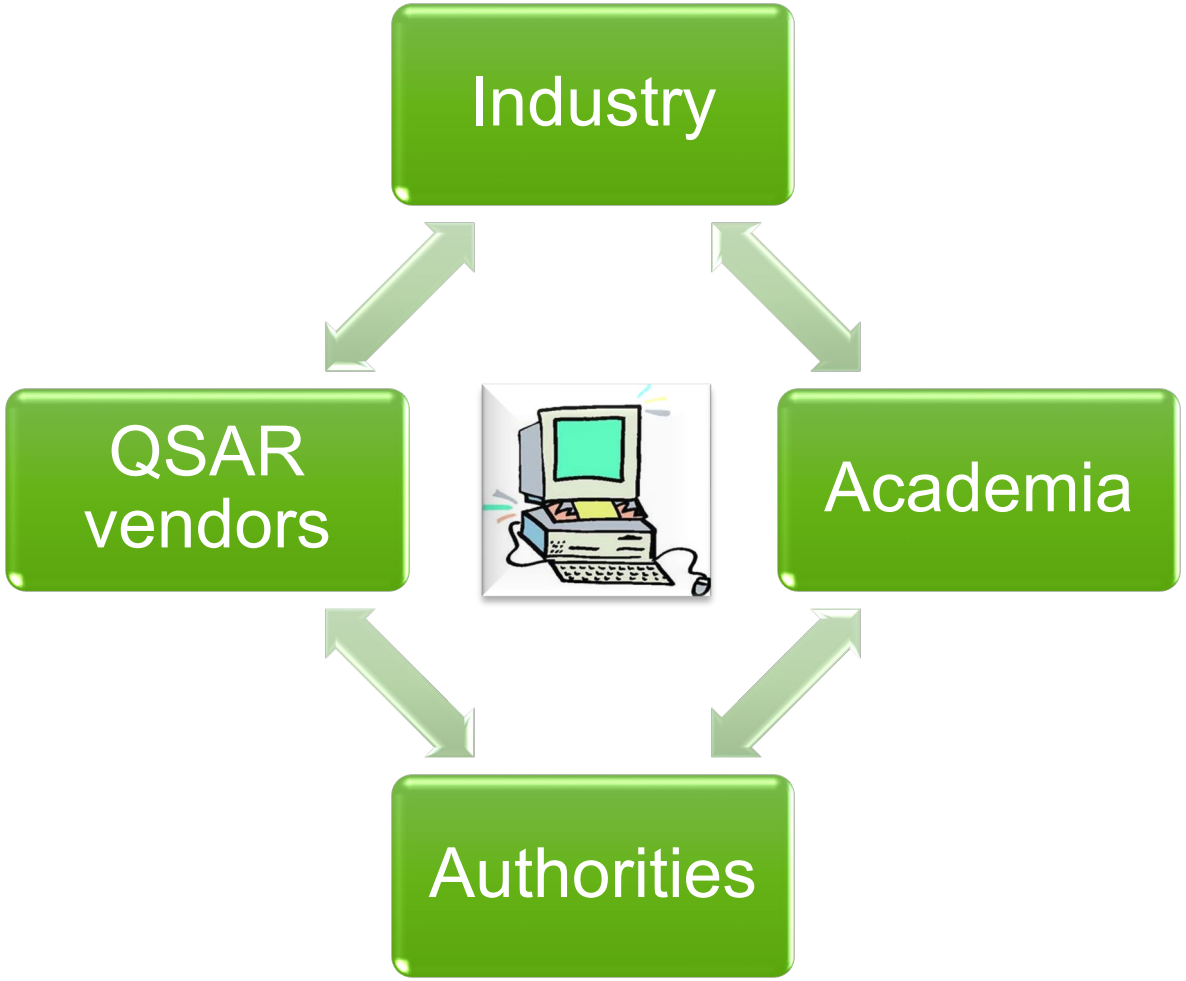


Use of QSAR/Read Across for the evaluation of pesticide metabolites and impurities

