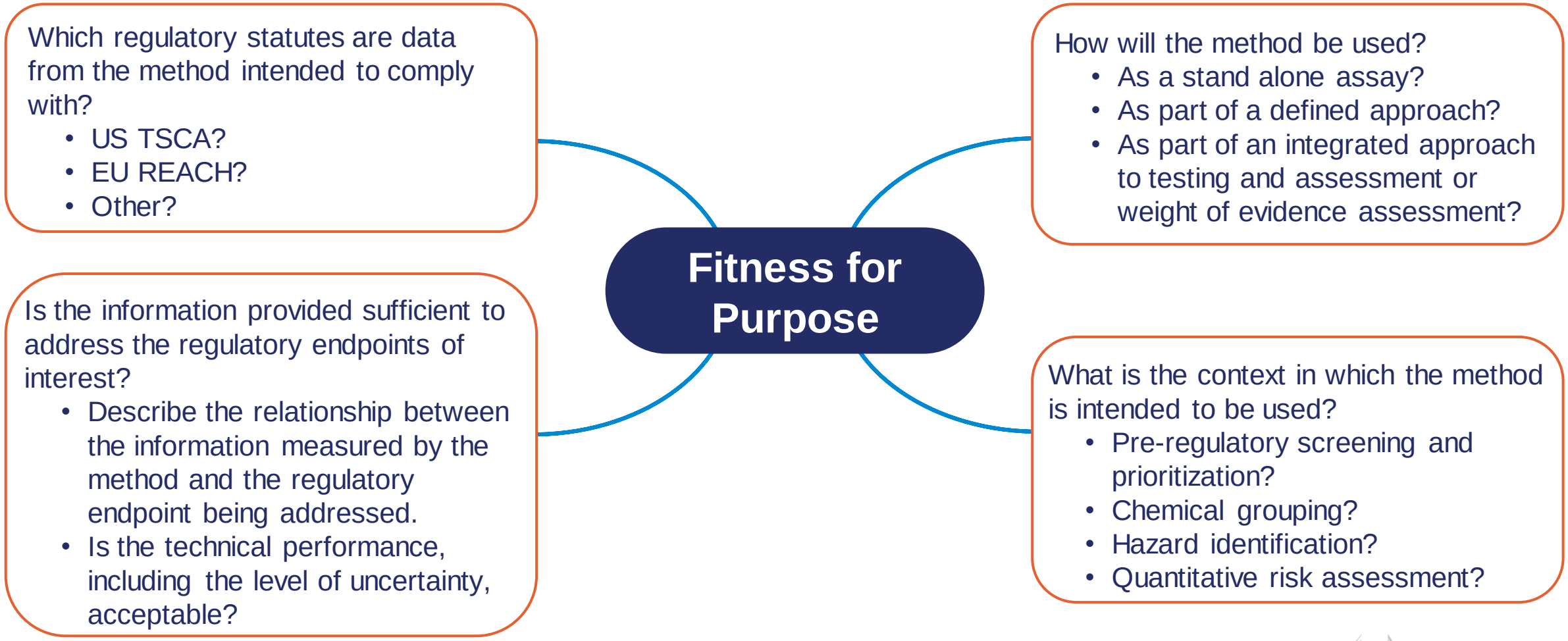
- Appropriate level of external review depends on the NAM and its intended use
- Might include publication in peer-reviewed journal or review by an independent scientific advisory panel
- International adoption by OECD typically needs formal peer review
- NAM developers may fund but should not manage peer review

