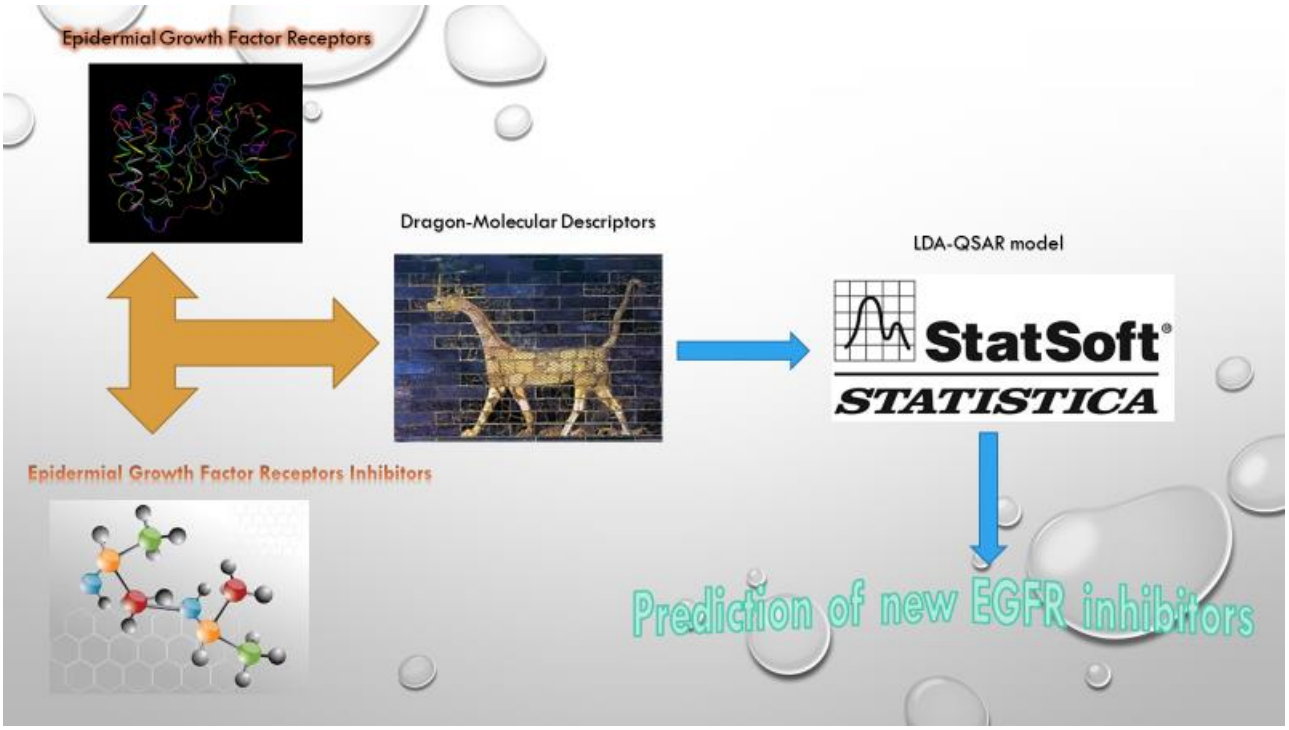
MOL2NET, International Conference Series on Multidisciplinary Sciences
http://sciforum.net/conference/mol2net-03

A novel QSAR model to predict epidermial growth factor inhibitors

Riccardo Concu1, Maria Natalia DiasSoeiro Cordeiro1*

1 REQUIMTE, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal.

** ric.concu@gmail.com*



Abstract.

