A cost-effective approach to new drug discovery and chemical safety

Modern lifestyle is highly dependent on the use of thousands of chemicals in the form of life-saving drugs as well as specialty substances like pesticides, herbicides, cosmetics, solvents, etc. Today the discovery of a new drug costs approximately US $ 1.0- 1.9 billion. One of the main causes of this skyrocketing cost is that the pharmaceutical companies have to make and screen many thousands of chemicals before they can find one useful drug. Commuter models based on statistics and mathematical chemistry can help in bringing down this exorbitant price of drug discovery by forecasting ahead of time which of the large number of candidate chemicals are promising and which ones have very little chance of success. In the area of environmental protection, the regulatory agencies need to know a large number of properties of the chemicals in order to carry out a proper hazard assessment of these substances and their daughter products. Because laboratory testing of a large list of chemicals can be prohibitively costly, properties computed from their structure can be used in lieu of laboratory test data. This decreases the price of drug discovery and chemical regulation. Mathematical chemistry can help us in generating such data in a cost effective manner.

Beginning at the middle of the twentieth century, the branch of science called mathematical chemistry, or more accurately, discrete mathematical chemistry, has made substantial progress during the past half century or so. A plethora of methods have been developed for the characterization of molecular structure and computation of properties. If we look stark differences in the field between 1950 and now, it may look like what Van Leeuwenhoek saw when he increased the magnification capacity of the microscope from 20x-30x to 270x based on the primitive microscope of Zacharias Jansen. But when we have hundreds or thousands of computed properties we may become confused like in the story of the elephant and the blind men. Statistics can help us in managing this quagmire efficiently.

The cost effective approach involving the strategic generation of some reasonable amount of lab data, mathematical chemistry, and statistical analysis ideally may proceed as follows:

> Test in the laboratory a properly selected number of samples from the large list of candidate chemicals.
> Calculate the mathematical descriptors of the molecules under investigation
> Develop robust and validated predictive models to estimate properties of interest
> Use the validated models to estimate properties of substances for which laboratory testing is expensive or unavailable.

At the current stage of technological development an approach outlined above has been possible because of two important factors: a) Novel applications of mathematics to chemical and biological systems, and ii) Availability of increasingly more computer power following Moore’s law and

associated software whereby hypothesis driven as well as discovery oriented research can be carried out within a reasonable time frame. This trend of research has led to numerous useful applications to scientifically, socially, technologically, and economically important areas such as drug discovery, protection of human as well as ecological health. The article by Basak and Majumdar gives a step-by-step road map of using mathematical and statistical methods in these areas.

Subhash C. Basak and Subhabrata Majumdar

Publication

The Importance of Rigorous Statistical Practice in the Current Landscape of QSAR Modelling.
Basak SC, Majumdar S
Curr Comput Aided Drug Des. 2015