(Human) biological relevance

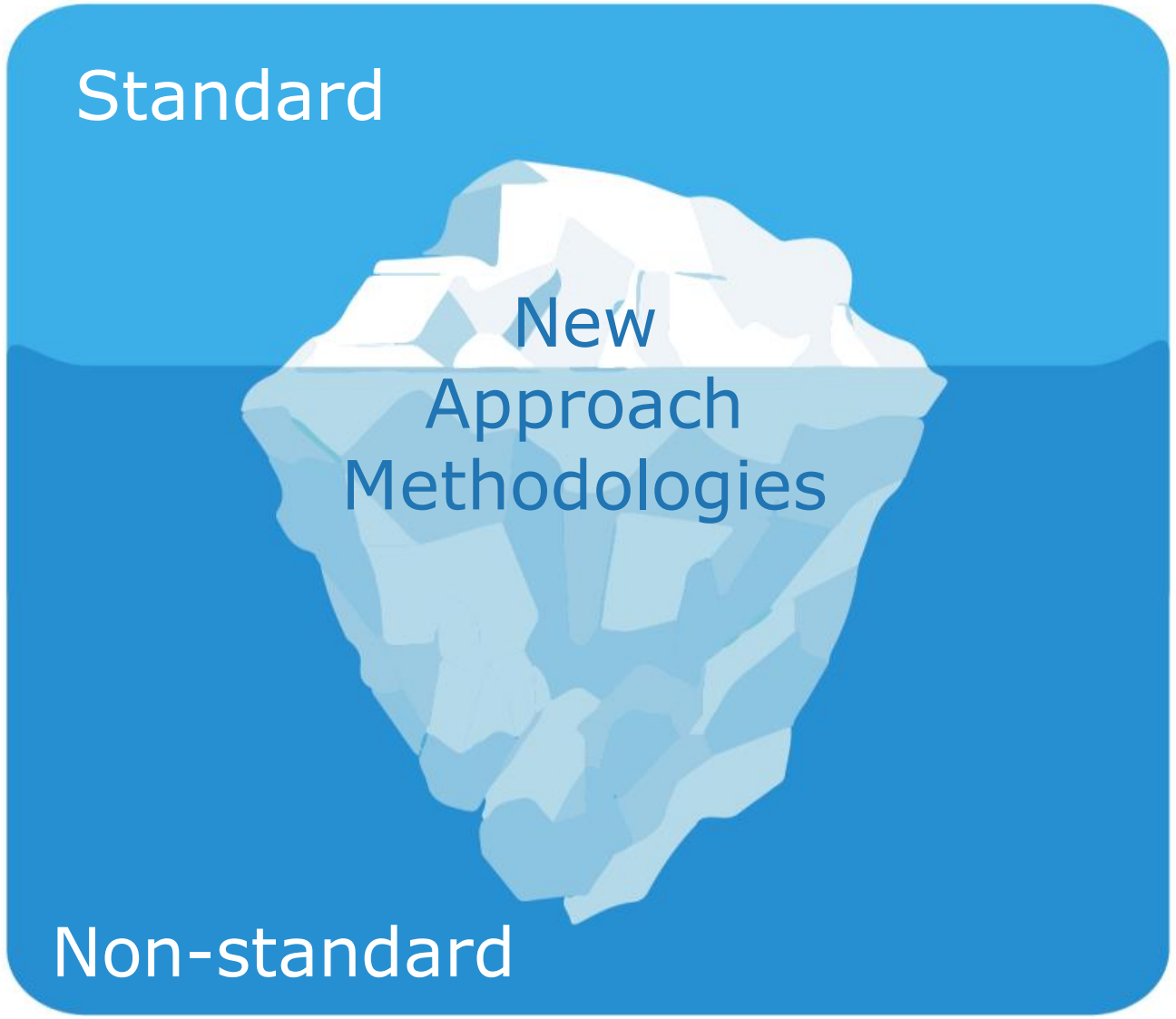
- Similarities between the physiology of, or the biology measured by, the test system, and human biology
- Concordance with human responses

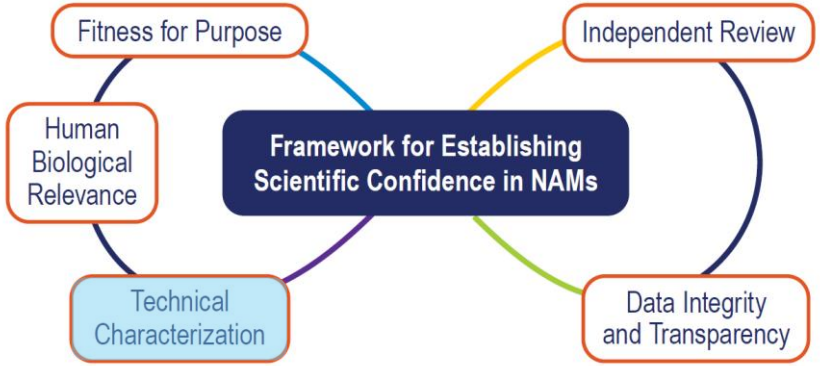
➢ **Establishing biological relevance of a NAM can be used to benchmark its performance**

