[P25] Docking-based classification models for estrogenic activity prioritization based on high-quality experimental data

Daniela Trisciuzzi[1], Domenico Alberga[2], Kamel Mansouri[3], Richard Judson[3], Giuseppe Felice Mangiatordi[1], Orazio Nicolotti[1]

[1]Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari 'Aldo Moro', Via E. Orabona, 4, I-70126, Bari, Italy.
[2]Dipartimento Interateneo di Fisica 'M. Merlin', Università degli Studi di Bari 'Aldo Moro', INFN, Via E. Orabona, 4, I-70126, Bari, Italy.
[3]National Center for Computational Toxicology, US Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park, 27711, NC, USA.

Exploratory toxicology is a new emerging research area whose ultimate mission is the protection of human health and environment from risks posed by chemicals. In this regard, the ethical and practical limitation of *in vitro* and *in vivo* approaches has encouraged the promotion of non-test methods for the fast screening of huge collections of chemicals available on the market [1]. The herein presented study is framed in this context and has been developed within the large-scale modeling project CERAPP (which stands for Collaborative Estrogen Receptor Activity Prediction Project) supervised by the US Environmental Protection Agency (US EPA) in partnership with 17 research groups in the USA and Europe [2]. Our task was the development of classification models discriminating estrogenic chemicals on the basis of campaigns of molecular docking, a structure-based approach largely used in drug discovery programs. In this respect, using eight estrogen receptor X-ray solved crystals available from the Protein Data Bank, we derived 24 reliable docking-based classification models trained on a US EPA collection of 1677 chemicals having high quality estrogenic experimental data [3]. Model performances were challenged by considering the early Enrichment Factor at 1% (EF1%) and the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. In addition, the Negative Predictive Values (NPVs) at two different sensitivity (SE) thresholds (SE=0.25 and SE=0.75) were also computed. The obtained docking-based models returned the following statistics: $0.63<AUC<0.72$, $2.5<EF_{1\%}<6.2$, $87.7<NPV<88.7$ (SE=0.25) and $89.7<NPV<94.2$ (SE=0.75). Within CERAPP, our models were thus successfully employed to prioritize a large chemical universe of >32,000 man-made chemicals for further testing.

Bibliography:

[1] N. O.; B. E. Drug Discov. Today 19 (2014) 1757–1768.
[2] M. K.; A. A. (2016). In press.
[3] J. R.S.; M. F.M. Toxicol. Sci. 148 (2015) 137-54.