QSAR & Combinatorial Science QCS

Computer-Aided Prediction of Rodent Carcinogenicity by PASS and CISOC-PSCT

Alexey Lagunin[a], Dmitrii Filimonov[a], Alexey Zakharov[a], Wei Xie[b], Ying Huang[b], Fucheng Zhu[b], Tianxiang Shen[b], Jianhua Yao[b], Vladimir Poroikov[a]*

[a] Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, Pogodinskaya Str., 10, Moscow, 119121, Russia
*e-mail: vladimir.poroikov@ibmc.msk.ru
[b] Shanghai Institute of Organic Chemistry (SIOC), Chinese Academy of Sciences, 345, Lingling Rd, Shanghai, 200032, China

Keywords: Structure-activity relationships, Rodent carcinogens, PASS, CISOC-PSCT, Consensus model, Toxicology

Presented at the EuroQSAR Conference

Received: December 15, 2008; Accepted: March 31, 2009

DOI: 10.1002/qsar.200860192

Abstract

Computer-aided prediction of rodent carcinogenicity for the external test set consisting of 293 chemicals was performed by PASS (Prediction of Activity Spectra for Substances) and by CISOC-PSCT. The set included 64 carcinogens from ISS Carcinogens Data Bank and 229 noncarcinogens from the Prestwick Chemical Library. We calculated the accuracy of carcinogenicity prediction by PASS and CISOC-PSCT in apart, and by the two programs together (the consensus model). Sensitivity, specificity and accuracy (concordance) were calculated for the external test set by PASS (0.81, 0.74, 0.76), by CISOC-PSCT (0.36, 0.89, 0.77) and by the consensus model (0.69, 0.86, 0.83). Thus, taking into account the prediction results of two computer programs for rodent carcinogenicity the consensus model increases the accuracy of prediction.

1 Introduction

Dozens thousands of chemicals are used and many more are being synthesized today. It is necessary to have the efficient methods for the assessment of action of these compounds on the environment and human health. Experimental testing is both time-consuming and rather expensive. A major driving force behind the success and growth of (Q)SARs for the prediction of toxicity is the implementation of the European Union REACH (Registration, Evaluation and Authorization of Chemicals) legislation [1]. It is anticipated that, under the REACH legislation, (Q)SARs will be more extensively used to reduce the need for in vivo toxicological assessment of existing chemicals. Many SAR methods were developed to predict the carcinogenicity (MultyCase, DEREK, OncoLogic and the others) [2]. Nevertheless owing to the complexity of carcinogenic effect there is no adequate method at the present time. The most of SAR methods have the accuracy of prediction less than 70% [2]. Thus, there is a pressing need in accurate in silico methods to predict the carcinogenicity. Currently the (Q)SAR community continues the discussion on advantages and disadvantages of consensus prediction with the final result of prediction formed by integration of the results predicted by several models [3, 4]. It is considered that the consensus models decrease variance of individual models. All individual models contain varying amounts of predictions with uncertainty and the averaging of them leads to more reliable predictions [5]. Nevertheless the other authors suppose that the consensus models do not offer significant improvements over single regression models and their accuracy is usually less than that of the best model [4]. The above mentioned statements on the consensus models were made on the basis of QSAR models. The purpose of this investigation is the evaluation of the consensus prediction for SAR models of carcinogenicity prediction. With this aim in view we evaluated the accuracy of the consensus model created on the basis of PASS [6] and CISOC-PSCT [7] prediction results. For

Abbreviations: CISOC-PSCT: computerized information system for organic chemistry – predictive system for carcinogenic toxicity; **IAP**: independent accuracy of prediction; **PASS**: prediction of activity spectra for substances; **SAR**: structure-activity relationship.

Supporting information for this article is available on the WWW under www.qcs.wiley-vch.de

more objective validation of prediction we created the external test set consisted of compounds not included into the training sets.