Fitness for purpose



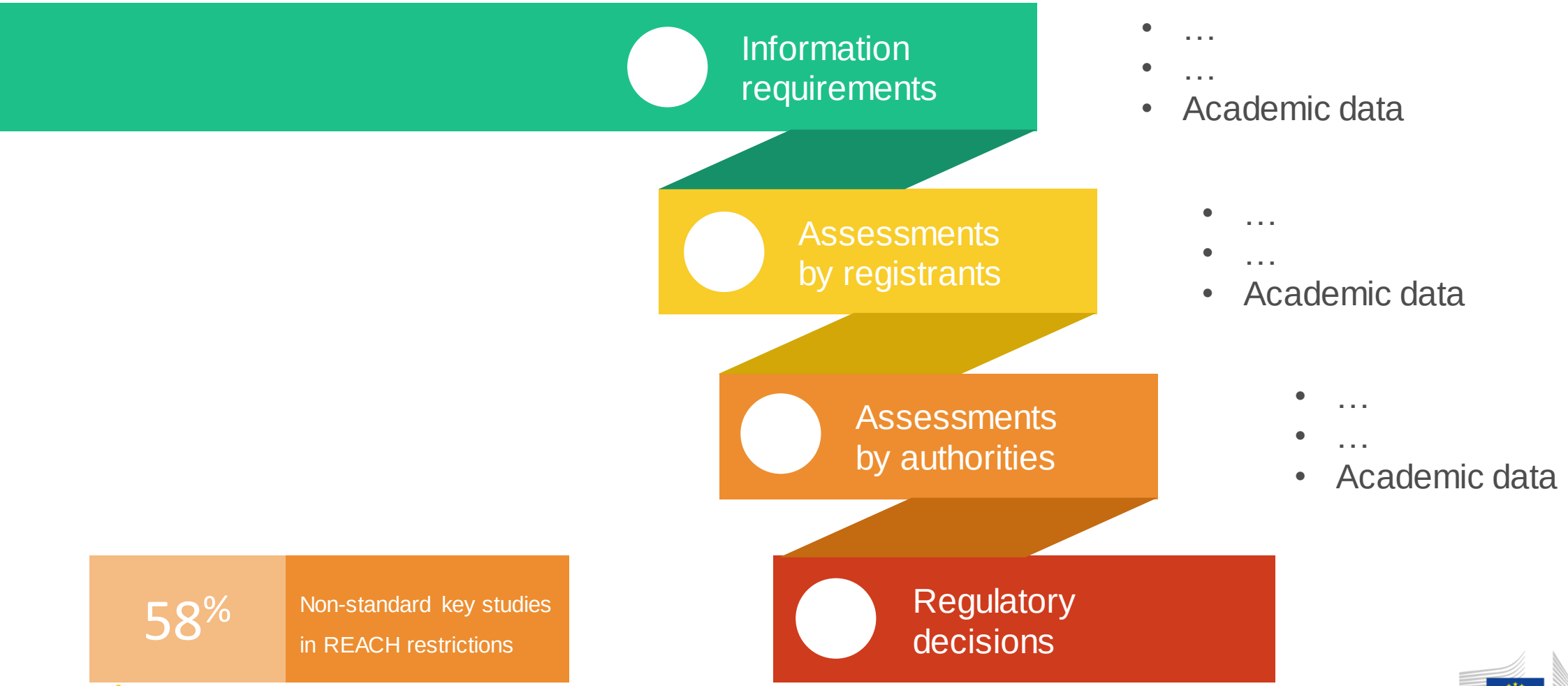
Technical validation of mechanistic NAMs





- Technical characterisation, including reproducibility and biological relevance, without having to establish regulatory application
- Acceptance of mechanistic NAMs that are not stand-alone and/or for which regulatory application is not yet clear

