

# Comparison and use of QSAR software to estimate Carcinogenicity

Roncaglioni A<sup>1</sup>, Milan C<sup>1</sup>, Schifanella O<sup>1</sup>, Benfenati E<sup>1</sup>

<sup>1</sup>Institute for Pharmacological Research “Mario Negri”, Department of Environmental Health Sciences, Laboratory of Environmental Toxicology and Chemistry, Via La Masa 19, 20156 Milan, Italy



## Introduction

The aim of REACH is to improve the protection of human health and the environment. To avoid increase in animal testing, to save time and resources and to raise the information robustness, REACH promotes the use of alternative methods. Among them, the use of *in-silico* methods and QSARs is encouraged by regulators as a tool for supporting and optimizing risk assessment strategies.

Several commercial and publicly-available models for carcinogenicity have been developed, offering tools that can be used in different situations.

The work here described gives an overview of the application of these computer programs for predicting carcinogenicity and their respective advantages and disadvantages.

## Materials and Methods

### Software

Software used to predict carcinogenic activity, their characteristics and output interpretation are reported in table below.

SOFTWARE	METHOD	AVAILABILITY	OUTPUT INTERPRETATION
<b>Caesar</b> v1.0.0.6	QSAR model based on CP-ANN	Freeware, online	Pos/Neg results automatically assigned
<b>Multicase</b> AF1 Module (Male rat, not proprietary)	Statistical method, fragment based + modulating factors	Commercial	Pos/Neg results automatically assigned, Borderline and Probably inactive results interpreted by user
<b>Toxtree</b> v1.60	Collection of knowledge based Structural Alerts (SA)	Freeware, downloadable	Presence of SA → Pos Absence of SA → Neg
<b>Derek</b> DFW_11.0.0	Collection of knowledge based rules	Commercial	Presence of Alert → Pos Nothing to report → Neg
<b>Lazar</b> Rat carcinogenicity (both sexes)	k nearest neighbours, excluding the identical compound	Freeware, online (batch not supported)	Pos/Neg results automatically assigned
<b>Hazard Expert</b> Module in Pallas 3.0	Collection of Structural Alerts + modulating factors for bioavailability	Commercial	None, uncertain (values up to 35) → Neg Possible, probable and highly probable (all other values) → Pos
<b>Topkat</b> v 6.1 Male rat v 3.2; Female rat v 3.2	Statistical method based on 2D descriptors (e-state, topological) with explicit definition of the Optimum Prediction Space	Commercial	Male AND female rat result < 0.5 → Neg Male OR female rat result ≥ 0.5 → Pos

### Datasets

Each software were tested using two datasets of compounds:

- **Berkeley Carcinogenic Potency Database (CPDB, publicly available)**  
805 chemicals with experimental data on rat carcinogenicity have been extracted from D55Tox CPDB database. Chemicals with experimental positive response for at least one sex were considered carcinogenic.
- **FDA 2009 SAR Carcinogenicity - SAR Structures (commercial)**  
It contains 2090 compounds. A total of 655 compounds were in common with the first dataset and the overall concordance for the experimental data was above 95%.

Once excluded chemicals already present in the CPDB dataset it was possible to select as second dataset 739 compounds with experimental data on rats (either on male or female rats). Chemicals with experimental positive response for at least one sex were considered carcinogenic.

### Statistical analysis

Sensitivity and specificity were calculated as follow:

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

where, TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative

$$\text{Specificity} = \frac{TN}{TN+FP}$$

### Receiver Operating Characteristic

Receiver Operating Characteristic (ROC) curve was used to graphically compare performances obtained. ROC graph represents an plots 1-specificity versus sensitivity. The point (0,1) is the perfect classifier, as all positive and negative cases are predicted correctly. The points (0,0) and (1,1) represent a classifier that predicts all cases to be negative and positive, respectively, whereas (1,0) is associated with a classifier that predicts ever wrongly. The closer is the model to the point (0, 1) the better is.