2 Materials and Methods

2.1 PASS (Prediction of Activity Spectra for Substances)

PASS is a computer program for the evaluation of general biological potential in a molecule on the basis of its structural formulae [6]. The list of predictable biological activities contains 3300 types (PASS 2007 version) including the main and side pharmacological effects (antianginal, analgesic, antiviral etc.), mechanisms of action (alpha-2 adrenaline antagonist, angiotensin-converting enzyme inhibitor, serotonin uptake inhibitor, etc.), specific toxicities (mutagenicity, carcinogenicity, teratogenicity, etc.) and metabolic terms (CYP1A2 substrate, CYP2D6 inhibitor, CYP3A4 inducer, etc.). The mean independent accuracy of prediction (IAP) calculated by leave-one-out cross-validation procedure is 94%. Earlier we have already shown the ability of PASS to predict rodent carcinogens [8]. PASS 2007 version includes 1210 of the known carcinogens. The other 116122 compounds from PASS training set are considered as noncarcinogens. IAP of carcinogenic prediction calculated by leave-one-out cross-validation is 92.8%. The results of PASS prediction are represented as a list of probable biological activity types, for which the probability to be active (Pa) and the probability to be inactive (Pi) is calculated. In the present study we used the difference between Pa and Pi as a value of PASS carcinogenicity prediction.

2.2 CISOC-PSCT

CISOC-PSCT is a SAR-based carcinogenicity prediction system, which consists of two principal parts: construction of the structure-carcinogenicity model and prediction of toxicity using this SAR model [7]. The training set includes 2738 of carcinogenic compounds extracted from the SYMYX MDL Toxicity database [9] and 4130 of noncarcinogenic compounds selected from the MDL CMC database [10]. The results of CISOC-PSCT prediction are represented as the list of values for Predictability, ToxicPossibility and ToxicImpossibility. Predictability reflects the belonging of a particular molecule to the model's applicability domain. It varies from 0% to 100% and compounds with predictability more than 50% are considered as appropriate for the prediction. Predictability in CISOC-PSCT is close to the distance to model [11]. ToxicPossibility and ToxicImpossibility mean a probability of presence and absence of carcinogenic effect in compounds. They vary from 0 to 1. In the present study we used the difference between ToxicPossibility (Tp) and ToxicImpossibility (Ti) as a value of CISOC-PSCT carcinogenicity prediction.

2.3 External Test Set

The external test set consisted of 293 chemicals and included 64 carcinogens and 229 noncarcinogens. The carcinogens were retrieved from ISS Carcinogens Data Bank [12]. The Data Bank on Carcinogens (Banca Dati Cancerogeni, BDC) is a factual data bank, available from the Istituto Superiore di Sanità website, aimed at supporting the risk management decision making by central and local administrators. Noncarcinogens were selected from the Prestwick Chemical Library [13]. The Prestwick Chemical Library contains 1120 small molecules: 90% of which are marketed drugs and 10% are bioactive alkaloids or related substances. All these molecules were selected for their high chemical and pharmacological diversity as well as for their known bioavailability and safety in human. In this study only those compounds were selected that were absent in CISOC-PSCT training set. Because all compounds from the external test set were initially in PASS training set they were excluded from PASS training set during the study. The charged structures were transformed to the neutral form. Only compounds with CISOC-PSCT Predictivity values more than 50% were included into the test set. Such compounds fall to the applicability domain of the model. The structures of compounds from the external test set are available in supplements.

3 Results and Discussion

The prediction of carcinogenicity for compounds from the external test set was performed by PASS and CISOC-PSCT on the basis of their training sets. The compounds having positive predicted values are considered as carcinogens and those with the negative predicted values are considered as noncarcinogens. Table 1 shows the values of sensitivity, specificity and accuracy of carcinogenicity prediction, and also the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) predicted for each method used in apart and with the two programs together (consensus prediction).

Table 1. Accuracy of prediction for the external test set using PASS, CISOC-PSCT, and consensus model. TP: true positive (the number of carcinogens predicted as carcinogens); TN: true negative (the number of noncarcinogens predicted as noncarcinogens); FP: false positive (the number of noncarcinogens predicted as carcinogens); FN: false negative (the number of carcinogens predicted as noncarcinogens).

	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
CISOC-PSCT	23	204	25	41	0.36	0.89	0.77
PASS	52	170	59	12	0.81	0.74	0.76
Consensus	44	198	31	20	0.69	0.86	0.83