An industry perspective

Dr. Markus Frericks, BASF SE

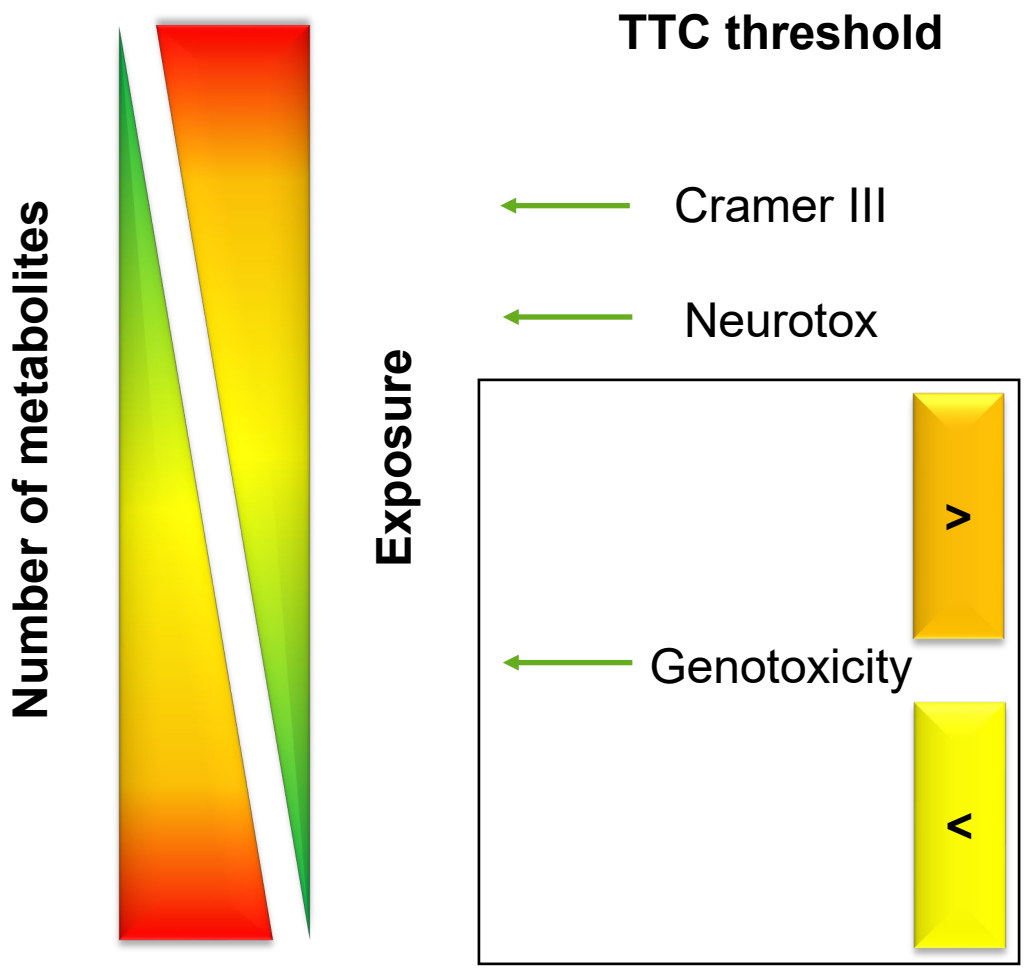
BASF
We create chemistry

What is the question

- Plant and life stock residues
- EFSA proposal EFSA Journal 2016; 14(12): 4549 → not accepted; moved to WHO
  - Current regulatory situation: Unclear
- Basic principle: Combination of evaluation of toxicological properties with dietary risk assessment
- Steps to be performed:
  - Assessment of genotoxic potential
  - Grouping of metabolites
  - Dietary risk assessment
  - Evaluation of general toxicity

Threshold of Toxicological Concern (TTC) for prioisation

Focus of this presentation: Genotoxic concern



- Indrusty supports use of TTC thresholds
- Full assessment of residues with significant exposure (e.g. above TTC threshold Cramer III or neurotox)
- Grouping and identification of key metabolites based on metabolism, exposure and potential toxicity (not part of this presentation)
- Use of QSAR/Read across for **genotoxicity**
- Exposure below TTC genotoxicity should be considered in WoE

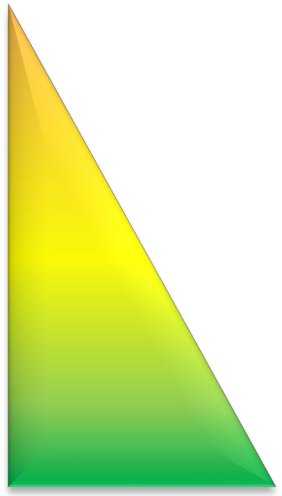
Proposed evaluation scheme by EFSA

- Evaluation needs to be done for:
  - Mutagenicity: **Ames**
  - DNA-Damage: **in vitro MNT**
  - If done by testing: ~120-150 k€ (synthesis, analytics, GLP-testing, in vivo follow up) per metabolite

- Methodology of evaluation:
  - Initial screening: Profiling for reactive elements (e.g. by OECD Toolbox profiler)
  - QSAR/Read Across: Based on ICH M7 for pharmaceutical impurities (Ames only)
    - Rule based expert system (e.g. Derek)
    - Statistical QSAR system (e.g. Chemtunes, Leadscope, Multicase)

Read Across requires expert review

- **QSAR/Read across depends on:**
  - Available data — Metabolite to
    - active igredient
    - plus all metabolites with data
    - metabolism information
    - other AI‘s same compound class
    - pharmaceuticals
    - substructures with data

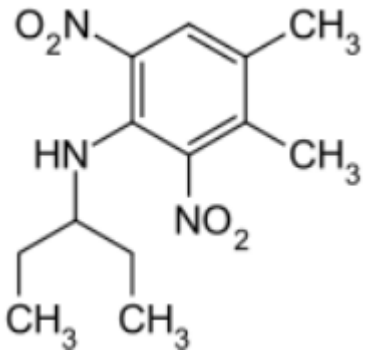


Amount of data

  - Does my chemistry react? — Presence/Relevance of structural alerts

  - How is my chemistry described and similarity assessed?
    - Descriptors and similarity matrices
    - Substructures/SMARTs

1st example: Dinitroaniline herbicides
Data quality

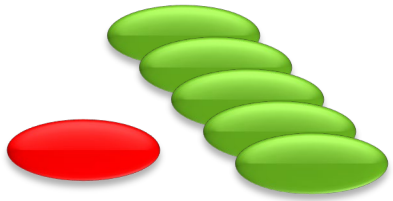


- Models often built on published data
  - Easy way out: Take worst case assumption positive
- Industry perspective: **Should be based on - weight of evidence**

  - up to date authority decision
  - Example: Dinitroaniline herbicide

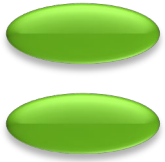
    Synthesis can result in nitrosamine impurities

Ames test:



Could be traced to impurity

Mammalian cells:

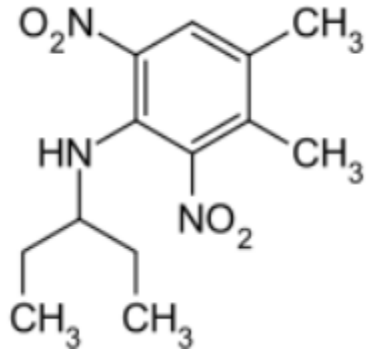


5 metabolites tested

Other dinitroanilines

Based on a weight of evidence: **negative**

1st example: Dinitroaniline herbicides
Structural alerts are context dependent



- Structural alerts in models are based
  - a.) expert knowledge
  - b.) statistical ratio positive/negative in training data
- Alerts are context dependent: Training data determines structural element and alert statistics
  - electron donating/withdrawing function in the molecule
  - Stabilisation via π-electron system
  - metabolic accessibility
  - position in the molecule

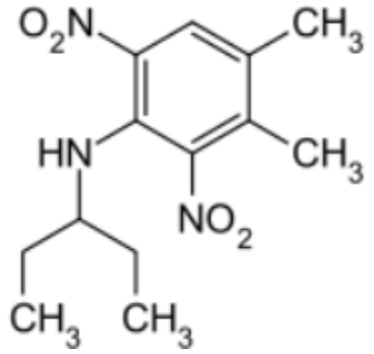
Example: Dinitroaniline

Structural alerts: 2x -NO_2-group and secondary aromatic amine

Experimental evidence: negative for parent, for metabolites and other dinitroanilides

