In-silico Prediction of Human Intestinal Absorption and Human Oral Bioavailability

Aixia Yan *

State Key Laboratory of Chemical Resource Engineering, Department of Pharmaceutical Engineering, P.O. Box 53, Beijing University of Chemical Technology, 15 BeiSanHuan East Road, Beijing 100029, P.R. China.

E-mail address: aixia_yan@yahoo.com

Prediction of Human Intestinal Absorption

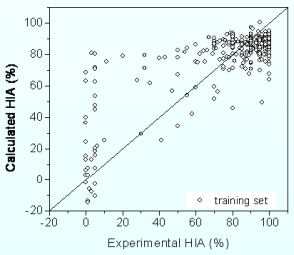
Human Intestinal Absorption (HIA) is an important property in drug design, which is not only one of the key steps during the drugs transporting to their targets but also influences bioavailability. Utilization of drugs in the human body is such a complicated process that HIA can hardly be analyzed precisely by statistical models, so in our work, QSAR (Quantitative Structure Activity Relationships) models for the prediction of human intestinal absorption (HIA) were built with molecular descriptors calculated by ADRIANA.Code, Cerius^2 and a combination of them. A dataset of 552 compounds covering a wide range of current drugs was investigated. A Genetic Algorithm feature selection method was applied to select proper descriptors. A Kohonen's self-organizing Neural Network (KohNN) map was used to split the whole dataset into a training set including 380 compounds and a test set consisting of 172 compounds. Nine descriptors were selected by a combination of ADRIANA.Code and Cerius^2 descriptors. The quantitative models for prediction of HIA were built by PLS (Partial Least Square) and SVM (Support Vector Machine). Nine combined descriptors from Adriana. Code and Cerius^2: Nrule5, LogP, Nrot, Jurs-FNSA-3, Jurs-RPCG, Hdon, LogS, MW, Acorr_Sigchg_3.

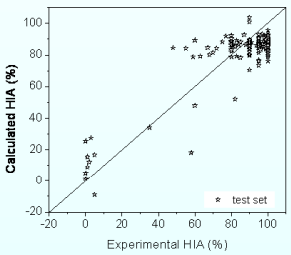
Self-organizing neural network classification model of HIA

compounds: 552
occupancy: 347
conflicts: 61
correctness: 89%

HIA
Low (blue): 0-29%
Middle (yellow): 30%-79%
High (red): 80%-100%

Prediction of HIA with nine combined descriptors by SVM





training set
r=0.81, s=12.50, MAE=12.74, n=380

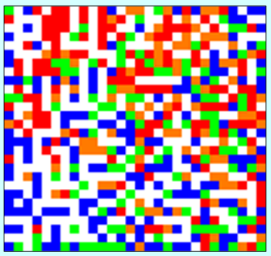
test set
r=0.88,s=9.14, MAE=10.71, n=172

Prediction of Human Oral Bioavailability

Bioavailability represents the percentage of an oral dose which is able to produce a pharmacological activity, i.e. the fraction of the oral dose that reaches the arterial blood in an active form. Bioavailability is related to several factors, such as gastrointestinal transition and absorption, intestinal membrane permeation, and intestinal/hepatic first-pass metabolism, etc.

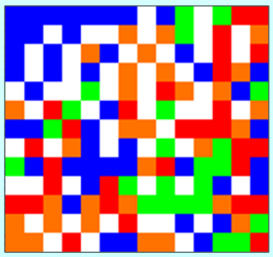
In this work, (1) Classification of the bioavailability of 772 compounds was done by Kohonen's self-organizing neural network (KohNN) using the SONNIA. (2) Classification of the bioavailability of 161 compounds with experimental HIA values were performed by a KohNN. (3) The bioavailability of 51 sulfonamide drugs was predicted by SVM. (4) The bioavailability of 29 β-lactam drugs was predicted by SVM. (5) The bioavailability of 58 central nervous system (CNS) drugs was predicted by SVM.

Self-organizing neural network classification model of Bioavailability



compounds: 772
occupancy: 516
conflict: 287
correctness: 62.8%

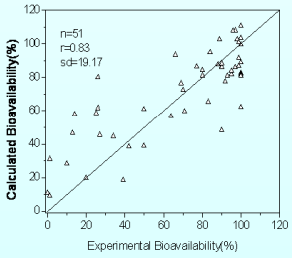
bioavailability
blue: 0-25%
green: 25%-50%
orange:50%-75%
red:75%-100%



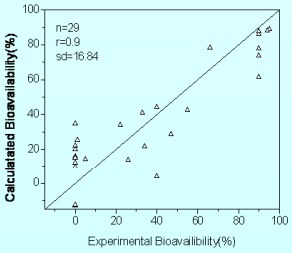
compounds: 161
occupancy: 125
conflict: 34
correctness: 79%

bioavailability (%)
blue: 0-25%
green: 25%-50%
orange: 50%-75
red: 75%-100%

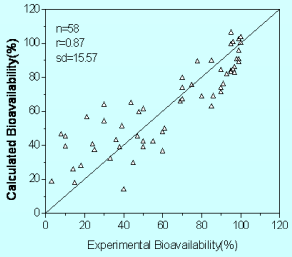
Prediction of bioavailability by SVM



51 sulfonamide drugs
Select 4 descriptors from ADRIANA.Code.
Acorr_TotChg_6,
Acorr_LpEN_9,
Acorr_PiEN_8,
Acorr_LpEN_4.



29 β-lactam drugs
Select 5 descriptors from ADRIANA.Code.
Acorr_SigChg_4,
Acorr_LpEN_9,
Acorr_TotChg_4,
Acorr_PiChg_7,
Acorr_SigChg_5.



58 central nervous system drugs
Select 4 descriptors from ADRIANA.Code.
Acorr_Ident_3,
Acorr_TotChg_2,
Acorr_Ident_7,
Acorr_PiEN_9.

References

1. Wang, Z., Yan, A.X.*, Yuan, Q.P., Gasteiger, J. Explorations into Modeling Human Oral Bioavailability, Euro. J. Med. Chem. , 2008, 43, 2442-2452.
2. Yan, A.X.*, Wang, Z., Cai, Z.Y. Prediction of Human Intestinal Absorption by GA Feature Selection and Support Vector Machine Regression, Int. J. Mol. Sci. 2008, 9, 1961-1976.

Acknowledgments: This work was supported by the Natural Science Foundation of China (20605003, 20975011), and the Fundamental Research Funds for the Central Universities (ZZ0911) of Beijing University of Chemical Technology. The authors thank Prof. J. Gasteiger and the Molecular Networks GmbH, Erlangen, Germany for providing the programs ADRIANA.Code and SONNIA for our scientific work. The authors also thank Dr. Tingjun Hou for providing the Human intestinal absorption and the bioavailability datasets, and thank Prof. Abraham for providing 241 compounds with SMILES files and HIA property values.