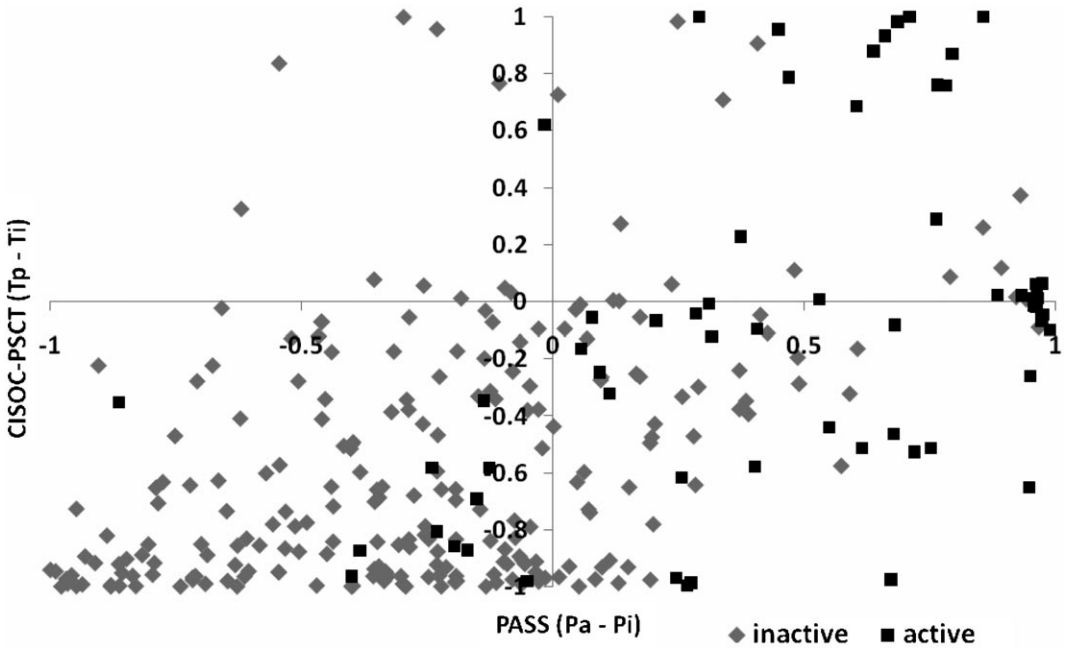

Figure 1. Comparison of PASS and CISOC-PSCT prediction results.

Carcinogenicity prediction by PASS displayed high sensitivity (0.81) and low specificity (0.74), while CISOC-PSCT showed high specificity (0.89) and low sensitivity (0.36). On the basis of these results one may suggest that the results of PASS and CISOC-PSCT complement each other. On the other hand, low correlation between prediction results obtained by PASS and CISOC-PSCT also promotes their joint application. Comparison of prediction results between PASS and CISOC-PSCT are shown in Figure 1. X axis corresponds to the PASS prediction results; Y axis corresponds to the CISOC-PSCT prediction results. Their correlation is about 18% for carcinogens and 20% for noncarcinogens. Thus, these data may be used as independent variables for further modeling and complementing each other.

Based on two previous prerequisites (significant difference in specificity and sensitivity of both programs and low correlation between their prediction results), we decided to create the consensus prediction. The consensus prediction is based on a simple unweighted consensus model [3, 4] that calculates the sum of prediction results of PASS and CISOC-PSCT. Positive value of the consensus model means that compound is a carcinogen, while the negative one of the consensus model means that compound is not carcinogen. The analysis of correlation between the prediction values of PASS and consensus model shows that for carcinogens it is 0.62 and 0.71 for noncarcinogens. Correlation between the prediction values of CISOC-PSCT and consensus model for carcinogens is 0.79 and 0.74 for noncarcinogens.

We have also calculated the sensitivity, specificity and accuracy for the results of consensus prediction (Table 1). It turned out that the accuracy of consensus prediction is the highest (0.83) in comparison to CISOC-PSCT (0.77) and to PASS (0.76). Sensitivity of the consensus model (0.69) is higher than that of CISOC-PSCT (0.36) and less than that at PASS (0.81). Specificity of the consensus model (0.86) is higher than that of PASS (0.74) and less than that at CISOC-PSCT. Sensitivity depends on the number of true positive and false negative prediction (TP/(TP + FN)). Specificity depends on the number of true negative and false positive prediction (TN/(TN + FP)). The total number of wrong predictions at consensus model consists of 51 predictions (31 false positives and 20 false negatives). It is far less than 66 wrong predictions at CISOC-PSCT (25 false positives and 41 false negatives) and 71 wrong predictions at PASS (59 false positives and 12 false negatives). Both the numbers of false negative and false positive predictions are very important for the environmental studies. To reveal all harmful compounds false negative predictions should be close to zero. On the other hand false positive predictions increase the number of animal experiments. Therefore, the developed methods should provide a possibility to optimize the rate of false negatives and false positives. Such possibility of the studied methods could be realized using the curves for sensitivity and specificity depending on the prediction results (Fig. 2). All values demonstrated in Table 1 were calculated with zero threshold.