Non-standard data in regulatory assessments



- …
- …
- Academic data

- …
- …
- Academic data

- …
- …
- Academic data

58% Non-standard key studies in REACH restrictions

Borchert et al 2022

European Commission

Technical characterisation

Describe:

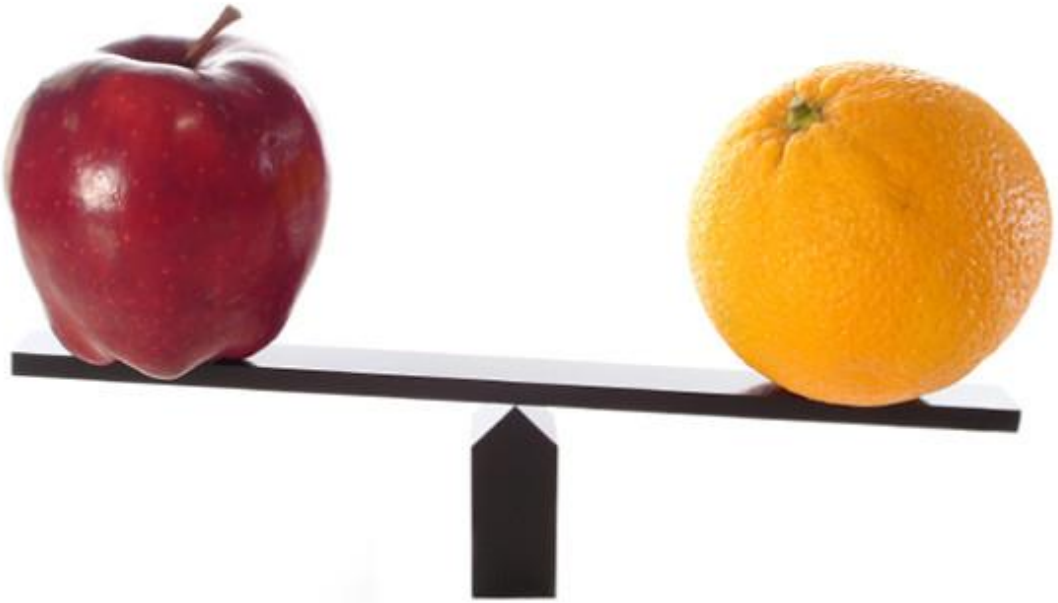
- accuracy
- intra-laboratory reproducibility
- transferability
- applicability domain
- reference chemicals and controls
- limits of detection and quantification

What is considered acceptable may depend on the NAM being evaluated and its intended use

Data reporting should allow for independent evaluation of the NAM, including:

- protocol
- equipment
- computational models being used

Relevance versus accuracy



While accuracy has historically been determined by comparing the results from a new method to results from animal methods, this should not be the default way to determine the relevance of a NAM

Toxicology and Applied Pharmacology
Volume 19, Issue 2, June 1971, Pages 276-360

ELSEVIER

Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests

Carrol S. Weil [a, b], Robert A. Scala [a, b]

Arch Toxicol (2017) 91:521–547
DOI 10.1007/s00204-016-1679-x

REVIEW ARTICLE

Cosmetics Europe compilation of historical serious eye damage/eye irritation in vivo data analysed by drivers of classification to support the selection of chemicals for development and evaluation of alternative methods/strategies: the Draize eye test Reference Database (DRD)

João Barroso[1,2] · Uwe Pfannenbecker[3] · Els Adriaens[4] · Nathalie Alépée[5] · Magalie Cluzel[6] · Ann De Smedt[7] · Jalila Hibatallah[8] · Martina Klaric[1] · Kirsten R. Mewes[9] · Marion Millet[10] · Marie Templier[10] · Pauline McNamee[11]

Arch Toxicol (2014) 88:701–723
DOI 10.1007/s00204-013-1156-8

IN VITRO SYSTEMS

Retrospective analysis of the Draize test for serious eye damage/eye irritation: importance of understanding the in vivo endpoints under UN GHS/EU CLP for the development and evaluation of in vitro test methods