Over-expression of the Epidermial Growth Factor Receptor (EGFR) is usually present in more than 90% of the Head and Neck Squamous Cell Cancer (HNSCC), due to this the selection of more selective and powerful inhibitors is a major point to threat this type of cancer. In fact, has been demonstrated that this over-expression is responsible of a more aggressive disease, increased resistance to chemotherapy and radiotherapy, increased metastasis, inhibition of apoptosis, promotion of neoplastic angiogenesis, and, finally, poor prognosis and decreased survival. Computational methods are a major tool while looking for new EGFR inhibitors since should help researchers selecting new and enhanced inhibitors in this area. In this contest, Quantitative structure activity relationship (QSAR) is one of the most and widely used computational technique to select new EGFR inhibitors. Here we will present a new QSAR approach aimed at the prediction of new EGFR inhibitors drugs using 1D molecular descriptors.

Introduction

Head and neck squamous cell carcinoma (HNSCC) is accounted for more than 500,000 cases every ear worldwide, according to the more recent statistics[1]. HNSCC is the sixth most common cancer in the USA, accounting for 2% of all new cancers diagnosed and 1% of all cancer deaths. Epidermal Growth Factor Receptor (EGFR) overexpression has been detected in over the 90% of the HNSCC[2-3] and is usually associated with more aggressive disease, increased resistance to chemotherapy and radiotherapy, increased metastasis, inhibition of apoptosis, promotion of neoplastic angiogenesis, and, finally, poor prognosis and decreased survival[4-8]. This family comprised four related receptors: the EGFR (ErbB1/EGFR/HER1), ErbB2 (HER2/*neu*), ErbB3 (HER3) and ErbB4 (HER4) which are all target for the treatment of the HNSCC. In this short communication we will present a novel 1D-QSAR model based on Dragon descriptors aimed at the prediction of novel compounds against the EGFR-3.

Materials and Methods

A total of 345 molecules were downloaded from the chemical repository Chembl with their correspondent IC50 values which were used as dependent variable. The IC50 values were reported in the nM scale in a range between 0.2 and >1000. The compounds with an IC50 higher than 200nM were labeled as inactive while the others were assigned to the active class. Using the Dragon 7.0© software we have calculated more than 2600 descriptors. Using a forward stepwise procedure we selected a total of 14 MD to build our final model. The final QSAR model was built using the Linear Discriminant Analysis (LDA) integrated in the STATISTICA® software. Finally, the model was validated using a classical cross validation model method, the Mathews correlation coefficient (MCC)[9].

Results and Discussion

The final LDA-QSAR model is represented by the following equation:

$$EGFRi = ZM1Per * -16.5 + ZM1MulPer * 12.76 + ZM2V * 14 + ZM2Per * -8.74 + DBI * 1.48 + GNar * -0.97 + Pol * -1.09 + MSD * -4.45 + AECC * 4.73 + Wap * 1.12 + PW4 * 1.77 + PW5 * -1.36 + BAC * -1.45$$

This model able to correct classify 130 out of 150 active inhibitors (SPEC=84.41%) and 141 out of 165 (SENS=87.57%) for an overall accuracy of 86.03%. More in depth, in the training set the model shows a SPEC of 83.62% (97 cases out of 110) and a SENS of 89.34 (109 out of 128); in the validation set the model correct classify 33 out of 40 (SENS=86.84%) and 32 out 37 (SPEC=82.05%). The MCC value is 0.72 which is in line with the other statistics which clearly indicates that the model is robust. In fact, as above reported, the performance of the model in the validation series is in line with the statistics in the training set. In addition, the MCC value also confirm the robustness of the model avoiding any kind of overfitting problem. Thus, the present model can be used to predict new EGFR inhibitors. Regarding the descriptors, we used only topological indices which are well-known MD able to codify the information within the chemical structure of the compounds that then, should be used to build up robust models.

Conclusions

The herein presented model is an LDA-QSAR model based on topological indices aimed at the prediction of new EGFR inhibitors. There is a strong need of computational tools in this area considering that, HNSCC is accounted for more than 500,000 cases every ear worldwide and is usually associated with considerable morbidity and mortality; the HNSCC is responsible for more than 300,000 deaths every year for instance. In this contest, QSAR models like the herein presented may play an important role while looking for new and powerful EGFR inhibitors.