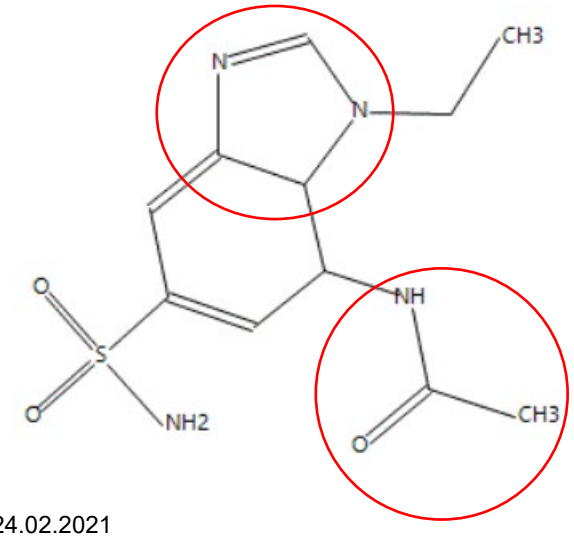
Weight of evidence: **Model not fit for purpose → targeted read across**

1st example: Dinitroaniline herbicides
Use of ADME information



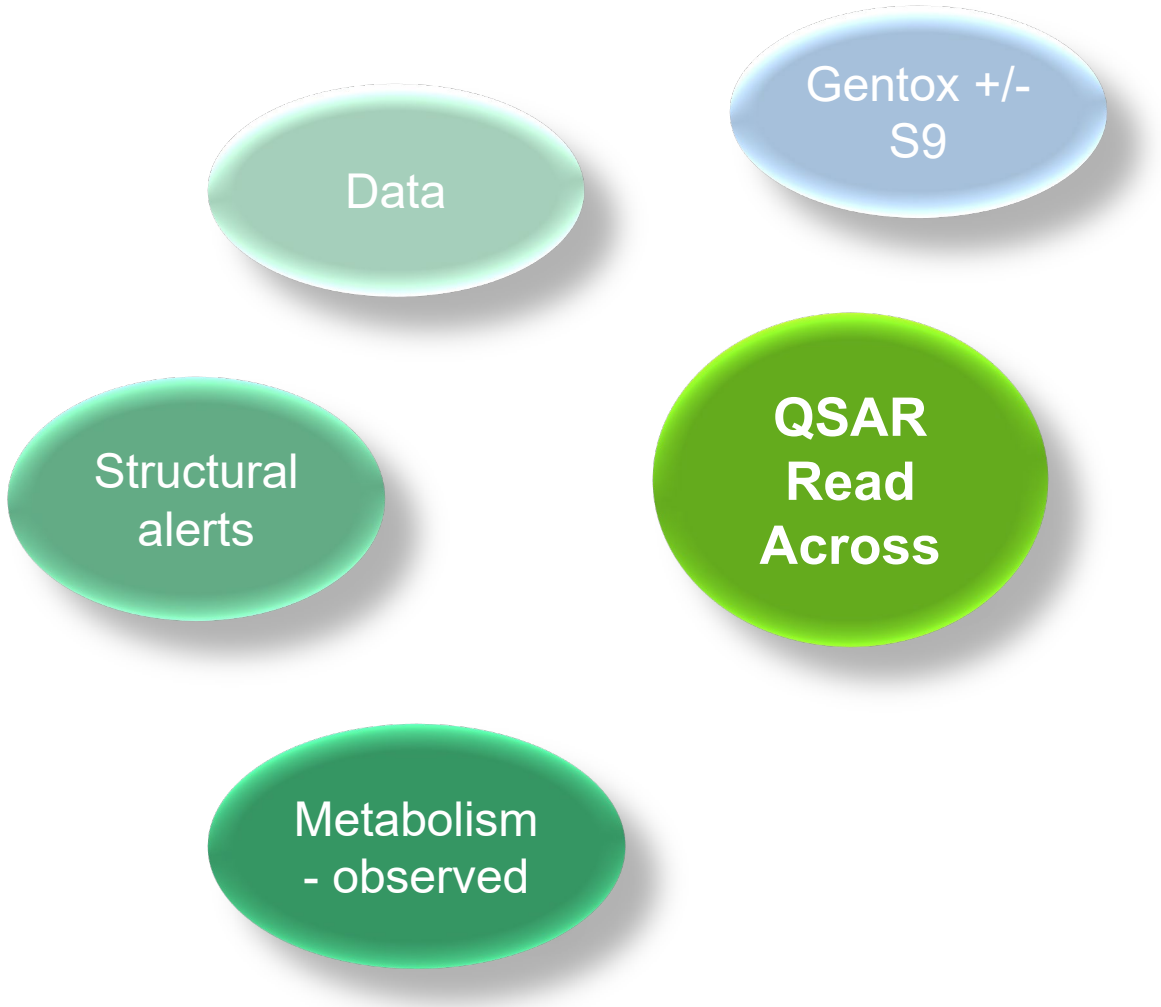
- Basic idea:
  - Genotoxicity assays have a metabolic activating system (S9 mix)
  - Main metabolites observed in bioavailable matrices can be used to expand chemical space
  - Relevant matrices - urine and bile (if site of toxicity)

- Case dinitroaniline herbicides:
  - side chain hydroxylation
  - Reduction of NO_2 to $-NH_2$
    - acid anhydrid formation
    - cyclization



Metabolic activation does <u>not</u> lead to genotoxicity

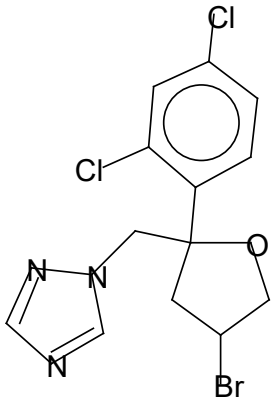
Important factors for QSAR/read across
Need for weight of evidence analysis