Els Adriaens · João Barroso · Chantra Eskes · Sebastian Hoffmann · Pauline McNamee · Nathalie Alépée · Sandrine Bessou-Touya · Ann De Smedt · Bart De Wever · Uwe Pfannenbecker · Magalie Taillhardat · Valérie Zuang

Analysis of Draize Eye Irritation Testing and its Prediction by Mining Publicly Available 2008-2014 REACH Data

Thomas Luechtefeld [1], Alexandra Maertens [1], Daniel P. Russo [2], Costanza Rovida [4], Hao Zhu [2,3] and Thomas Hartung [1,4]

Regulatory Toxicology and Pharmacology 122 (2021) 104920

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

ELSEVIER

Analysis of variability in the rabbit skin irritation assay

John P. Rooney [a,*], Neepa Y. Choksi [a], Patricia Ceger [a], Amber B. Daniel [a], James Truax [a], David Allen [a], Nicole Kleinstreuer [b]

Toxicology in Vitro 34 (2016) 220–228

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

ELSEVIER

Analysis of the Local Lymph Node Assay (LLNA) variability for assessing the prediction of skin sensitisation potential and potency of chemicals with non-animal approaches

Coralie Dumont, João Barroso, Izabela Matys, Andrew Worth, Silvia Casati *

Concept Article

Uncertainties of Testing Methods: What Do We (Want to) Know About Carcinogenicity?

Martin Paparella [1], Annamaria Colacci [2] and Miriam N. Jacobs [3]

Toxicological Sciences

Evaluation of Variability Across Rat Acute Oral Systemic Toxicity Studies

Agnes L. Karmaus*, Kamel Mansouri†, Kimberly T. To*, Bevin Blake†,1, Jeremy Fitzpatrick‡,2, Judy Strickland*, Grace Patlewicz‡, David Allen*, Warren Casey†, and Nicole Kleinstreuer†

Review

A Section 508–conformant HTML version of this article is available at http://dx.doi.org/10.1289/ehp.1510183.

A Curated Database of Rodent Uterotrophic Bioactivity

Nicole C. Kleinstreuer,[1] Patricia C. Ceger,[1] David G. Allen,[1] Judy Strickland,[1] Xiaoqing Chang,[1] Jonathan T. Hamm,[1] and Warren M. Casey[2]

Reprod Toxicol. 2018 October ; 81: 259–271. doi:10.1016/j.reprotox.2018.08.016.

DEVELOPMENT OF A CURATED HERSHBERGER DATABASE

P Browne[a], NC Kleinstreuer[b], P Ceger[c], C Deisenroth[d], N Baker[e], K Markey[f], RS Thomas[d], RJ Judson[d], W Casey[b]

EPA Public Access
Author manuscript
Comput Toxicol. Author manuscript; available in PMC 2021 August 01.

About author manuscripts | Submit a manuscript

Published in final edited form as:
Comput Toxicol. 2020 August 1; 15(August 2020): 1–100126. doi:10.1016/j.comtox.2020.100126.

Variability in *in vivo* studies: Defining the upper limit of performance for predictions of systemic effect levels

Ly Ly Pham[1,2], Sean Watford[1,3], Prachi Pradeep[1,2], Matthew T. Martin[1,4], Russell Thomas[1], Richard Judson[1], R. Woodrow Setzer[1], Katie Paul Friedman[1]