## Results

The performance obtained by the software used for the whole set of compounds tested are shown in figure 1. Results are given in percentage. Multicase is the best software for the prediction of actual negatives (83% of specificity) while Caesar seems the best software for the prediction of actual positives (72% of sensitivity). In figure 2 the same results are given in a graphical way using a ROC curve.

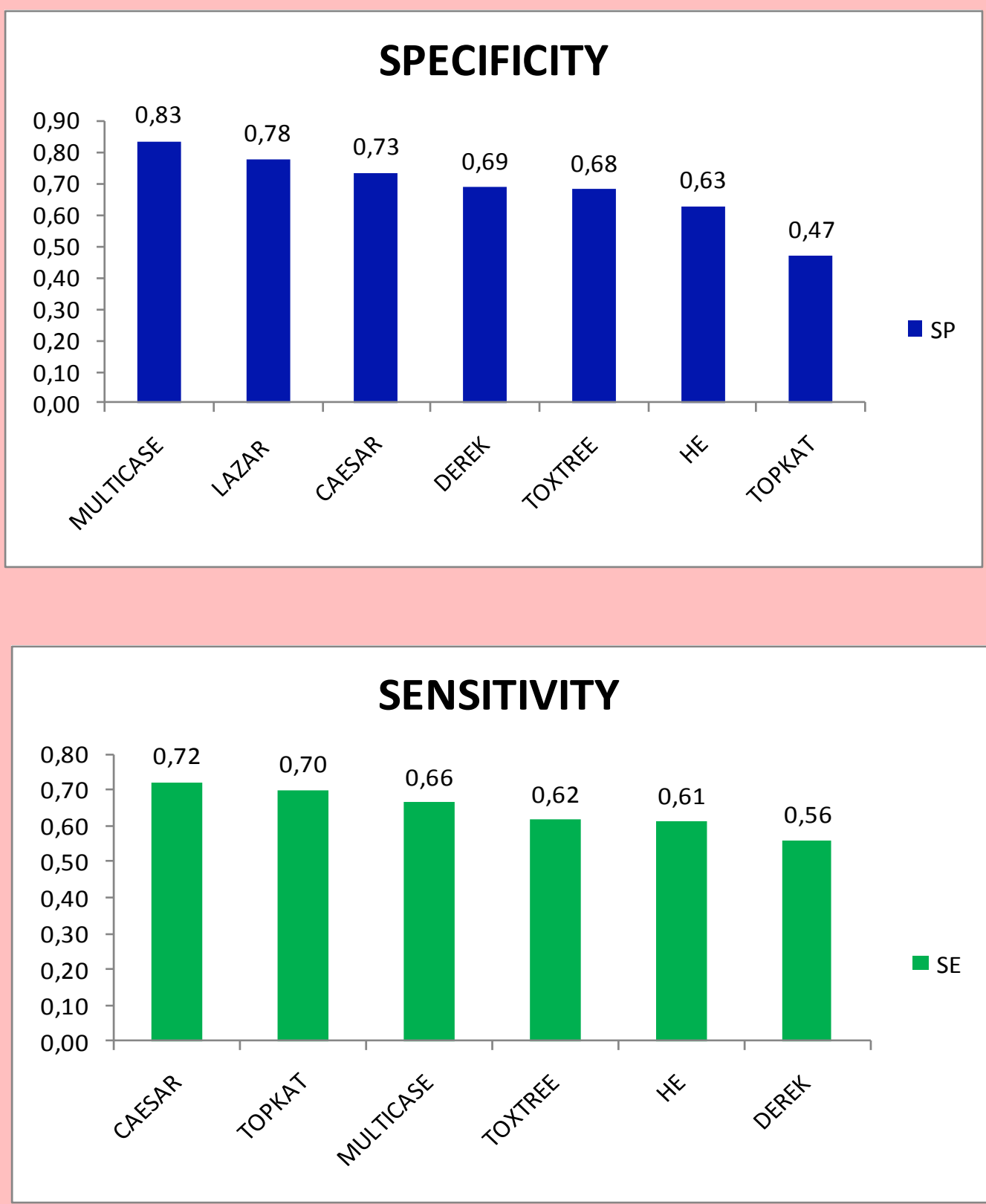


Fig.1

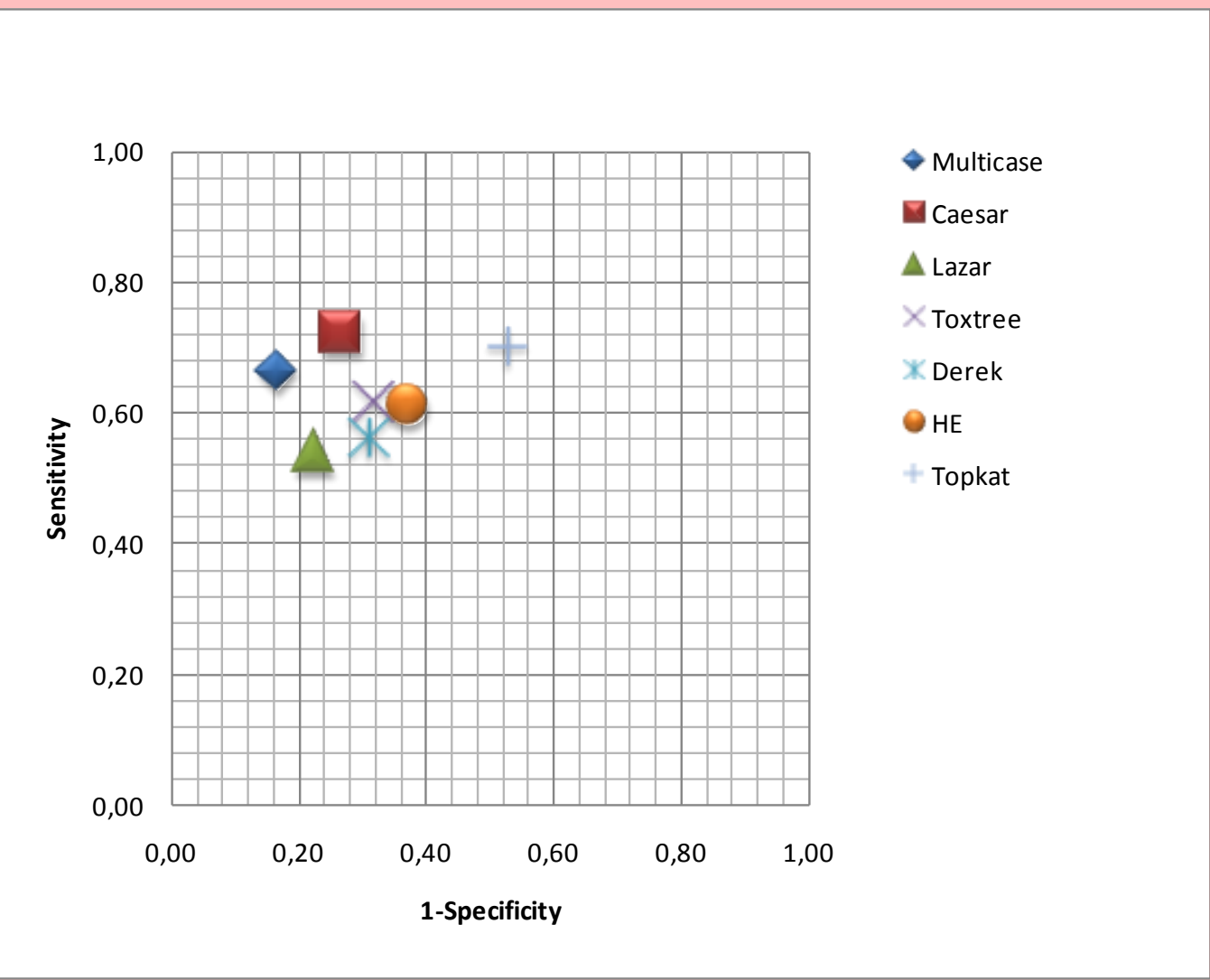


Fig.2

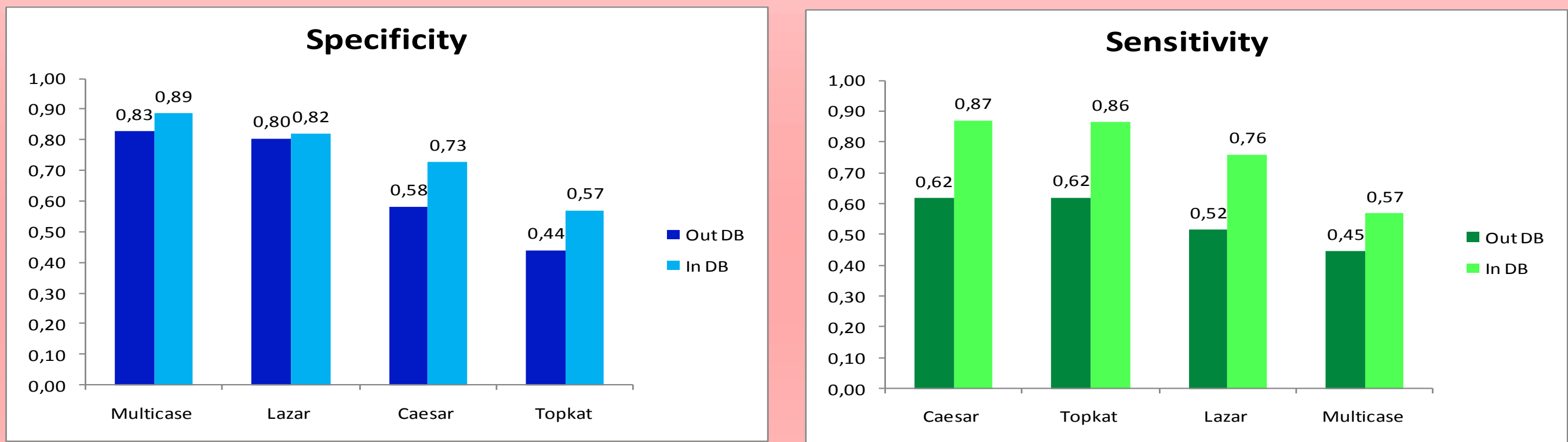


Fig.3

To better analyze these results data were split, whenever possible, considering separately chemicals that were already present in the internal database of each software or those that could be considered “new” chemicals for the software under evaluation. Four software were evaluated under this perspective (Multicase, Lazar, Caesar and Topkat). For the knowledge based Structural Alerts this kind of analysis is not applicable.

In this way we have a better information about the real predictive performance of each software (fig. 3). As expected, the quality of prediction for unknown compounds is inferior than the behavior observed for chemicals in each specific internal dataset. Again, Multicase has the highest specificity (83%) while Caesar and Topkat perform better for the prediction of actual positives (62% of sensitivity).

## Conclusions and future perspectives

Carcinogenicity is surely a very critical endpoint to evaluate with *in silico* methods and current performances of several software here tested are not optimal. Currently the *in-silico* approach can't be an alternative to the *in-vivo* test, but it may be useful as additional information on the possible carcinogenic effects as a pre-screening test or to prioritize chemicals for regulatory purposes.

Further analysis will focus on:

- The definition of the applicability domain or domain boundaries for the different models
- A deeper evaluation of model's behavior on specific chemical classes
- An evaluation of possible combination of different models to enhance performances

Financial support from  
ANTARES (LIFE08  
ENV/IT/000435) and OSIRIS  
(Contract no. GOCE-CT-2007-  
037017) projects is gratefully  
acknowledged