ACKNOWLEDGEMENTS

This work received financial support from Fundação para a Ciência e a Tecnologia (FCT/MEC) through national funds and co-financed by the European Union (FEDER funds) under the Partnership Agreement PT2020, through projects UID/ QUI/50006/2013, POCI/01/0145/FEDER/007265, and NORTE-01-0145-FEDER-000011 (LAQV@REQUIMTE). RC acknowledges also FCT and the European Social Fund for financial support (Grant SFRH/BPD/80605/2011). To all financing sources the authors are greatly indebted.

References

1. Global Burden of Disease Cancer, C., Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: A systematic analysis for the global burden of disease study. *JAMA Oncology* **2017,** *3* (4), 524-548.
2. Goke, F.; Franzen, A.; Hinz, T. K.; Marek, L. A.; Yoon, P.; Sharma, R.; Bode, M.; von Maessenhausen, A.; Lankat-Buttgereit, B.; Goke, A.; Golletz, C.; Kirsten, R.; Boehm, D.; Vogel, W.; Kleczko, E. K.; Eagles, J. R.; Hirsch, F. R.; Van Bremen, T.; Bootz, F.; Schroeck, A.; Kim, J.; Tan, A. C.; Jimeno, A.; Heasley, L. E.; Perner, S., FGFR1 Expression Levels Predict BGJ398 Sensitivity of FGFR1-Dependent Head and Neck Squamous Cell Cancers. *Clin Cancer Res* **2015,** *21* (19), 4356-64.
3. Mehra, R.; Cohen, R. B.; Burtness, B. A., The role of cetuximab for the treatment of squamous cell carcinoma of the head and neck. *Clin Adv Hematol Oncol* **2008,** *6* (10), 742-50.
4. Rieke, D. T.; Klinghammer, K.; Keilholz, U., Targeted Therapy of Head and Neck Cancer. *Oncol Res Treat* **2016,** *39* (12), 780-786.
5. Nakano, T.; Yamamoto, H.; Nakashima, T.; Nishijima, T.; Satoh, M.; Hatanaka, Y.; Shiratsuchi, H.; Yasumatsu, R.; Toh, S.; Komune, S.; Oda, Y., Molecular subclassification determined by human papillomavirus and epidermal growth factor receptor status is associated with the prognosis of oropharyngeal squamous cell carcinoma. *Hum Pathol* **2016,** *50*, 51-61.
6. Pollock, N. I.; Wang, L.; Wallweber, G.; Gooding, W. E.; Huang, W.; Chenna, A.; Winslow, J.; Sen, M.; DeGrave, K. A.; Li, H.; Zeng, Y.; Grandis, J. R., Increased Expression of HER2, HER3, and HER2:HER3 Heterodimers in HPV-Positive HNSCC Using a Novel Proximity-Based Assay: Implications for Targeted Therapies. *Clin Cancer Res* **2015,** *21* (20), 4597-606.
7. Machiels, J. P.; Schmitz, S., Epidermal Growth Factor Receptor Inhibition in Squamous Cell Carcinoma of the Head and Neck. *Hematol Oncol Clin North Am* **2015,** *29* (6), 1011-32.
8. Nijkamp, M. M.; Span, P. N.; Terhaard, C. H.; Doornaert, P. A.; Langendijk, J. A.; van den Ende, P. L.; de Jong, M.; van der Kogel, A. J.; Bussink, J.; Kaanders, J. H., Epidermal growth factor receptor expression in laryngeal cancer predicts the effect of hypoxia modification as an additive to accelerated radiotherapy in a randomised controlled trial. *Eur J Cancer* **2013,** *49* (15), 3202-9.
9. Boughorbel, S.; Jarray, F.; El-Anbari, M., Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PLoS One* **2017,** *12* (6), e0177678.