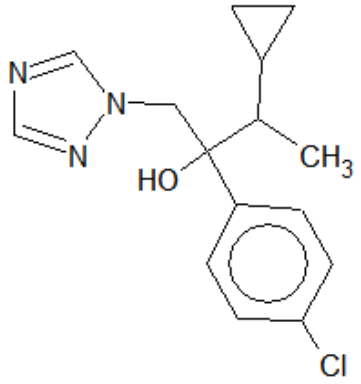


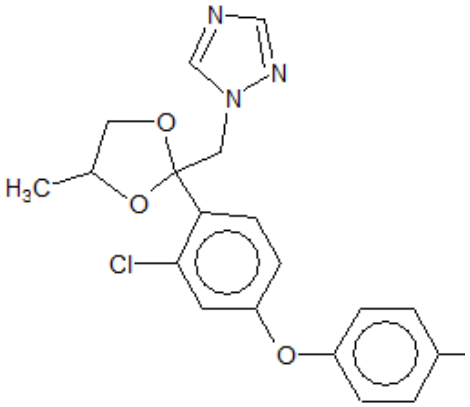
A 2nd real life example

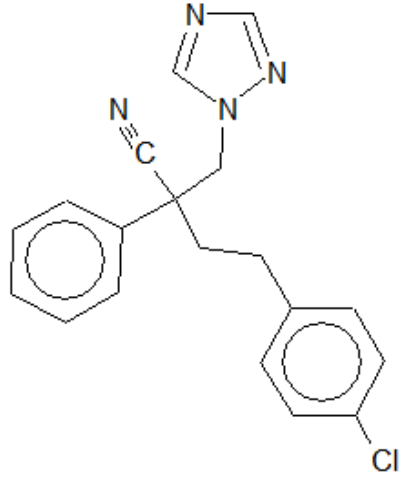
Mutagenicity of azole fungicides

Bromuconazole	c1cc(c(cc1Cl)Cl)C2(CC(CO2)Br)Cn3cncn3
Cyproconazole	CC(C1CC1)C(Cn2cncn2)(c3ccc(cc3)Cl)O
Difenoconazole	CC1COC(O1)(Cn2cncn2)c3ccc(cc3Cl)Oc4ccc(cc4)Cl
Fenbuconazole	c1ccc(cc1)C(CCc2ccc(cc2)Cl)(Cn3cncn3)C#N
Fluquinconazole	C1=CC2=C(C=C1F)C(=O)N(C(=N2)N3C=NC=N3)C4=C(C=C(C=C4)Cl)Cl
Penconazole	CCCC(Cn1cncn1)c2ccc(cc2Cl)Cl
Tebuconazole	CC(C)(C)C(CCc1ccc(cc1)Cl)(Cn2cncn2)O
Epoxiconazole	c1ccc(c(c1)[C@@H]2[C@@](O2)(Cn3cncn3)c4ccc(cc4)F)Cl
Prothioconazole	C1CC1(C(CC2=CC=CC=C2Cl)(CN3C(=S)N=CN3)O)Cl
Prothioconazole-desthio	C1CC1(C(CC2=CC=CC=C2Cl)(CN3C=NC=N3)O)Cl
Mefentrifluconazole	CC(Cn1cncn1)(c2ccc(cc2C(F)(F)F)Oc3ccc(cc3)Cl)O
Metconazole	CC1(CCC(C1(Cn2cncn2)O)Cc3ccc(cc3)Cl)C
Tetraconazole	c1cc(c(cc1Cl)Cl)C(Cn2cncn2)COC(C(F)F)(F)F
Triticonazole	CC1(CCC(=CC2=CC=C(C=C2)Cl)C1(CN3C=NC=N3)O)C









A 2nd real life example
Available genotoxicity data

- Data extracted from pesticide regulatory documents and pharmaceutical data

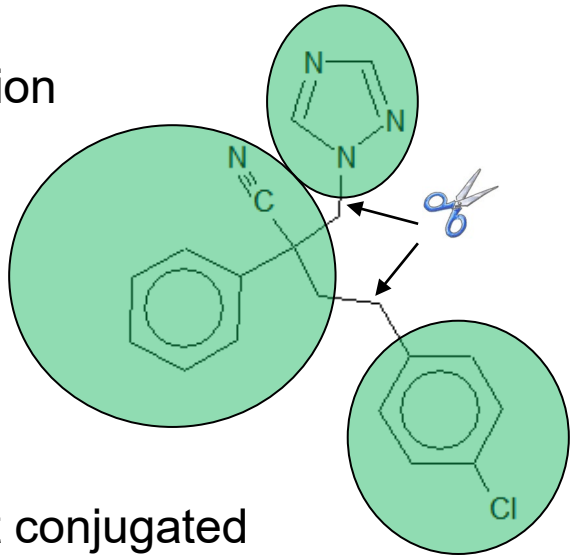
Compound class	Ames	In vitro CA/MNT (positive)	In vivo MNT
Pesticide AI	13	13 (5)	13
Metabolites	24	6	6
Pharmaceuticals	12	10 (1)	11
Total	**49**	**29 (6)**	**30**

- Database strongly skewed towards negative
- How to increase confidence?
- Can we expand the chemical space?

A 2nd real life example:

Use of ADME data:

Bridge and side chain:
- Gets cleaved by oxidation
- Further oxidation and conjugation



Triazole metabolites:
- alcohol, amine, acid
- Negative for mutagenicity

Substituted halogenated phenols:
- Score positive for mutagenicity in some QSAR base
  based on other phenolic compounds

Short lived dihydro-diols that get conjugated

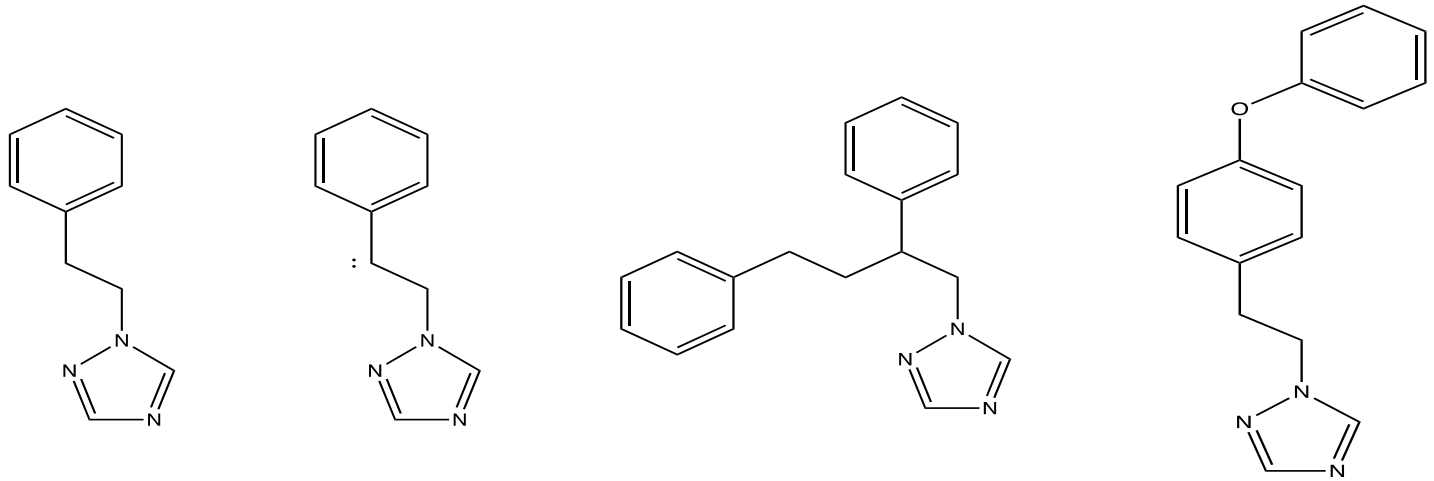
Conjugation: Glucuronidation
Sulfatation