Accuracy

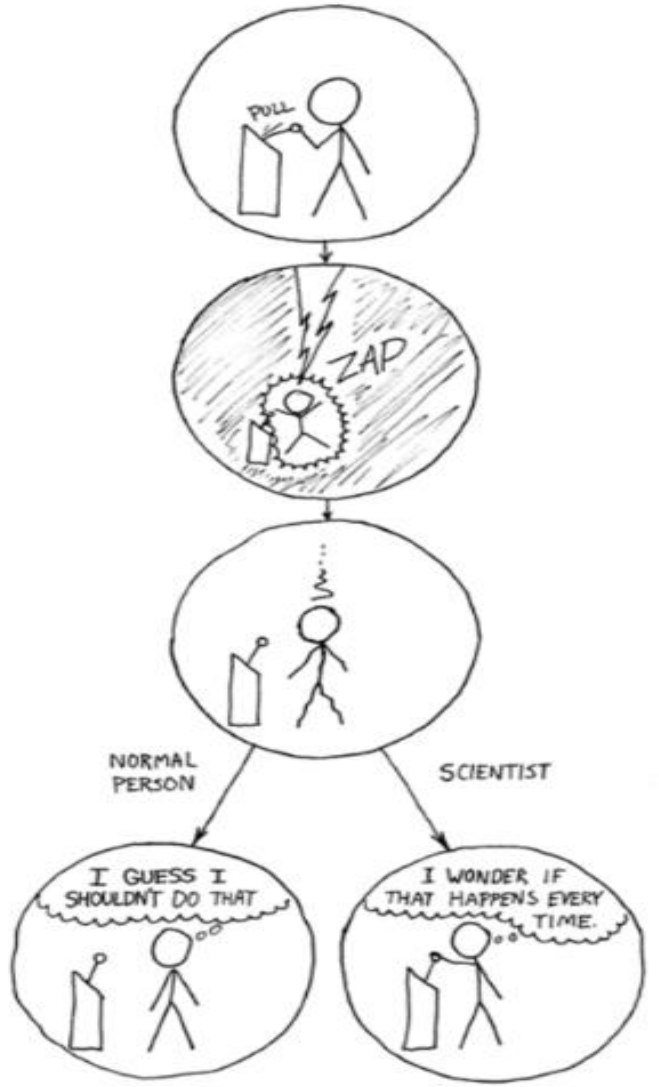
Traditional animal test methods should not be assumed to provide data relevant to human biology or mechanisms of toxicity and be the “right” answer to determine if another method is valid.

Instead, accuracy can be demonstrated by considering:

- Consistency across NAMs
- Ability to identify positive and negative reference chemicals
- Greater emphasis on biological relevance and reproducibility

Reproducibility and ring trials

- Demonstrating reproducibility is essential
- Ring trials are the most time-consuming and expensive part of a validation study and are often more a reflection of laboratory quality or expertise than of a NAM's reproducibility
- Properly designed training and transfer studies are essential and informative
- Proficiency testing adds confidence on capacity of a laboratory to perform test



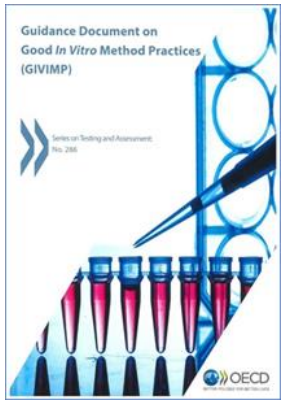
WLR and BLR of validated *in vitro* methods

Method (eye irritation)	WLR	BLR
EpiOcular EIT	95%	93%
SkinEthic HCE	92%	95%
LabCyte EIT	96%	87%
MCTT HCE EIT	93%	90%
SkinEthic HCE TTT	85-100%	90-100%
Vitrigel	80-100%	92%
Ocular Irritection	80-90%	84-86%

Method (skin sensitisation)	WLR	BLR
DPRA	85%	80%
ADRA	100%	100%
kDPRA	96%	88%
h-CLAT	80%	80%
U-SENS	90%	84%
IL-8 Luc	88%	88%
GARDskin	82-89%	92%

Thyroid Validation Study, a collaborative effort!

- 15 **EU-NETVAL** labs
- 18 *in vitro* methods
- 14 method developers
- Transfer & optimise
- Reproducibility (WLR)
- Relevance (30 chem)
- Data for IATA/DA



EURL ECVAM
European Union Reference Laboratory for Alternatives to Animal Testing

European Commission

Final thoughts

- Validation is essential to facilitate acceptance and ensure sound science-based decisions
- Validation needs to keep pace with rapid scientific progress, e.g. emergence of Defined Approaches (data integration), computational models, new technologies such as Organ-on-Chip
- Important to maintain scientific integrity, credibility and usefulness while making process more efficient
- Frame validation as a process to characterize and reduce uncertainty rather than a ring trial to demonstrate "toxicological equivalence“
- Important to characterize (human) relevance and uncertainty of reference *in vivo* method
- Validation ≠ regulatory acceptance and use

Thank you

joao.barroso@ec.europa.eu

European Commission