Evaluation of the applicability of existing QSAR models for predicting the genotoxicity of pesticides

E. Fioravanzo[1], C. L. Battistelli[3], A. Bassan[1], R. Benigni[2], C. Bossa[3], A. Ciacci[1], M. Fuart Gatnik[1], R. Serafimova[4], O. Tcheremenskaia[3]

[1]S-IN Soluzioni Informatiche, Vicenza, Italy. [2]Alpha-Pretox, Rome, Italy. [3]Istituto Superiore di Sanità, Roma, Italy. [4]EFSA, Parma, Italy.

2104 / P441

Background & Objectives

The guidance document issued by EFSA Panel on Plant Protection Products and their Residues (doi: 10.2903/j.efsa.2016.4549) establishes a process for the definition of pesticide residues, including the evaluation of the potential risk based on toxicity and potential for dietary exposure. While a comprehensive toxicological dossier is developed for active substances, often none or only limited information about toxicological properties of their metabolites is available. Thus, the use of QSAR models and read across is proposed for the assessment of genotoxic potential of all identified metabolites as a first step in the residue definition procedure. Within an ongoing EFSA-funded project (OC/EFSA/PRAS/2016/01) the applicability and reliability of existing 52 commercial and freely-available QSAR models were critically evaluated for the prediction of genotoxicity of pesticides.

Methodological workflows & Tools

Preparation of the datasets

- Assays selection
- Structures selection
- Experimental data analysis
- Chemical space analysis

Predictions

- Training set analysis
- Applicability domain analysis
- Predictions analysis

Commercial

Derek Nexus v.5 and Sarah Nexus v.2.0.1 by Lhasa Limited
CASE Ultra 1.6.2.1 by MultiCASE Inc.
Leadscope Model Applier v2.2.1.1 by Leadscope Inc.
ChemTunes ToxGPS by Molecular Networks GmbH
ACD/Percepta 2016 (Build 2911) by ACD/Labs Inc.

29 models
24 statistical based
5 expert rule-based

Free

Lazar v. 1.1.0 by In Silico Toxicology GmbH
ToxTree v. 2.6.13 by Ideaconsult Ltd and JRC
Vega v.1.1.4 by IRCCS

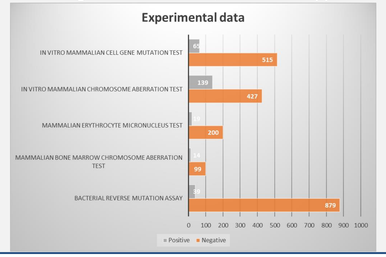
5 models
2 statistical based
3 expert rule-based

TEST TYPE	GUIDELINE	METHOD	Models	statistical based	expert-rule based
Bacterial Reverse Mutation Assay	OECD Guideline 471 and 472	in vitro	18	12	6
Mammalian Bone Marrow Chromosome Aberration Test	OECD Guideline 475	in vivo	3	2	1
Mammalian Erythrocyte Micronucleus Test	OECD Guideline 474	in vivo	6	4	2
In vitro Mammalian Chromosome Aberration Test	OECD Guideline 473	in vitro	7	6	1
In vitro Mammalian Cell Gene Mutation Test	OECD Guideline 476	in vitro	6	5	1

Preparation of the dataset

Experimental data were reviewed by an expert. The analysis aimed to review the quality of experimental data, based on the information present in the genotoxicity EFSA database (https://zenodo.org/record/582137#.WoQTxPmnGUI) and to conclude on the overall outcome of experimental data per compound. The criteria, that were applied for the expert review are:

- The general criteria applied was a conservative, worst case approach;
- *In vitro* assays, without and with metabolic activation: positive if either result is positive (for Ames test, consisting of several strains without and with metabolic activation: positive if one strain result is positive);
- *In vivo* assays: no metabolic activation is applicable. A special case is that of the Mammalian erythrocyte micronucleus test: negative results without evidence of interaction with target cells were not considered;
- Experiments repeated with more systems (e.g., in vitro mammalian cells gene mutation), or repeated with the same system(s), or contradictory (discordant) results: judgement by Expert inspection according to a conservative, worst case approach. Decisions were taken on a case-by-case.