Main metabolites

A 2nd real life example

Building of a substructure space

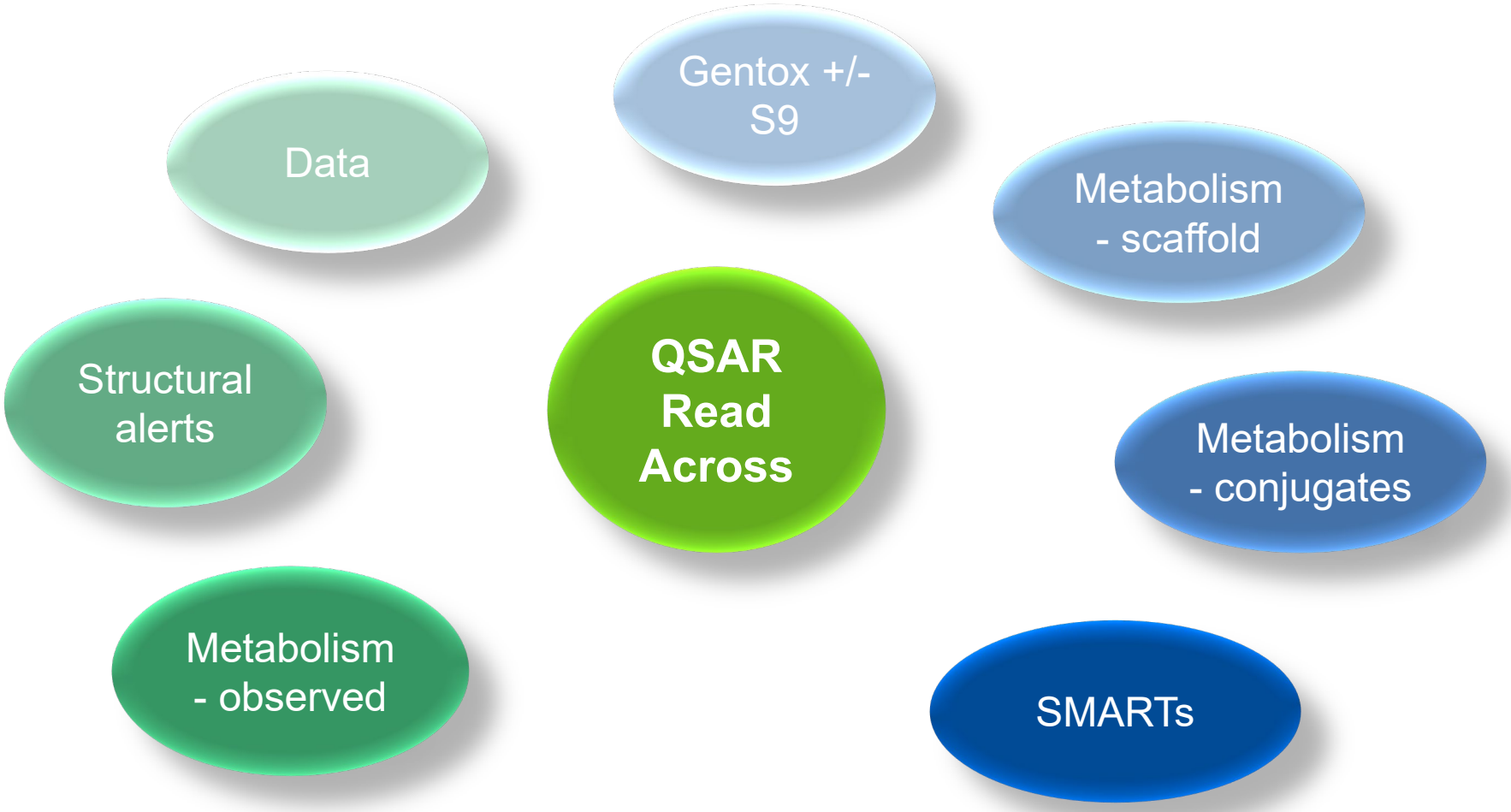
- We used the Bemis and Murcko method to generate sub-structures
- This allowed us to define an initial set of 30 sub-structures
- Substructures can than be used to querry known genotoxicity databases and calculate the respective regression terms for each



Safe space for read across **Substructures + observed metabolism**

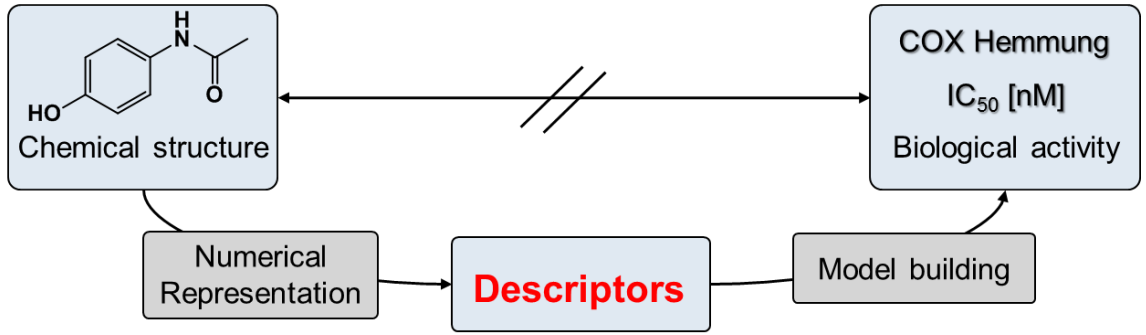
BASF
We create chemistry

Important factors for QSAR/read across
Need for weight of evidence analysis



Alternatives

Chemical descriptors



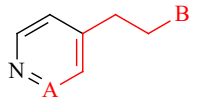
- Chemicals can be computed in different ways
  - Analogy:
    - English
    - Chinese
    - Inuit
    - Pictograms