In most cases, the curves of sensitivity and specificity of the consensus model lie between those of sensitivity and specificity of PASS and CISOC-PSCT that lead to more balanced results of carcinogenicity prediction. Sensitivity and specificity may be changed by the threshold of predic-

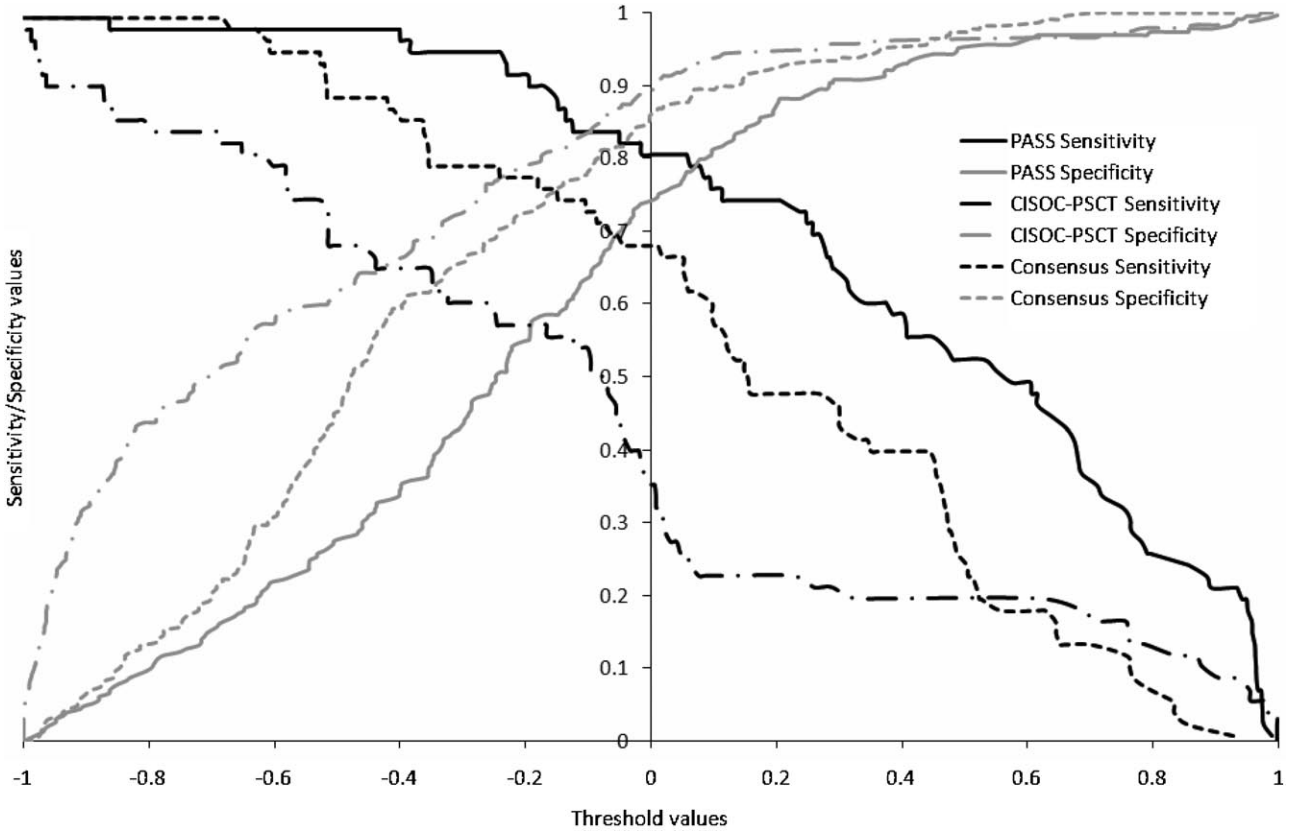

Figure 2. Sensitivity and specificity of PASS, CISOC-PSCT and Consensus model depending on the estimates of carcinogenicity prediction.

tion values. For all studied models, shifting of the threshold to the left increases the sensitivity and decreases the specificity. On the contrary, threshold shifting to the right leads to specificity increase and sensitivity decrease.

