QSAR Models for Identifying Potential Estrogenic Activity of Phenolic Compounds

Kamel Mansouri[1], Regina Politi[2], Sherif Farag[2], Eugene Muratov[2], Alexander Tropsha[2], Richard Judson[1], Qingda Zang[3], David Allen[3], Warren Casey[4], Nicole Kleinstreuer[3]

[1]EPA/ORD/NCCT, RTP, NC, USA, [2]UNC-CH, Chapel Hill, NC, USA; [3]ILS/NICEATM, RTP, NC, USA; [4]NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

Computational tools can be used for efficient evaluation of untested chemicals for their ability to disrupt the endocrine system. We have employed previously developed global QSAR models that were trained and validated on the ToxCast/Tox21 ER assay data for virtual screening of a large library of ~30,000 industrial chemicals. Compounds predicted to have estrogen receptor (ER) binding and agonist activity included a high proportion of phenolic compounds, consistent with prior knowledge on the influence of this structural moiety on chemical interaction with the ER. However, global models did not accurately predict specific activity and relative potency of various phenols. Thus, we constructed local QSAR models focused on this chemical category only in the ToxCast/Tox21 data. Models developed with random forest and partial least squares discriminant analysis methods were trained and tested on the data from both ToxCast/Tox21 assays and well-curated literature sources. Local models consistently yielded higher balanced accuracies, sensitivity, and specificity than global models as evaluated on the external test sets. Our results suggest that these models can be used as reliable support tools for evaluating the endocrine disrupting potential of environmental phenolic chemicals. *This work does not reflect EPA policy. This project was funded in whole or in part with Federal funds from the NIEHS, NIH under Contract No. HHSN273201500010C.*

Keywords:

1. Endocrine disruptor chemicals

2. Computational toxicology

3. ToxCast/Tox21