Chemistry descriptors of structural similarity

- Atom types — H, Cl, Br, Na, C, O, N, P, S, …
- Atomic charges — Na^+, O^-, Fe^{2+}, …
- Hybridizations — sp^3, sp^2, sp, …
- Bonds

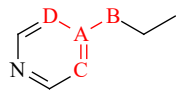
Topological descriptors for structural similarity

- Distance
  Atom pairs



Distance AB is 5

- Neighbors
  Atom-centered fragments



First neighbors of A are B, C and D

Other descriptors of similarity

- Molecular weight
- Solubility (e.g. $logK_{ow}$)
- Biological activity

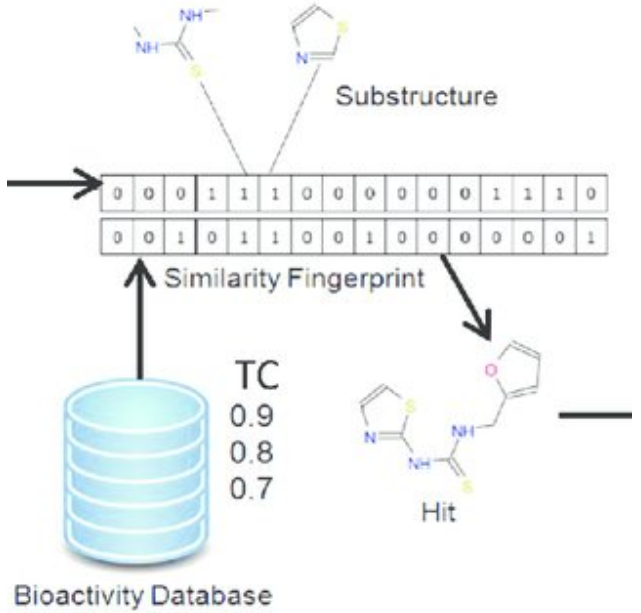
Structural alerts = **funtional groups** («profilers»)

Each method has pro and contra arguments:
Good for one chemistry might fail for others

Industry: **Retain freedom to operate**

Alternatives
Similarity

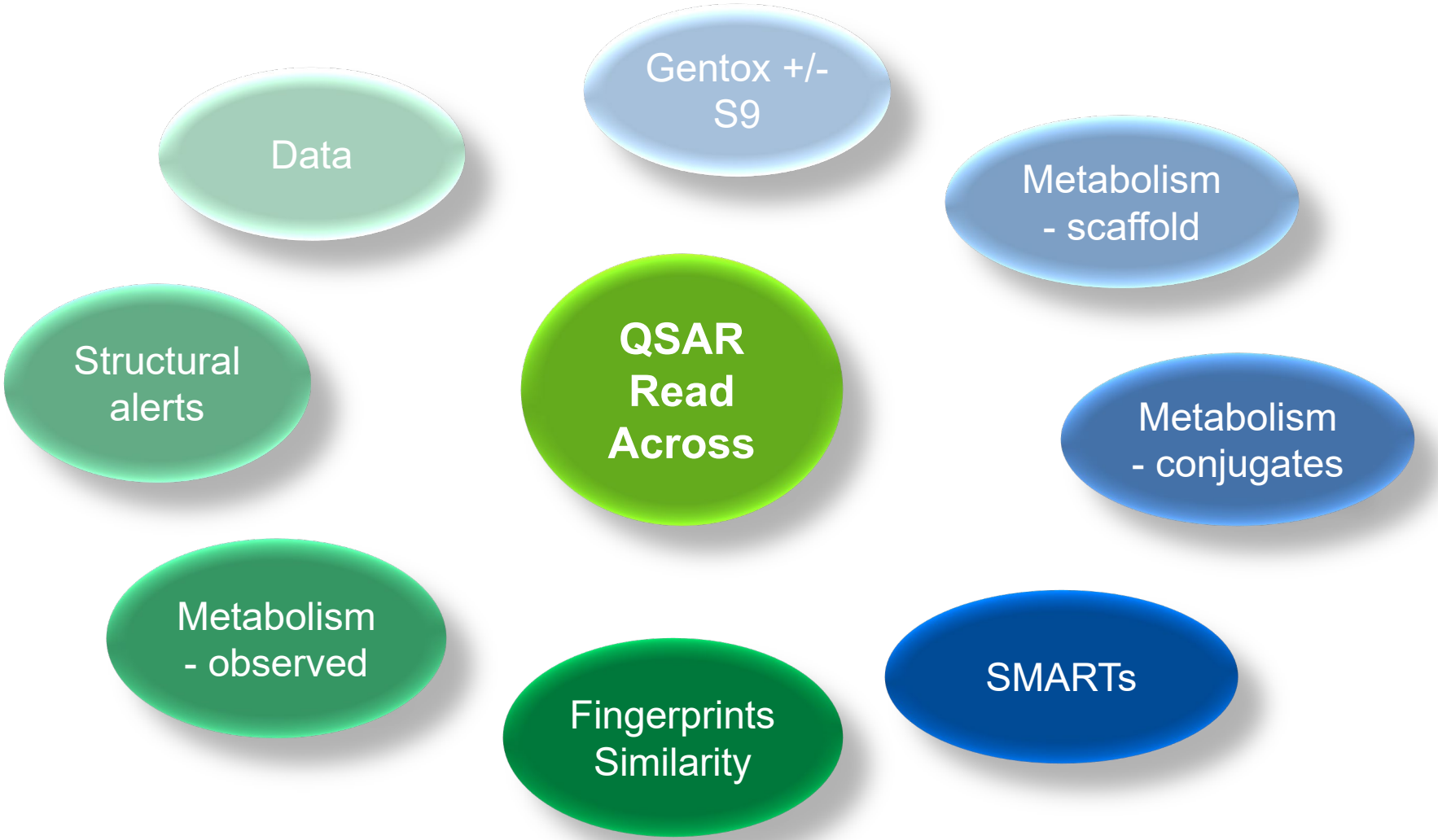
- Chemical descriptors are transformed into binary vectors (fingerprints)
- Similarity matrices like Tanimoto compare those
- Higher overlap → higher similarity



- For biological activities, where a substructure (pharmacophore) drives the effect a weighted assessment needs to be done
  - Example: Organophosphate → neurotoxicity depends on acetylcholin esterase
    Direct receptor interaction