Reliability and nonreliability are the other important characteristics of prediction results. The molecules with nonreliable predictions should be experimentally tested. In contrast to PASS, CISOC-PSCT outputs the value of Predictability that displays the adequacy of the model to predict the appropriate compound. It was considered that if Predictability is less than 70% the results of prediction are inconclusive. There are four compounds in the test set with Predictability less than 70% (No 11 for carcinogens; No 60, 104 and 121 for noncarcinogens), see supplementary materials. On the other hand, the prediction results may be considered inconclusive when the value of probability to be carcinogen is close to the value of probability to be noncarcinogen. In this case the difference between these values is close to zero. We calculated how many compounds have prediction values fall to the interval [−0.05:0.05]. Twenty four compounds with PASS prediction results fall to the interval. For CISOC-PSCT and consensus model this number is 18 and 14 compounds, respectively. Thus, consensus model may decrease the number of compounds with inconclusive prediction.

4 Conclusions

"Which is better: the best (Q)SAR model or the consensus?" This question often arises at the analysis of several (Q)SAR models. In this study we have found the parameters of SAR models that may help to answer this question. In our opinion the difference between sensitivity and specificity of initial models and their low correlation may be the reason for creation of the consensus model. Thus, on the basis of these prerequisites we used the well known consensus approach to combine the results of rodent carcinogenicity predictions made by PASS and CISOC-PSCT programs. The consensus model increased the accuracy of prediction from 0.77 (the best prediction accuracy obtained by a single computer program) to 0.83. It confirms the reasonableness of the consensus model based on the results of PASS and CISOC-PSCT for rodent carcinogenicity prediction.

5 Acknowledgements

The study was supported by RFBR-NSFC Grant #06-03-39015 and NSFC-RFBR Grant #20711120201.

6 References

[1] *White Paper on a Strategy for a Future Chemical Policy* (COM(2001)-88final); Commission of the European Communities: Brussels, Belgium **2001.** Accessible from: http://europa.eu/scadplus/leg/en/lvb/l21275.htm.

[2] R. Benigni, R. Zito, *Mutat. Res.* **2004**, *566*, 49 – 63.

[3] M. Hewitt, M. T. Cronin, J. C. Madden, P. H. Rowe, C. Johnson, A. Obi, S. J. Enoch, *J. Chem. Inf. Model.* **2007**, *47*, 1460 – 1468.

[4] H. Zhu, A. Tropsha, D. Fourches, A. Varnek, E. Papa, P. Gramatica, T. Oberg, P. Dao, A. Cherkasov, I. V. Tetko, *J. Chem. Inf. Model.* **2008**, *48*, 766 – 784.

[5] J. S. Geman, E. Bienenstock, R. Doursat. *Neural Comp.* **1992**, *4*, 1 – 58.

[6] V. Poroikov, D. Filimonov, in: *Predictive Toxicology* (Ed: C. Helma), Taylor & Francis, LLC, Boca Raton **2005**, pp. 459 – 478.

[7] Q. Liao, J. H. Yao, F. Li, S. G. Yuan, J. P. Doucet, A. Panaye, B. T. Fan, *SAR QSAR Environ Res.* **2004**, *15*, 217 – 235.

[8] A. A. Lagunin, J. C. Dearden, D. A. Filimonov, V. V. Poroikov, *Mutat. Res.* **2005**, *586*, 138 – 146.

[9] SYMYX MDL Toxicity Database: http://www.symyx.com/products/databases/bioactivity/toxicity/index.jsp

[10] SYMYX MDL Comprehensive Medicinal Chemistry (CMC) Database: http://www.symyx.com/products/databases/bioactivity/cmc/index.jsp

[11] I. V. Tetko, I. Sushko, A. K. Pandey, H. Zhu, A. Tropsha, E. Papa, T. Oberg, R. Todeschini, D. Fourches, A. Varnek, *J. Chem. Inf. Model.* **2008**, *48*, 1733 – 1746.

[12] R. Binetti, F. Ceccarelli, F. M. Costamagna, A. D'Angiolini, A. Fabri, M. Ferri, G. Riva, P. Roazzi, D. Trucchi, I. Marcello, *Ann. Ist. Super. Sanità.* **2008**, *44*, 31 – 42.

[13] Prestwick Chemical Library: http://www.prestwickchemical.fr/.