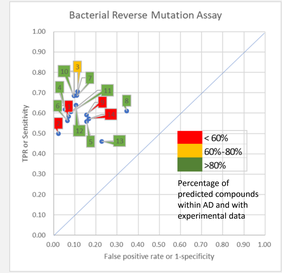


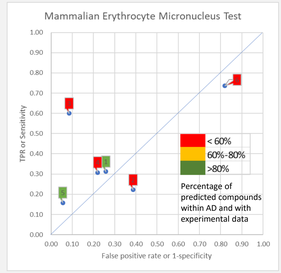
The experimental results are largely unbalanced, with a strong prevalence of negative genotoxicity results

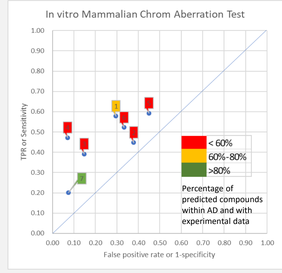
Conclusions

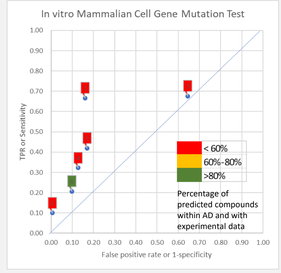
Overall the results point to a substantial difference between the prediction of the Ames test and that of the other experimental assays. All (Q)SAR models for Ames test are in the top left ROC area, thus pointing to statistically significant predictions. Sensitivity ranges between 46% (Toxtree) and 71% (a model from Leadscope), Specificity between 66% (Lazar) and 98% (Percepta). On the contrary, reliability of the (Q)SAR models for assays different Ames is still far from optimality and the ROC graph represents random results. **Why the Ames test is predicted better?** A first hypothesis is that a much larger historical database of Ames results is available for training the (Q)SAR models; this is linked to its central role in regulations. Other differences may be in the nature itself of the assays and assay results. The Ames test has a clear scientific basis (each strain has been designed and constructed as to be able to respond to certain types of potentially mutagenic chemical structures). On the contrary, the other *in vitro* genotoxicity assays have demonstrated to be prone to false positive results, whereas the *in vivo* assays have shown limited sensitivity (mainly due to ADME reasons) (Zeiger, E. 2004. History and rationale of genetic toxicity testing: an impersonal, and sometimes personal, view. Environ.Mol.Mutagen. 44:363-371). Thus, the Ames test appears to be a “cleaner” (less noisy) tool to identify DNA-reactive chemicals; as a consequence, the relationship between the biological effect and the causative chemical features are expected to be more easily identified.

Predictions

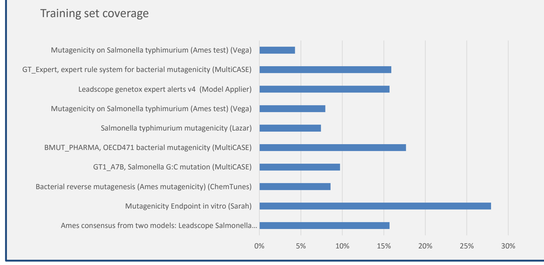


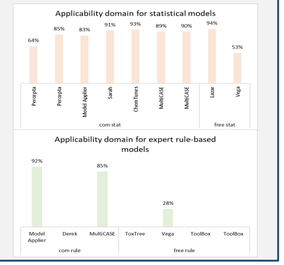






Ames models: Training set coverage and AD analysis





Disclaimer: The authors declare that the results shown here have been produced under a contract with EFSA (OC/EFSA/PRAS/2016/01) and that the opinions expressed do not represent EFSA’s official position.