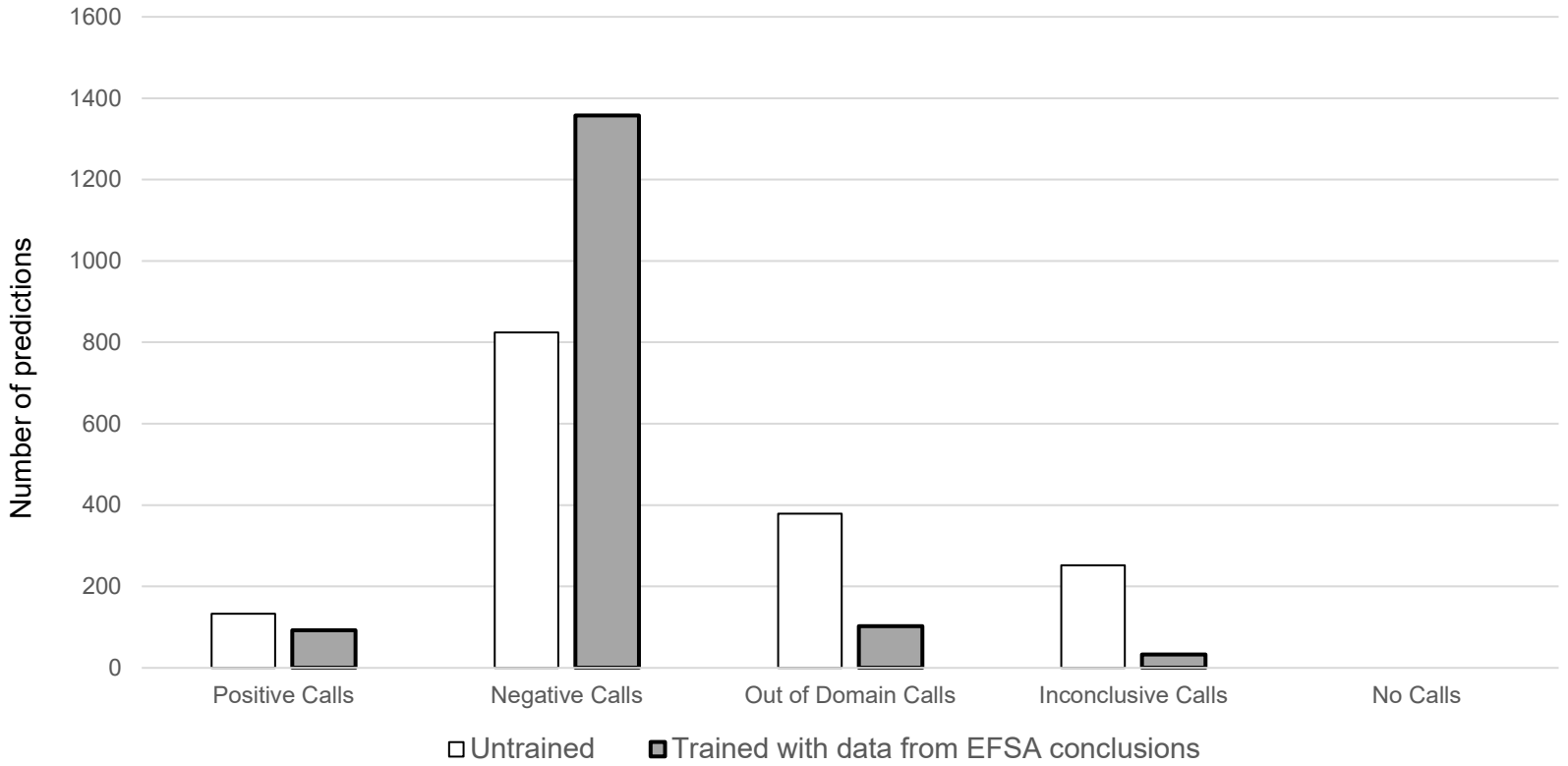
Each method has pro and contra arguments
Good for one chemistry might fail for others
Industry: Retain freedom to operate

Important factors for QSAR/read across
Need for weight of evidence analysis



Is the training database adequate?

- Most QSAR models are built on pharmaceuticals or chemicals
  - Pesticides usually not part of the training database
  - Example: Case Ultra Ames (old version) based on 1500 AI‘s and metabolites



- **Training data:**
  ~ 700 Ames tests from PPP
- **Untrained**
  - **50% positive, inconlcusive, out of domain**
- **Trained**
  - **85% in domain**

