**Assignment 1**

**Q1. Explain: Chemoinformatics and its history**

* “The use of information technology and management has become a critical part of the drug discovery process. Chemoinformatics is the mixing of those information resources to transform data into information and information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and organization”.
* Above is the definition for chemoinformatics given by Dr. Brown in 1998.
* The current scope of chemoinformatics has exceeded past just drug discovery.
* Although the term chemoinformatics is more than two decades old, the roots of the field date back much earlier.
* The first and most important journal covering this field was established in 1961 named *Journal of Chemical documentation* (JCD) which was then renamed to *Journal of chemical Information and Modelling* (JCICS) to better reflect the contents of the journal which was papers related to chemoinformatics.
* Chemoinformatics is a field of IT that uses computers to facilitates the following for large quantities of chemical data:
  + Collection
  + Storage
  + Analysis
  + Manipulation
* This chemical data includes:
  + Chemical formulas
  + Chemical structures
  + Chemical properties
  + Chemical spectra
  + Biochemical and Biological activities
* The most common uses of chemoinformatics are:
  + Virtual screening
    - This involves the creation of large *in silico* virtual libraries of compounds, which are then submitted to a docking programming in order to identify the most active members
  + Quantitative Structure-Activity Relationship (QSAR)
    - This is used to predict the activity of compounds from their structures.

**Q2. Why is it required to study chemoinformatics? Describe.**

* Chemoinformatics plays a key role to maintain and access enormous amount of chemical data, produced b chemist, by using a proper database.
* The field of chemistry also needed a new technique for knowledge extraction from data to model complex relationships between the structure of the chemical compound and biological activity or the influence of reaction condition or chemical reactivity.
* Recent chemical developments for drug discovery are generating a lot of chemical data. (Information explosion)
* This has created a demand for effectively collect, organize, analyse and apply the chemical information in the process of modern drug discovery and development.
* Entanglement of chemistry and information management started in mid 1970s and the demand for chemoinformatics grew exponentially (Even though the term was coined much later in 1998.
* The modern drug discovery and development pipeline process, starts with disease selection, target identification, lead identification, lead optimization, pre-clinical trial testing, clinical trial testing, approval and circulation.
* This takes a lot of time to do *ex-silicon* hence the involvement of computers (Chemoinformatics) in the early stages speeds up and reduces cost of drug design.
* Currently, Target identification, Lead identification, Lead optimization is done *in-silico.*

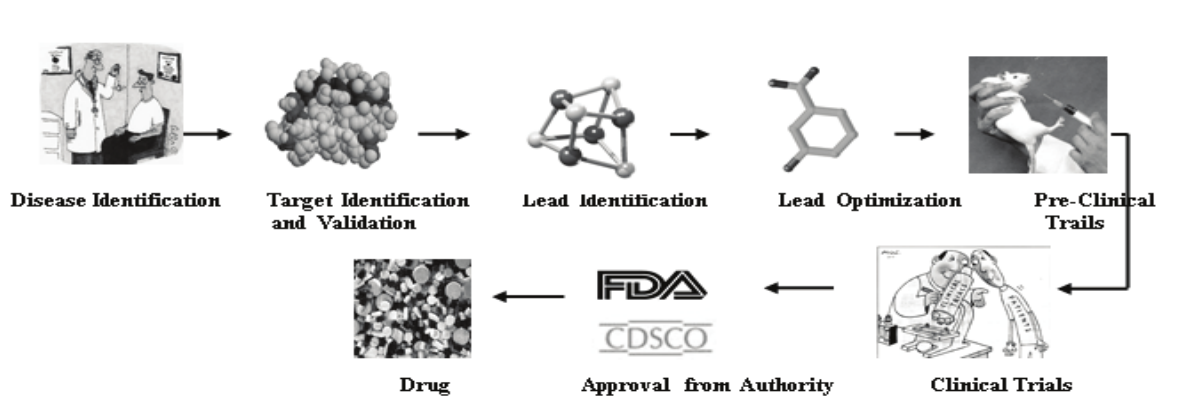


Fig: Modern Drug designing process

**Q3. Write a note on history of chemoinformatics**

* Chemoinformatics has been an active field in various guises since the 1970s and earlier, with activity in academic departments and commercial pharmaceutical research and development departments.
* The term chemoinformatics was defined in its application to drug discovery by Dr. F. K. Brown in 1998
* “Chemoinformatics is the mixing of those information resources to transform data into information and information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and optimization”.
* Since then both spellings “Cheminformatics” and “Chemoinformatics” has been used interchangeably.
* In 1946 King et al, publicised an article illustrating the use of IBM's business accounting machines in carrying out the construction of the rotational spectra of asymmetric rotors by the evaluation of mathematical equations for line position and line intensity. This is regarded as the first work involving computer technology in chemistry and the birth of Chemoinformatics.
* In 1957 IBM made the first transistor-based computers as well as FORTRAN, a high-level programming language which made computers a generally available commodity for scientists.
* 1951 marked the invention of the first database system developed by Chemical abstract service (CAS) of the American Chemical Society.
* The following years from 1950s to 1990s saw exponential growth in both the capability of computers as well as the data available about chemistry.
* By the end of 1990s advancements in chemoinformatics had led to the invention and use of Database Systems, Case expert Systems, CASD expert systems and 3D structure builders.
* This was the turning point for chemoinformatics, which then led to the cascade of events that led to the current state of chemoinformatics / Drug design.

**Q4. Comment on Chemoinformatics vs Cheminformatics**

* Cheminformatics and Chemoinformatics have both been used interchangeably in research papers.
* Cheminformatics is being used more frequently (2.5 times more) than chemoinformatics since 2009.

**Q5. Describe application of chemoinformatics in detail**

* Applications of Cheminformatics are:
  + Strong data generated through experiments or from molecular simulation Retrieval of chemical structure from chemical database
  + Prediction of physical, chemical and biological properties of chemical compounds
  + Elucidation of the structure of a compound based on spectroscopic data.
  + Structure, substructure, similarity and diversity searching from chemical database
  + High throughput screening (HTS) is the integration of technologies (lab automation, assay tech, micro plate based instrumentation
  + Docking – Interaction between two macromolecules
  + Drug Discovery
  + Molecular Science, Materials Science, Food Science, Atmospheric chemistry, Polymer chemistry, Textile chemistry, Combinatorial organic synthesis

**Q6. Elaborate on types of learning approach used in Chemoinformatics**