QSAR model for carcinogenicity (v1.0)

ProtoTOX

ProtoTOX is a computational (*in silico*) tool focused on the prediction of endpoints related with the toxicity of chemical substances. It includes a variety of *in vitro* and *in vivo* tests in humans, animals, microorganisms and cell lines.

ProtoTOX mainly includes, but is not limited to, endpoints used by REACH, a European Union regulation, adopted to improve the protection of human health and the environment from the risks that can be posed by chemicals, while enhancing the competitiveness of the EU chemicals industry.

Endpoint

Human health effects: Carcinogenicity. Carcinogenicity Studies

Chemicals are defined as carcinogenic if they induce tumours, increase tumour incidence and/or malignancy or shorten the time to tumour occurrence.

Metrics

Training set

Experimental values	QSAR predictions	
	Non-carcinogen	**Carcinogen**
Non-carcinogen	286	14
Carcinogen	13	174

Validation set

Experimental values	QSAR predictions	
	Non-carcinogen	**Carcinogen**
Non-carcinogen	74	27
Carcinogen	20	44

Parameters	Training	Validation
Accuracy	0.94	0.72
Sensitivity / recall	0.93	0.69
Specificity	0.95	0.73
Precision	0.93	0.62
Negative predictive value	0.96	0.79
F-score	0.93	0.65
Matthews Correlation Coefficient	0.88	0.41
Critical Success Index	0.87	0.48
Area under the ROC	0.94	0.71