How can we achieve high acceptance?
Data sharing

Current Status
Data base not adequate

Positive → Test

Negative
→Prediction **out of domain**
→Not valid
→ Test

Include PPP data

Future:
Data base adequate

Positive → Test

Negative
→prediction **in domain**
→Valid
→ No testing

Issue: **Data protection**
Regulatory acceptance

Industry proposal:
Data sharing initiative

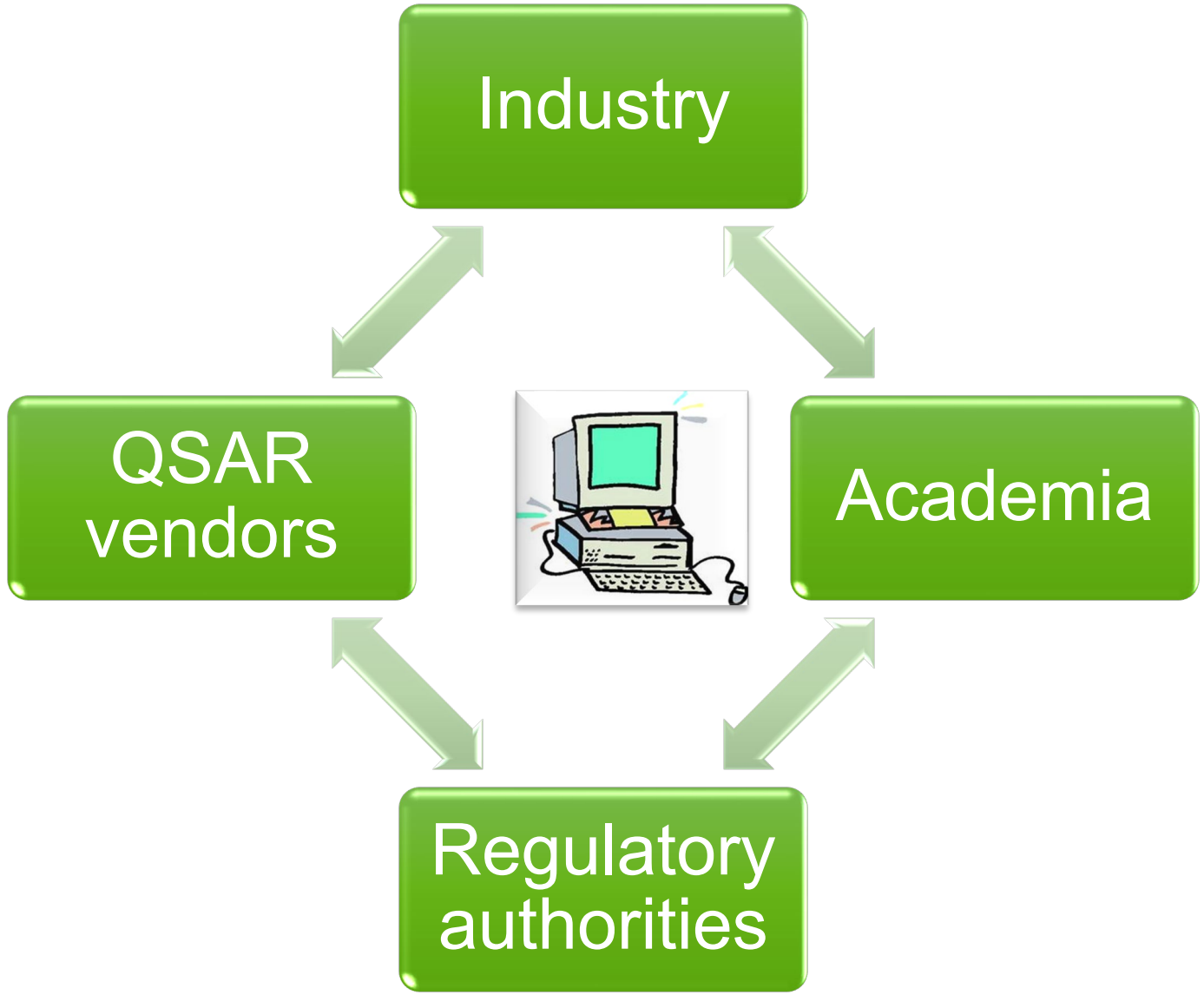
Required stakeholders:
Industry :
Tentative **Yes** **CLE**
QSAR vendors: **Yes** from all major vendors
via **honest broker** (e.g. Lhasa)

Regulatory authorities: **EFSA support**
BfR interested

Main issues: Funding and IP protection

Summary and conclusion:

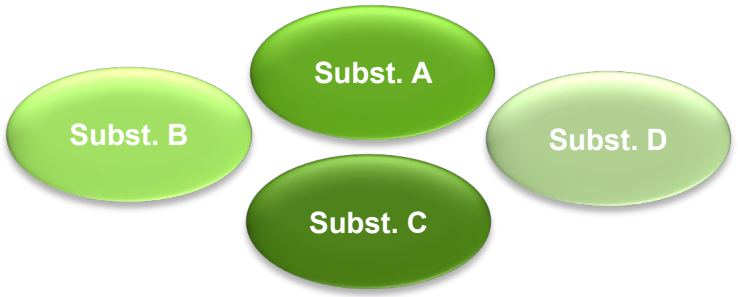
- Industry supports the use of QSAR and Read Across
- TTC genotoxicity should be part of evaluation
- QSAR/RA needs to be fit for purpose
- Current models limited since plant protection chemistry is not included
- WoE should allow alternative approaches



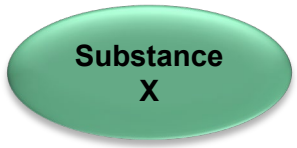
We create chemistry

QSAR models Applicability domain

Data know



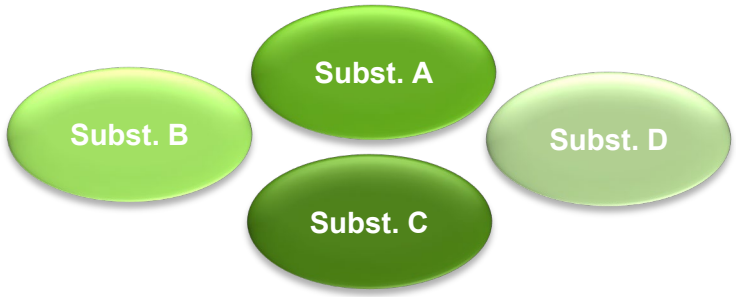
Prediction



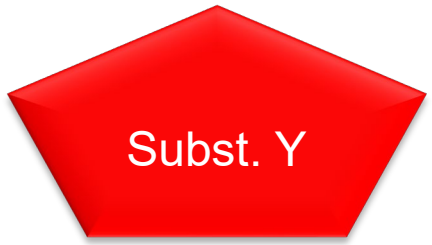
- Prediction **possible**, if data base is **adequat** for substanz X
  - Prediction „**in Domain“ (applicability domain)**

QSAR models Applicability domain

Data unknown



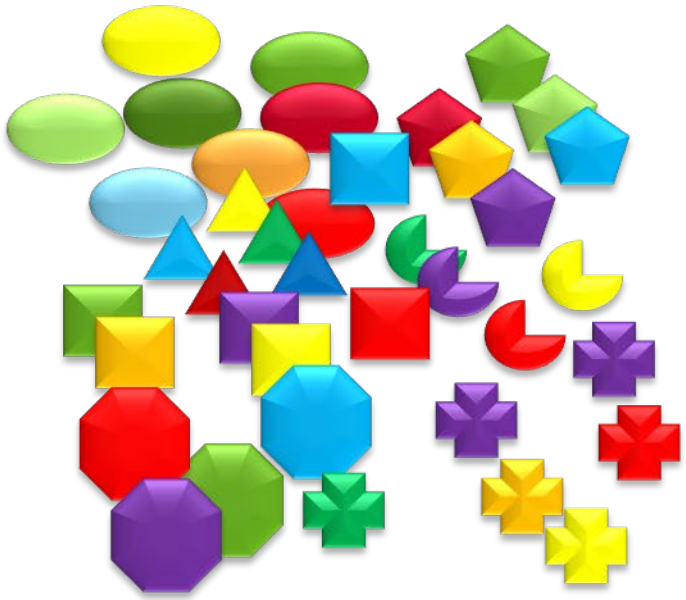
prediction



- Prediction **not possible**, if data base is **not adequat** for substanz Y
  - Prediction „**out of Applicability Domain**“

QSAR models Applicability domain

Data known



Trained Model

Prediction